FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Chasing shadows: the interplay of privacy and (digital) identification – toward an identifiability-based framework for privacy impact assessment

DISSERTATION

submitted in partial fulfillment of the requirements for the degree of

Doctor of Social and Economic Sciences

by

Stefan Strauß
Registration Number 09855841

to the Faculty of Informatics
at the Vienna University of Technology

Advisor: Prof. Dr.Wolfgang Hofkirchner

The dissertation has been reviewed by:

(Prof. Dr. Ingrid Schneider) (PD Dr. Michael Nentwich)

Wien, 25.07.2017

(Stefan Strauß)

FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Chasing shadows: the interplay of privacy and (digital) identification – toward an identifiability-based framework for privacy impact assessment

DISSERTATION

zur Erlangung des akademischen Grades

Doktor der Sozial- und Wirtschaftswissenschaften

eingereicht von

Stefan Strauß
Matrikelnummer 09855841

An der
Fakultät für Informatik der Technischen Universität Wien

Betreuung: Prof. Dr.Wolfgang Hofkirchner

Diese Dissertation haben begutachtet:

(Prof. Dr. Ingrid Schneider) (PD Dr. Michael Nentwich)

Wien, 25.07.2017

(Stefan Strauß)

Erklärung zur Verfassung der Arbeit

Stefan Strauß
Weißgerberlände 50, 1030 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

_____ _____

(Ort, Datum) (Name Verfasser)

Ackknowledgements

I would like to thank Wolfgang Hofkirchner for supervising this thesis and his advice and support during my work. I would also like to express my gratidude to all my colleagues and friends at ITA for the vidid discussions and great working atmosphere.

Finally, special thanks goes to my family and friends and particularly to Angelika for her patience, motivation and all the great support. Last not least I woud like to thank our little daughter Mathilda, who enriches life and contributed a lot to the finalisation of this thesis.

Abstract

Digital technology deeply affects the natural interplay between privacy, personal identity and identification. Privacy implications of digital identification practices are apparent but privacy studies tend to neglect either technical or societal factors in this regard. Only few studies take multiple angles into account. To narrow this gap, this research applied an interdisciplinary approach, located in the field of technology assessment; informed by general systems theory as an analytical lens. The results reveal a privacy control dilemma of digital identification shaped by several interrelated socio-political, economic and technical factors. Existing tensions between privacy, security, surveillance and transparency aggravate. Increases in implicit and explicit forms of (digital) identification, partially overlapping with surveillance practices, reinforce this dilemma: further growth in digitally networked environments complicates the detection of privacy risks and the creation of appropriate safeguards. Thus a core problem of contemporary privacy protection is uncontrolled identifiability, aggravating information asymmetries and agency problems inherent to the control dilemma. Entailed is increasing demand for privacy by design (PbD) and privacy impact assessment (PIA), also stimulated by the new European data protection regulation. Easing the dilemma requires more transparency of privacy-affecting information processing and thus PIA. Based on a review of existing PIA approaches, a refined approach was developed. The proposed identifiability-based PIA framework contributes to improve the theoretical understanding of privacy impacts with practical relevance. Included is a typology of identifiable information with explicit consideration of technical identifiability. The typology enables a more systematic analysis of privacy-relevant information processes as integral part of PIA. This can also contribute to the development of more effective PbD implementations. A prototypical PIA process sketches a potential practical adoption of the identifiability-based approach, supportive as a guideline for PIA implementations in institutions. Progressing digital automation and semi-autonomous systems make a further expansion of identifiability likely and thus additional demand for effective PIA and PbD.

Kurzfassung

Digitale Technologien haben massive Auswirkungen auf das Verhältnis zwischen Privatsphäre, Identität von Personen und deren Identifizierung. Privatsphäre-Implikationen von digitalen Identifikationsmechanismen sind offensichtlich, in der Forschung werden dabei allerdings entweder gesellschaftliche oder technische Aspekte häufig vernachlässigt. Diese Arbeit versucht mit einem interdisziplinären Ansatz im Bereich der Technikfolgen-Abschätzung, gestützt durch die allgemeine Systemtheorie, diese Lücke zu verringern. Die Ergebnisse offenbaren ein Kontroll-Dilemma der Privatsphäre bezüglich digitaler Identifizierungspraktiken, beeinflusst von einer Reihe miteinander verwobener, sozio-politischer, ökonomischer und technischer Faktoren. Bestehende Spannungen zwischen Privatsphäre, Sicherheit, Überwachung und Transparenz nehmen weiter zu. Die damit verbundene Zunahme impliziter und expliziter Identifizierung, begünstigt durch digitale Technologien, teilweise überlappend mit Überwachungspraktiken, verschärft das Dilemma. Das Erkennen von Privatsphäre-Risiken und die Entwicklung geeigneter Schutzmaßnahmen werden dadurch erheblich erschwert. Ein Kernproblem des Schutzes der Privatsphäre ist daher die unkontrollierte Identifizierbarkeit von Individuen, wodurch Macht- bzw. Informationsasymmetrien weiter zunehmen. Dementsprechend wächst der Bedarf nach Privacy by Design (PbD) und Privacy Impact Assessment (PIA), wie auch in der neuen Europäischen Datenschutzgrundverordnung vorgesehen. Um das Dilemma zu entschärfen ist mehr Transparenz von (datenschutzrelevanten) Informationsprozessen erforderlich und daher PIA. Anhand bestehender Ansätze wird ein neuer konzeptueller Rahmen für PIA mit Fokus auf Identifizierbarkeit vorgeschlagen. Das Modell leistet einen Beitrag zum theoretischen Verständnis der Folgen digitaler Technologien für die Privatsphäre mit Relevanz für die Praxis. Der Entwurf einer Typologie verschiedener Arten von Identitätsinformation berücksichtigt explizit auch technische Identifizierbarkeit. Das kann zu einer systematischeren Analyse von personenbezogenen Informationsprozessen als Kernbestandteil von PIA beitragen sowie zur Entwicklung effektiverer PbD-Ansätze. Anhand eines prototypischen Prozesses wird die mögliche praktische Anwendung des PIA-Modells in Institutionen kurz veranschaulicht. Mit Zunahme digitaler Automatisierung und teil-autonomer Systeme ist mit steigender Identifizierbarkeit und daher weiterem Bedarf nach PIA und PbD zu rechnen.

Contents

List of figures

List of tables

CHAPTER 1

Introduction

"We are but whirlpools in a river of ever-flowing water. We are not stuff that abides, but patterns that perpetuate themselves. A pattern is a message, and may be transmitted as a message". (Norbert Wiener (1954: 96).

Privacy is a heavily threatened socio-technical concept and among the "endangered species" of the information society. Information and communication technologies (ICTs) essentially transformed the organisation and functioning of society which, besides their numerous benefits, enabled novel forms of privacy intrusion. As, e.g., Wright and De Hert (2012a: 3) put straight: "If privacy is a cornerstone of democracy, then democracy is in trouble." This critical appraisal addressed the numerous threats to privacy resulting from technology and surveillance practices; one year before whistleblower Edward Snowden informed the world about the yet unknown extent of mass surveillance. At latest since the Snowden revelations in 2013, there is hard evidence for surveillance programs at global level exploiting personal information writ large (Greenwald 2014). Since then, the US National Security Agency (NSA) became more or less synonymous to surveillance in the public discourse. The case as such raises many serious questions which are yet only rudimentary explored and addressed, e.g., about the legitimacy of surveillance practices and privacy intrusion for law enforcement, their effectiveness and threats to human rights, the accountability of security authorities and national intelligence etc. Irrespective of its explosive political impact, the Snowden case teaches the important lesson that contemporary technology provides numerous ways to intrude into privacy, which evidently serve various forms of surveillance. Apparently, though, privacy threats are not limited to the NSA or other security agencies mentioned in the Snowden files. In fact, privacy-intrusive practices exploiting digital technology have been a critical issue long before the Snowden files went public. Nevertheless, the revelations intensified the demand to reinforce privacy protection as well as transparency of information processing. A certain "Snowden effect" is observable as privacy awareness seems to have increased since then: for instance, privacy issues gain importance in science and research: a simple search in the web of science[1] on privacy (conducted on June 28 2017) leads to more than 22,500 results including the years between 2000 until 2017. Nearly the half of these results (more than 11,100) concerns the years between 2013 and 2017. But also among businesses as well as in the public, privacy issues gain in importance: companies started to invest more in protecting their data by e.g., fostering encryption of online services (cf. Kuchler 2014; Finley 2014) and several studies indicate increasing public concerns about surveillance (cf. Madden 2014; Lyon 2014; Kerr 2015). The on-going data protection reform of the

[1] webofknowledge.com

European Union is another indicator that society tries to cope with privacy issues. However, on the other hand, privacy and security experts observe a certain fade out of the Snowden effect as surveillance practices continue and serious privacy threats are yet unsolved (cf. Weinger 2016). While privacy is not dead it requires considerable action to become revitalised. Updating legal frameworks is highly important but the effectiveness of regulation strongly depends from their practicability in socio-technical practices. Irrespective of the Snowden case, protecting privacy is increasingly challenging in contemporary society. The processing of personal information is often opaque and it is often unclear, to what extent privacy is really affected. To strengthen privacy protection thus essentially requires coming toward a deeper understanding of privacy impacts and the very mechanisms inherent to socio-technical practices enabling privacy intrusive activities. Today, digital information flows can include not merely one but various applications, often appearing as conglomerate of multiple interwoven technologies. Given the complexity of digital technology, there is a certain risk to get lost in technological conflation when analysing privacy impacts. Therefore, it can be challenging to grasp the extent to which an application bears privacy risks. As a consequence, it is equally difficult to implement effective protection mechanisms. In total, privacy suffers from its generally abstract conceptualisation and a broad range of socio-technical threats. A basic motivation of this research is thus to shed light on the theoretical understanding of privacy impacts. This can facilitate impact assessment as well as on the longer run the implementation of privacy safeguards in the realm of privacy by design (PbD). Both issues are of utmost importance for the General Data Protection Regulation (GDPR) of the European Union, becoming effective in May 2018. This new regulation transforms the privacy regime in Europe and strengthens the general role of privacy impact assessment (PIA). In accordance with the new regulation, there are new approaches necessary to implement PIA and reinforce the level of privacy protection. This requires an analysis of the core functions of privacy and the main issues challenging their effectiveness. This research ties in here by putting emphasis on the interplay of privacy and (digital) identification, because privacy and identity are essentially linked, sharing a naturally close relationship. The extent to which this relationship is affected by ICTs and the related socio-technical practices is explored to grasp the emergence of privacy impacts as well as approaches to improve corresponding safeguards. Based on the results, a novel framework for PIA is proposed to contribute to the theoretical understanding and practical implementation of privacy protection.

1.1 Digital identification in the network society

In the early days of the Internet and the World Wide Web, a popular cartoon[2] from Peter Steiner published in 1993, claimed that “on the Internet, nobody knows you’re a dog”. Today, this cartoon may be reworded to “on the Internet, everybody knows you, your dog as well as why, when and how you got him”. In fact, online anonymity, as the cartoon implicitly hints at, is far more complicated than it used to be during the 1990s. From a wider view, the popularity Steiner’s cartoon received is a good example for the powerful

[2] https://en.wikipedia.org/wiki/On_the_Internet,_nobody_knows_you%27re_a_dog

modalities of the Internet and ICTs to spread information across multiple contexts, enriched with a vigorous self-dynamic. Today we would say that Steiner's cartoon got "viral" so to speak. While spreading a cartoon online differs significantly from distributing personal information, the very mechanisms are the same, resulting from the dynamics of ICTs: information is easily reproducible, more or less unbound from spatial and temporal limits. These dynamics accelerate not least as today's technologies are – if not already– nearly ubiquitous. This has consequences for the ways our identities are represented and processed. In line with Wiener's notion of humans as self-perpetuating patterns (as outlined in the opening citation), our identities may be perceived as unique patterns representable by information. This is not to be misunderstood as reductionist approach (as partially suggested by classical cybernetics assuming analogies between human beings and machines). Apparently, identity is more than a unique pattern of (computable) information. Identity is a multifaceted phenomenon with various meanings and functions in society, hardly explainable by machine analogies or similar mechanistic views. But irrespective of its multiple functions, the peculiarities of an identity are representable by unique pieces of information enabling to recognise that one entity differs from others. Against the background of an increasingly networked information society, co-shaped by technology, the notion of identity as a pattern represented by information is of special relevance. ICTs created new ways of gathering and processing information about individuals serving a variety of social, economic and political purposes. Using ICTs may generate various forms of information suitable to identify a particular person.

ICTs are not merely technical tools but integral parts of society serving various societal functions; they represent socio-technical systems which shape society and vice versa. With their rapid progress and widespread diffusion, ICTs deeply pervade a broad array of societal domains and every-day-life contexts. This pervasion entails what Moor (1998: 15) called "informational enrichment" (or informatisation) of societal activities as well as their conceptions. This means that ICTs enabled new options to digitally represent and process information about societal entities such as organisations, domains, objects, or people and the activities involved. Consequently, ICT usage also affects the representation of identities, which today can be easily embedded in networking structures. This can reinforce identifiability as contemporary technology offers various ways of direct and indirect identification. In this regard, increasing identifiability is an important side-effect of the (digital) information age. In the "age of identification", as Hildebrandt (2008: 56) once stated, our personal identities are embedded in and exposed to a magnitude of digital environments. Technology alters the way identities are represented, organised and handled by individuals as well as groups and institutions (Whitley et al. 2014). These developments are phenomena of the "network society" as described by Manuel Castells (2000: 5) as a "social structure characteristic of the Information Age", triggered by globally networked ICTs. The structural setting of society changed with technology, adding a specific informational layer to the social structure. As "information processing is at the source of life, and of social action, every domain of our eco-social system is thereby transformed" (Castells 2000: 10). Castells early realized the deep structural shifts of society resulting from informatisation. Today networking structures occur within and between offline and online environments enabled and reinforced by ICTs. Hence, these shifts did not merely

affect how information is structured and processed in digital environments. Society has more and more entered a stage of convergence between analogue and digital environments with information as a driving force in our ecosystem (cf. Floridi 2010; Hofkirchner 2010/2013). This ongoing socio-technical transformation proceeds quickly and is hard to grasp; or in other words: "Our technological tree has been growing much more widely, rapidly and chaotically than its conceptual, ethical and cultural roots" (Floridi 2010: 5). Floridi uses the term "infosphere" to describe this transformation embracing the (ontologically) powerful nature of information. The infosphere "denotes the whole informational environment constituted by all informational entities (thus including informational agents as well), their properties, interactions, processes and mutual relations" (Floridi 2010: 6). This infosphere constantly alters with ICTs and describes a highly dynamic environment comprising analogue and digital settings, linking online as well as offline domains.[3] Hence, socio-technical change reaches a new quality including an incremental shift of the boundaries between society and technology, physical (or analogue) and digital environments. As Verbeek (2011: 30ff.) pointed out, humans are "profoundly technologically mediated beings" and "technology is part of the human condition". Technologies basically represent socio-technical systems that affect the human condition and vice versa. ICTs highlight and enforce this interplay: their rapid diffusion and usage entails increasing connectivity and permanent availability of always-on computer devices, employed in various domains. There is a dramatic increase in digitally networked environments observable, reinforcing further growth in the amount of digital information. While in the year 2001, the number of global internet users was about 500 million, today there are about 3.5 billion internet users worldwide (ITU 2016). Network providers predict the amount of global internet traffic will soon exceed one zettabyte per year (10^{21} byte and about one trillion gigabyte). Especially mobile computing is on the rise. In 2016, about 50 per cent of global internet traffic results from wireless and mobile devices. By 2021, over 60 per cent are expected. The number of networked devices is assumed to be then three times higher than the world population, i.e., more than 26 billion devices (Cisco 2016). Against this background, visions of a globally networked information society including notions of pervasive or ubiquitous computing, ambient intelligence etc. (cf. Weiser 1991; ITU 2005) take more concrete shape with developments in the realm of "smart" technologies, the Internet of Things and the like. The general boost in digital networks is accompanied by a further expansion of digital information processing. With these developments, individuals and thus their identities are increasingly interconnected and represented by their digital information, prominently highlighted by, but not limited to Web 2.0 and social media platforms. Digital identities are already involved in a broad variety of interactions (e.g., information exchange, communication, collaboration, sharing and creating content); amongst others fulfilling (and stimulating) the societal need to communicate and exchange with others. But entailed are also further growth in the amount of personal information, personalisation and uncontrolled information disclosure. Trends in the realm of big data (Mayer-Schönberger/Cukier 2013; Strauß 2015a), aiming at

[3] It is conceptually broader than cyberspace, which primarily addresses online environments. The Merriam Webster Dictionary defines cyberspace as "the online world of computer networks and especially the Internet" http://www.merriam-webster.com/dictionary/cyberspace

exploiting data from everyday life for novel services further amplify the processing of digital information in multiple application contexts. This has consequences for the representation and processing of individual identities as well.

1.2 Privacy vs. (digital) identification?

Altogether, these developments amplify digital networking structures which deeply affect society in general, including social practices, norms and values. Among the variety of societal impacts, serious threats and challenges concern the notion and protection of privacy. Various tensions result from the need "to fit the technological and socio-political evolutions" which generate "new threats for the individuals' capacity for 'self-development' of their personality" (Rouvroy/Poullet 2009: 55). Hence, threats to privacy in the end also affect identity building (Hildebrandt 2006). In the light of the proceeding socio-technical transformations, the need for "a radical re-interpretation of informational privacy, one that takes into account the essentially informational structure of human beings and of their operations as social agents" (Floridi 2013: 243) is more topical than ever. In other words, there is a certain demand to re-conceptualize privacy with respect to the informational nature of humans and the representation of their (digital) identities. On the one hand, because threats to privacy can threaten identity-building of the individual concerned as well. On the other hand, because ICTs also transformed the way identities are represented and processed, i.e., identification. In this regard, ICTs have impact on the interplay of identification and privacy.

The dynamics of ICTs further intensify the challenge to effectively protect privacy and to adapt existing practices to the changed requirements. Digital information can flow across many different contexts. As a consequence, the boundaries between personal and non-personal information, private and public spheres can be strained. In between these boundaries, in the point of intersection, identity becomes a particular (informational) matter. An expansion of digital information processing affects the nexus between privacy and identity in manifold ways. Given the peculiarities of digital information, it can create a sort of permanent link to the identity of an individual person. This has effects on the privacy of this person as well; because "one's informational sphere and one's personal identity are co-referential, or two sides of the same coin" (Floridi 2013: 245). Therefore, protecting privacy includes the protection of personal identity. Conversely, identification can be privacy-intrusive. Hence, to some extent, privacy and identification can be mutually exclusive. For instance, when identification is related to security and surveillance practices. There is a certain tension between privacy and security which mirrors in the discourse on digital identity (cf. Halperin/Backhouse 2008; Strauß 2011). This tension can challenge the effectiveness of privacy protection as security considerations often dominate the implementation of identification processes. This is observable in the broad scope of digital identification processes being directly and indirectly involved in socio-technical systems and practices. Many technologies and online applications process different forms of personal (or identity) information. To ease the handling of this information, the field of

identity management (IDM) emerged in research and development. The increasing relevance of IDM indicates a societal demand to deal with issues of identity in the information society (cf. Halperin/Backhouse 2008; Rannenberg et al. 2009; Aichholzer/Strauß 2010a; Strauß 2011; Kubicek/Noack 2010a/2010b; Whitley et al. 2014). The basic aim of IDM is to unify identification and authentication processes. Different technological artefacts such as digital devices (e.g., online profiles, electronic ID cards, smartcards, smartphones) can be used a carrier devices for managing individual identities. These artefacts can be seen as a sort of "strong tie" between analogue and digital environments as they can support identification mechanisms in both: online and offline contexts. There are numerous application contexts where digital identification processes are involved ranging from transactions in e-government and e-commerce, multiple online services, as well as social media platforms. The implementation IDM concepts or of digital identification systems primarily aims at fostering efficiency and security of online services. In contrast to that, privacy protection plays a rather marginal or implicit role. Most IDM approaches provide unique identification without any features of anonymous and pseudonymous usage (Strauß 2011). This hampers the applicability of IDM for privacy protection. Besides explicit forms of identification, there are implicit forms of identification as well. Hence, identification is not merely a formal process (e.g., between citizens and the state, customers and businesses etc.) but also occurs in the form of profiling activities, where information is gathered to create extensive profiles of the individuals concerned. Moreover, explicit and implicit identification may overlap. In total, socio-technical identification practices not merely affect the handling of personal information but the values and norms it rests upon. In particular privacy and the individuals' ability to control their information, i.e., informational self-determination, suffer from uncontrolled identification practices. The processing of identifiable information as well as the option to be identified (identifiability) can lead to an imbalance of control and information asymmetries at the cost of the individual concerned. Consequently, individuals have very limited options to effectively protect their privacy as well as to take action in case of privacy abuse.

To tackle these problems which are likely to aggravate with further pervasion of ICTs, there is an increasing demand for novel approaches to early detect and stem threats to privacy. This demand refers to two crucial concepts: privacy impact assessment (PIA) and privacy by design (Pbd). For many years, privacy advocates argue for making both concepts mandatory in order to raise the general level of protection (cf. Wright/De Hert 2012b; EGE 2014). Finally, the new EU data protection framework fosters both concepts and foresees obligatory PIA under certain conditions. PbD aims at incorporating privacy protection already in the development process of technologies (cf. Cavoukian 2012a; Klitou 2014). PIA is an instrument to explore privacy risks and the extent to which technologies are in accordance with privacy and data protection (cf. Clarke 2009; Wright/De Hert 2012b; EGE 2014). Both approaches are interrelated and complement each other: conducting a PIA is vital for the implementation of privacy safeguards and on the longer run, it can facilitate the development of privacy-friendly technology. Conversely, when technologies have integrated PbD features, this can ease privacy impact assessment. As the creation of privacy safeguards requires knowledge about the amount and processing

of personal information, PIA is a precondition of PbD. PIA is not a new concept, early approaches date back to the 1990s. However, ICTs and the entailed socio-technical transformations request for a refinement of PIA conceptualisations. To come to enhanced approaches, it is crucial to gain deeper insights into the socio-technical transformation of privacy and how it is affected by technology. While there is broad scientific and societal consensus about privacy being threatened, there is a lack of common conceptual understanding of the way socio-technical systems enable privacy intrusions. As will be shown, (digital) identification and the processing of identity information play an important role in this regard. However, this particular role yet gained only little attention for privacy impact assessment. This research contributes to close this gap and sheds light on the interplay between privacy and (digital) identification to refine approaches of PIA. The results can improve the theoretical understanding of privacy impacts as well as support institutions in the practical implementation of PIA processes.

CHAPTER 2

Problem description and research design

2.1 Aims and methodological approach

There are four basic aims of this research: to explore the interplay of privacy, identity and identification; to shed light on the emergence and peculiarities of digital identification; and analyse its impacts on privacy, including its overlaps with surveillance practices. Finally, based on the results of this analysis, the functions and scope of PIA will be examined to develop a refined approach. This is particularly relevant in the light of the new European data protection regulation, the entailed transformations of the privacy regime and of socio-technical practices. While the regulation stimulates PIA on a general level, it does not provide guidance on its implementation. There is thus demand for conceptual approaches to ease the practical implementation of PIA in institutions. The proposed framework can be supportive in this regard although it has no particular focus on legal issues. It thus may be seen as an ethical PIA approach which ideally contributes to legally motivated assessments as well.

The starting point of this research concerns the interplay of privacy, identity and identification. Although this interplay seems apparent at first glance, it gained little conceptual attention for the assessment of privacy impacts. Making this relationship more explicit and revealing its dynamics in the realm of ICTs is thus considered important to allow for more effective privacy protection. Besides the various challenges resulting from technology and corresponding usage practices, also the complexity of privacy as such complicates its protection. Privacy has an abstract character and multidimensional scope on the one hand, and there are narrow conceptualisations on the other hand. Traditional notions of privacy frame it as “right to be let alone” (Warren/Brandeis 1890). This notion is still popular and mirrors in public discourse as well as in technology development. Partially similar is the view on privacy as data protection which is dominant in the European privacy regime. Without doubt, these framings are relevant, particularly in a legal sense. However, as a crucial societal value, privacy protection is relevant beyond legal obligations as well. A narrow conceptual framing of privacy hampers the detection of impacts and creation of effective safeguards. Therefore, more differentiated views on privacy and on personal information are required. There are some promising conceptualisations in this regard such as Nissenbaum’s (2010) theory of contextual integrity, which underlines the public and private value of privacy. Contextual integrity suggests that the contexts of personal information processing, i.e., informational norms determine the extent to which privacy is affected. In the view of Nissenbaum (2010), a breach of informational norms results in a breach of privacy. However, in the light of interconnected technologies and applications, the examination of norms as well as the

detection of breaches can be demanding. It is thus important to gain a deeper conceptual understanding of the emergence of privacy impact to come toward more effective approaches of privacy protection. A general problem is that technology and usage practices challenge to determine what needs be protected. While personal information is obviously essential, it can be challenging to grasp the extent to which personal information is factually processed in a privacy-affecting manner. Conceptual approaches such as the seven types of privacy suggested by Finn et al. (2013) to grasp the multiple dimensions of privacy, or Solove's (2006) taxonomy of privacy-affecting activities can be useful in this regard. However, also these more recent approaches are rather diverse and difficult to integrate in impact assessments. To ease this problem I argue for a stronger focus on issues of identifiability and identification. The basic assumption is that a privacy impact is generally determined by the processing of information referring or relating to an individual, i.e., some form of identification. Identification is thus assumed to have a connecting role that enables interlinkage of different socio-technical systems. This work is an attempt to shed light on this connecting role from a wider, systemic perspective. The emergence, function and societal impacts of (digital) identification in relation to the protection of privacy are of main interest.

There is various research about the general privacy impacts of new technologies and surveillance from sociological, political, economic, legal, ethical and technological perspective (for example Haggerty/Ericsson 2000; Lyon 2003; Clarke 2006; Solove 2006; Hildebrandt 2006; Nissenbaum 2010; Bennett 2011; Wright/De Hert 2012a; Finn et al. 2013; Lyon 2014; Wright et al. 2015). However, fewer studies explore socio-technical privacy impacts from different, interdisciplinary angles. Studies of privacy issues in the social sciences often neglect the peculiarities of technology, and studies in the realm of engineering or computer sciences focus on technical approaches while neglecting relevant social sciences perspectives on privacy. Hence, there is a certain research gap, particularly as regards the interplay between identity and privacy from a systemic perspective. This research contributes to narrow this gap by applying an interdisciplinary approach. It is located in the research domains of technology assessment as well as science and technology studies (STS), where multiple aspects, including socio-political, technological and ethical aspects are taken into account. The research is guided by the following questions:

- What are the basic concepts and functions of identity, identification and privacy, and how are they interrelated?
- What are the relevant functions, drivers and dynamics of the emergence of digital identification?
- To what extent is digital identification related to the security and surveillance practices and how does this affect the protection of privacy and informational self-determination?
- What are the prospects and perils of privacy controls (privacy by design) and privacy impact assessment? What are relevant conceptual approaches to overcome existing barriers?

To find answers to these questions, this research is based on a combination of different methods: the analysis of relevant literature serves as basis to gain a detailed overview on the state-of-the-art whereas an interdisciplinary approach incorporates conceptual, empirical and technical investigations from different research perspectives (including privacy, security and surveillance studies, technology assessment, identity management, as well as system and information science). The analysed literature involves research books and papers, policy documents, legal frameworks, official standards, surveys, technical reports and specifications, as well as media reports. A system-theoretical approach serves as research heuristic to explore the socio-technical transformations related to ICTs (as described in the next Section).

The study is mainly theoretical but combined with empirical investigations from the mentioned literature. Furthermore, the data material includes empirical findings from research projects I was involved in: particularly about the innovation of governmental identity management systems (Aichholzer/Strauß 2010a/2010b; Strauß/Aichholzer 2010; Strauß 2011), societal impacts and privacy issues of cloud computing and social network sites (Strauß/Nentwich 2013; Leimbach et al. 2014) as well as the European research project SurPRISE[4] that dealt with the interplay between privacy, security and surveillance technologies (Strauß 2015b; Strauß 2017a). A central part of the SurPRISE project was a large-scale citizen participation process with 1780 participants in total, conducted in nine European countries based on an interactive survey. As a work package leader, I was inter alia involved in designing the analytical framework, conducting a security policy analysis, and concerned with the analysis of the empirical results of the Austrian participation process as well as with the synthesis of the overall results of the large-scale participation (for more details see Section 5.2). The methods applied in the mentioned projects include: policy analysis, surveys, qualitative interviews and workshops with experts from different fields (privacy and security researchers and practitioners, policy makers, technology developers, legal experts and individual citizens), practical software tests of identity management systems and of social media platforms. Findings from each of the mentioned projects flew into this research, updated and enriched with additional theoretical and empirical investigations.

2.2 Systems theory in a nutshell

The dynamics of ICT-induced socio-technical transformation, which affect privacy and identification, are explored through the analytical lens of systems theory, serving as a research heuristic. The term system here primarily means a socio-technical system, i.e., a technology or a set of technologies and its interplay with societal entities, structures,

[4] SurPRISE is the acronym for "Surveillance, Privacy and Security: A large scale participatory assessment of criteria and factors determining acceptability and acceptance of security technologies in Europe". The project received funding from the EU's seventh framework programme for research, technological development and demonstration under grant agreement no. 285492.

functions and practices.[5] The main focus is on ICTs, here understood as information processing metasystem.

A systems approach enables to look beyond the limits of mechanistic views to analyse theoretical problems as well as issues related to modern technology (Bertalanffy 1969; Hofkirchner 2005). It offers a general frame of inquiry to put emphasis on the interplay of interrelated components and their dynamics (Laszlo/Krippner 1998). A general asset of a system-theoretical approach is thus its provision of a meta-perspective through abstraction which enables an analytical view from multiple angles. This allows the cross-disciplinary investigation of socio-technical phenomena (such as ICTs pervading society) which are accompanied by high (and increasing) complexity. Systems theory provides a methodological approach to cope with this complexity and thus allows grasping the bigger picture of socio-technical change (Hofkirchner 2005; Hofkirchner 2013).

The scientific paradigm of systems thinking inherent to systems theory emerged during the 1950s. General systems theory originates from biologist Ludwig von Bertalanffy who outlined his approach in the British journal for the philosophy of science (Bertalanffy 1950). Bertalanffy (1972: 424) described general system theory as "a model of certain general aspects of reality. But it is also a way of seeing things which were previously overlooked or bypassed, and in this sense is a methodological maxim. And like every scientific theory of broader compass, it is connected with, and tries to give its answer to perennial problems of philosophy."

Systems thinking entered several disciplines such as mathematics, cybernetics, organization and management theory, biology, sociology, philosophy, psychology, engineering, system science, up to innovation and transition management (cf. Wiener 1954; Bertalanffy 1969; Ackoff 1971; Varela et al. 1974; Giddens 1984/1997; Parsons 1991; Luhmann 1991; Laszlo/Krippner 1998; Kneer/Nassehi 2000; Geels 2004; Fuchs/Hofkirchner 2005; Hofkirchner 2005; Hofkirchner/Schafranek 2011; Hofkirchner 2013). Although each of these domains deals with different types and categories of systems (e.g., biological/living; ecological; social; technical; cognitive; physical; chemical; informational; open; closed; autopoietic; formal; adaptive etc.) there are some general features of a system irrespective of the domain. In all living systems, processes of autopoiesis occur, which means that these systems can reorganize and reproduce themselves (Varela et al. 1974). Important in each case is the dynamics of a system which refers to self-organisation, i.e., transformation processes emerging within the system altering its structure and form of organization (Hofkirchner 2013).

The roots of systems theory can be traced back to Aristotle's statement, "the whole is more than the sum of its parts" which provides a basic system characteristic. It points out that a system is not to be reduced to a set of quantitative elements as not merely the different components, but also the relations and interactions between them affect the properties and dynamics of a system. Hence, the constitution of a system is not merely explainable by an isolated view on its elements (Bertalanffy 1969; Bertalanffy 1972). This general understanding is the basic idea of systems theory with the premise that "in order to

[5] The author is aware that there are differences between open vs. closed and isolated systems. However, a discussion about these issues is less relevant in the context of this research. As a socio-technical system is generally co-referential with its environment it is here generally assumed to be an open system.

understand an organized whole we must know both the parts and the relations between them" (Bertalanffy 1972: 411). Thus systems theory is not limited to questions regarding the quantity of system elements but also how they are interrelated and what the characteristics of these interrelations are. Its offers a degree of generality or abstraction that enables to comprehend "general characteristics partaken by a large class of entities conventionally treated in different disciplines" (ibid: 416). In this regard, systems theory has an inter- and transdisciplinary stance. A system represents "a model of general nature, that is, a conceptual analog of certain rather universal traits of observed entities" and can be defined "as a set of elements standing in interrelation among themselves and with the environment" (ibid: 416f.). In the broadest sense, a system is "a complex of interacting components together with the relationships among them that permit the identification of a boundary-maintaining entity or process" (Laszlo/Krippner 1998: 48). Figure 1 sketches the basic characteristics of a system.

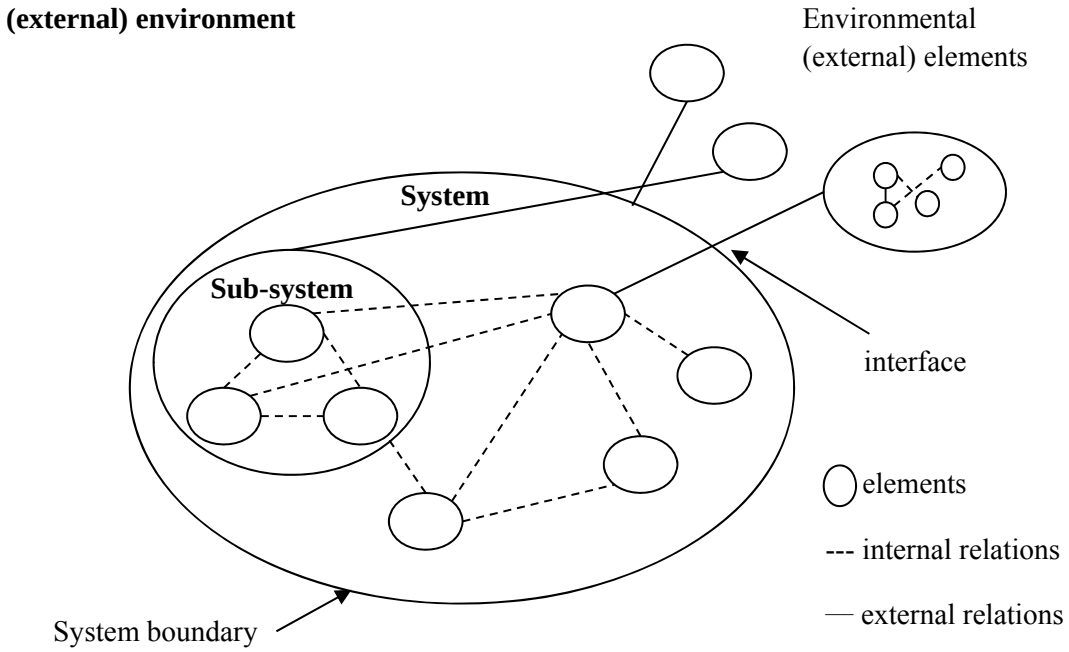


Figure 1: Basic system characteristics (own representation, inspired by Hofkirchner 2013)

As illustrated, a system consists of different elements, i.e., the items that are parts of the system. These elements have relations which are the connections among the systems' elements that enable its dynamics. Elements with relatively stable relations may together be seen as a sub-system of a larger system.[6] The boundary of a system is the edge that makes the elements and relationships of a system distinguishable from the external environment outside the system. A system is embedded in an environment which implies that there are external systems, components etc. outside its boundaries. External relations can occur, e.g., through interfaces. Finally, an interaction of a system with others can be

[6] For example, an organisation may be seen as a sub-system of an enterprise. The departments of this organization may be seen as sub-systems of this organization, which is determined and affected by the individuals working in these departments etc. A simple technological example is a web browser, which may be seen as a (software) sub-system of a computer system.

described as a process with an input and an output, meaning that things flow into the system (e.g., matter, energy, goods, information etc.) as well as out of it into its environment (e.g., into another system or its components). Input and output can be accompanied by feedback loops, meaning that the processes triggered by a system or its elements may loop back. Or in other words: a system is dynamically interrelated with its environment (Bertalanffy 1969; Bertalanffy 1972; Laszlo/Krippner 1998; Hofkirchner 2013). Hence, a system is mostly nothing static or mechanical but dynamic by nature and has emergent properties. Given these dynamics, a system and its properties can transform as it features a transformative capacity, i.e., it has inherent transformations due to its relations and interactions. This transformative capacity is closely related to processes of self-organization, which affect "the way evolutionary systems come into existence or change their structure, state or behaviour and the way they maintain themselves (their structure, state or behaviour) (Hofkirchner 2013: 115). A system thus has no strictly determined, immutable order or organization but alters and transforms. These transformation patterns may have effects on the system itself as well as on its environment. For example, the introduction of a new technology affects existing user practices and the way users interact with the technology may have further societal impact, which may affect the technology, e.g., by adaptations in design.

Considering its dynamics, a system can be seen as an evolutionary system, where relations between elements emerge which enable interactions that provide synergetic effects. This can lead to a self-reinforcing dynamic as certain relations begin to dominate their interaction (Hofkirchner 2013: 105). This dynamic process has thus two basic properties: emergence and dominance, i.e., elements and interactions between these elements emerge which then have effects on the state of the system. As the interactions and processes within a system are not strictly deterministic but dynamic, relations can emerge and also dominate an interaction process that may cause multiple effects. The organisational structure itself may be dynamic and thus each element can be subject to feedback loops where emergence and dominance alternate bottom-up as well as top-down (Fuchs/Hofkirchner 2005; Hofkirchner 2013). Similar dynamics exist in societal structures which emerge from the actions of individual actors (societal agents) but also affect them at the same time. Put shortly, in social systems, the individuals represent interacting agents at micro level, which generate social, cultural, economic, political structures, values, norms, organizations, institutions, technologies, processes, practices etc., which then emerge at the macro level. These generated items can be subsumed under the term socio-technical regime (cf. Kemp et al. 2001; Smith et al. 2005). The interactions between elements at the macro level can loop back to the micro level (Fuchs/Hofkirchner 2005). The figure below outlines this interplay:

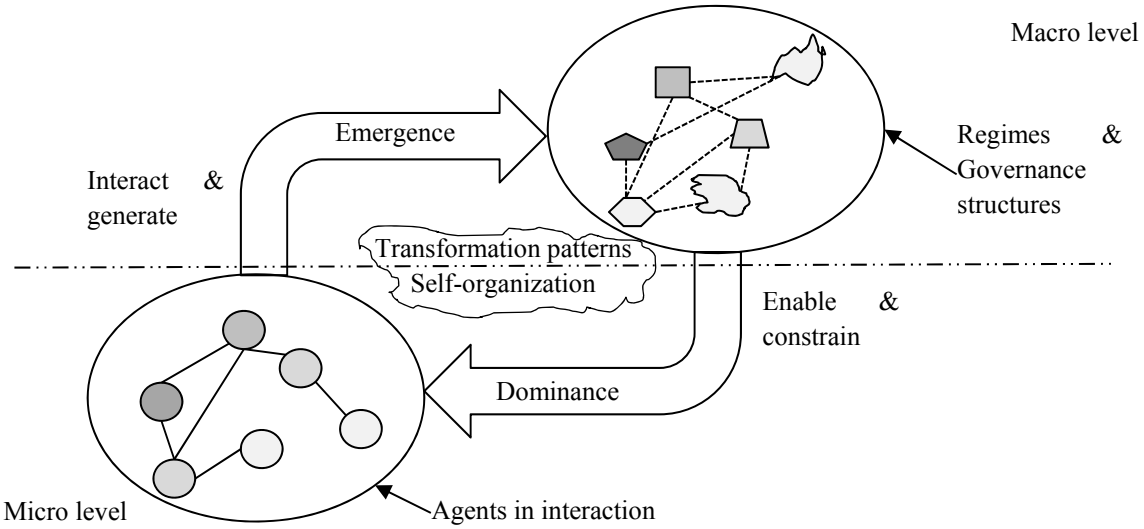


Figure 2: Model of systemic interplay between micro- and macro level (own representation, adapted from Hofkirchner 2013).

This setting partially refers to Gidden's (1984/1997) structuration theory which assumes a dialectical interplay of agency and structure: social structures emerge from as well as result in social actions. Societal structures entail enabling and constraining effects which may have some influence on individual action, though not in a direct, causal sense.[7] Or in other words: there are governance structures resulting from regimes, which enable and constrain societal functions, policies, processes, practices etc. Similar dynamics can be found in socio-technical systems as the creation and use of a technology e.g., involve organizations, structures, policies, formal and informal norms, individual practices etc. which influence each other.

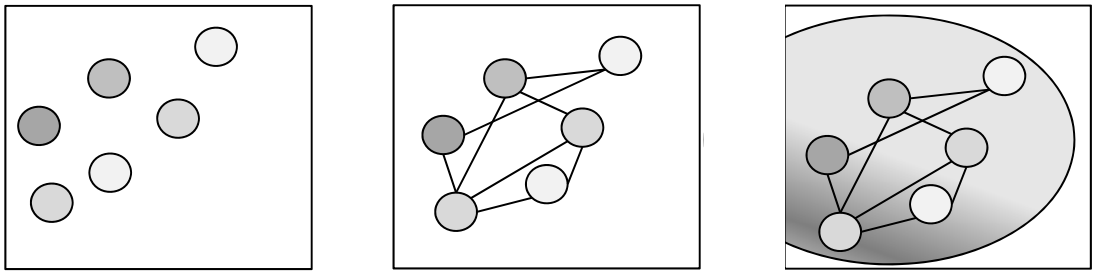
2.2.1 Metasystem transition

Self-organization processes and transformation patterns of an evolutionary system can be generally grasped by the concept of metasystem transition – MST (Hofkirchner 2013). Physicist and cyberneticist Valentin Turchin (1977: 98) described MST as the "quantum of development", i.e., a common feature of development processes where incremental progress takes place resulting in more complex forms of organization. It "creates a higher level of organization, the meta-level in relation to the level of subsystems being integrated" (Turchin 1977: 56). A simple example is the following: "When a human being applies tool B to objects of a certain class A, this tool, together with objects A, forms a metasystem, in relation to subsystems A. (…) Thus, the appearance of a tool for working on certain objects that had not previously been worked on is a metasystem transition within the production system" (Turchin 1977: 93).

Turchin's original approach is not without controversy because of its assumed higher order control system which gives the misleading impression that a centralized metasystem

[7] For example, the organization of a company has formal and informal rules and structures which have some impact on the employees behaviour and actions.

would emerge with the ability to control other subsystems. However, this notion of control is not to be misinterpreted in that way but rather meant in the sense of a more complex form of organization where different (formerly less structured) elements build an integrative, structural or functional entity. Hence, MST addresses the emergence of new systemic properties with a more complex form of organization inherent to the dynamics of evolutionary systems (Hofkirchner 2013). It describes a process of three main stages (individuation, interaction, integration), where a higher level of organization, perceivable as a metasystem, occurs through increasing interactions between homogenous subsystems (Turchin 1977; Heylighen/Campbell 1995; Fuchs/Hofkirchner 2005; Hofkirchner 2013). Figure 3 below illustrates the different transition stages of MST:



(1) Individual phase (2) Interaction phase (3) Integration phase

Figure 3: Phases of a metasystem transition (adapted from Hofkirchner 2013: 116ff.)

(1) The individual phase is characterized by a number of entities (or isolated, unrelated elements) that have no connections yet. In other words: there are different systems that are rather isolated from each other. Internal information processes within these components are dominating this stage. (2) In the interaction phase, occasional relations between the different entities emerge as they interact with each other. Hence, individual systems are increasingly connected and are in interaction. These interactions are not stable and can change or diminish. Processes in this phase are less affected by path dependencies and reversible. (3) In the third phase of integration, the interactions among the entities expand further, become more stable and the entities (subsystems) become elements of the emerging system. Specialisation occurs as a subsystem becomes adapted to the new structure which alters its functionality. The changes taking place in this integration phase are hardly reversible as the established structure and organization of the system are stable and thus do not easily change. Though, this does not mean that the relations between the elements are not changeable at all. To the contrary, the system as a whole and its interactions remains highly dynamic (Hofkirchner 2013). Therefore, the integration phase is not to be misunderstood as final development but rather as the occurrence of significant changes in quality and organization of a system. Furthermore, these transition phases can recur at different organizational levels. Turchin (1977) called this the “stairway effect”, where small transitions may induce larger transitions, such as the production of tools serving the production of further, more complex tools (ibid: 89ff.). Another example is a technology scaling up from its originating market niche and entering the regime level, e.g., becoming a product of widespread diffusion. In this regard, this effect is related to the

economies of scale. However, the term "stairway effect" as such is misleading as it suggests a simple hierarchical order which is hardly given in complex networking effects.

Processes akin to MST are observable in many different real-world phenomena. Besides evolutionary processes such as the emergence of multicellular organisms, biological processes in animal and human development etc., Turchin (1977) outlined a number of examples such as the creation of language, cultural achievements as the production of metal tools, community building and social integration, the division of labour, and other economic and technological developments perceivable as metasystem transitions (Turchin 1977; Heylighen/Campbell 1995). Indeed, a three stage model alone cannot grasp the high complexity of real-world phenomena such as transformations of social systems. Nevertheless, the major benefit of the MST perspective is that it facilitates to abstract this complexity and allows grasping the dynamics of systemic change along its three basic transition phases. This is useful to heuristically gain a simplified view on the dynamics of real-world phenomena from a wider perspective, which is supportive to reveal according socio-technical transition paths and trends. On this basis, further, more detailed explorations can be conducted.

The analytical lens MST provides is particularly supportive to explore socio-technical change. In general, this change is not an isolated or deterministic process, but it involves several transformation processes resulting from a complex interplay of actors, organizational institutional contexts, markets, technologies and infrastructures (Kemp et al. 2001). This set of transformation processes can be subsumed under the term system transition. Socio-technical transitions are transformation processes that entail wider changes in society or a societal subsystem. These processes are co-evolutionary, unfold within long timescales, involve multiple actors and include changes that encroach upon existing technology and user practices (cf. Kemp et al. 2001; Geels 2004; Geels/Schot 2007). The inherent dynamics of socio-technical transitions are determined by interactions between systemic components which entail or foster network effects (cf. Katz/Shapiro 1994). At the same time, path dependencies and technological lock-ins emerge (Arthur 1989), which benefits further diffusion of a technology but also aggravates the emergence of alternatives to the prevailing socio-technical system and its related usage practices. The diffusion and usage of a socio-technical system has consequences for society. A transition can be understood as multi-level transformation, whereas a typical distinction in three interrelated hierarchical levels, i.e., sociotechnical niche, regime and landscape is made in science and technology studies (Kemp et al. 2001; Geels 2004; Smith et al. 2005; Geels/Schot 2007). The rationale of these levels is that novel technologies, applications, functions etc. emerge in technological niches, which are limited domains, serving as sort of "incubation chambers" for novelties supported by a small group of actors to gain internal momentum. Niche developments are influenced by regimes, i.e., the cluster of existing technologies, organizations, knowledge, rules and norms, socio-technical practices etc. The landscape represents the exogenous environment whereas niches and regimes have no direct influence upon (Geels 2004; Geels/Schot 2007). These levels can be seen as equivalents to micro level, macro level and system environment. Niches are rather isolated and have little or no interactions with other entities at the regime level. When a niche development gains momentum, its relations and interactions increase and it may become

integrated into the regime level; where it may gain further momentum and have wider impacts on society, i.e., the landscape level. These dynamics at multiple levels have certain similarities with a metasystem transition. For real-world systems, their settings and characteristics are apparently not always easy to define, particularly not when social systems are involved. It is thus an analytical decision where e.g., to draw the boundary of a system and which elements are seen as system components.

2.3 Employment of systems theory in this research

The outlined characteristics of a system underline that it is more than a set of elements triggering particular mechanisms with specific (determinable) effects. A system can serve multiple purposes with multiple, also unintended side-effects. For instance, if an element of a system changes (e.g., due to an environmental impact), this can have effects on the system as a whole as well as on its environment.[8] When a socio-technical system emerges it becomes a (contextually embedded) part of societal reality which consequently has an impact on society. As a socio-technical system is a sub-system of society, the way it is implemented and used has effects on those societal processes, functions, practices etc., the system relates to and vice versa. There is thus a mutual process of co-shaping at different levels. A systemic perspective offers a research heuristic to explore transformation patterns of socio-technical phenomena including this complex interplay. Through different layers of abstraction, it enables to put emphasis on the system, its constituent elements and their relations, structural settings and the resulting interactions and dynamics. On the one hand, the conceptualisation of a phenomenon as system allows focusing on its (internal) components and dynamics. On the other hand, it also supports to investigate how the system is related to external factors of its environment. This is supportive for the assessment of societal impacts resulting from socio-technical systems, such as represented by ICTs. In general, a new technology can enable innovation, new strategies to solve existing problems and improvements of socio-technical settings, processes, functions, applications, practices etc. But it can also constrain those existing items and practices which can induce transformation patterns as societal actors try to cope with these challenges. For instance, mobile phones enabled mobile communications and permanent accessibility but constrained possibilities of solitude and being unavailable. The same dynamics occur in other socio-technical systems as well. In our case, the focus is on two vital societal concepts and functions and their very dynamics in relation to ICTs: privacy and (digital) identification. Put simply, ICTs enabled and reinforced the latter but led to constraints of the former.

In this research, the outlined system-theoretical framework serves as a heuristic to grasp the socio-technical transformations of ICTs with emphasis on digital identification and its dynamics as an integral part thereof, from a wider, meta-perspective. ICTs are

[8] A drastic example is the nuclear disaster of Fukushima in 2011, where an earth quake triggered a series of critical effects on components of the nuclear power plant which led to a nuclear meltdown with accordingly critical environmental and societal consequences. The emergence of the Internet bears various positives examples, such as it fostered free access to information worldwide at individual as well as organizational and societal level.

framed as socio-technical metasystem through the lens of a MST in order to gain a basic structure of its main socio-technical transition paths. This is important to investigate the particular dynamics of the emergence of digital identification and its driving forces. Based on this, a central aim is to analyse those mechanisms related to digital identification, which constrain the effectiveness of privacy protection. A basic assumption of this study is that the processing of information related to an individuals' identity is the main determinant of a (informational) privacy-affecting activity concerning this individual. This does not necessarily imply that privacy is violated. But this nexus between identity information and privacy is crucial for the assessment of the associated impacts. The processing of identity information has an enabling effect for a privacy impact to emerge. The basic characteristics of a system are employed to this phenomenon in order to explore its impact on privacy and develop a framework for privacy impact assessment with a focus on identifiability.

From a systemic perspective, the interplay between privacy and identity in relation to ICTs can be grasped as follows: Information processing is generally essential for the dynamics of a socio-technical system, which depend from the relations and interactions between its different elements (cf. Fuchs/Hofkirchner 2005; Hofkirchner 2013). An evolutionary system has three (hierarchically) interrelated information-generating functions (Hofkirchner 2013: 184ff.): cognition, communication and co-operation. Cognition is "the individual, (internal) generation (and utilisation) of information by a system". Communication "is the interactional, interfacial generation (and utilisation) of information by (co-)systems". Co-operation is "the collective, external generation (and utilisation) of information by (co-)systems in conjunction" (ibid). An interaction between two entities (or systems) requires initial processing of information (enabling the interaction). Information thus allows establishing links between different systems. Interaction may stimulate the generation of further information. Hence, an increasing degree of interactions and networking stimulates a growth in information processing. ICTs represent a multifunctional conglomerate of different socio-technical subsystems which is perceivable as an information metasystem. The essence of the interactions and dynamics of this metasystem is the processing of (digital) information which (amongst various other societal impacts) affects the digital representation and availability of individual identities as well as identification practices, which has consequences for the functionality of privacy and data protection. In the scope of this research, identification is basically seen as form of (re-)cognition which enables further information generation and processing. However, a further discussion about these general functions is not intended in this study.

Identity is framed as the informational representation of an individual (human) entity (A). Identification involves the processing of identity-related information by another entity (an individual or institution – system B). In its simplest form, there is a relation between the systems A and B, whereas B processes identity information representing A. As soon as a technology is involved, an additional entity C as an intermediary occurs which creates additional relations and interactions. Figure 4 provides a simple illustration of this setting which may be seen as an identification system whereas A, B and C are sub-systems:

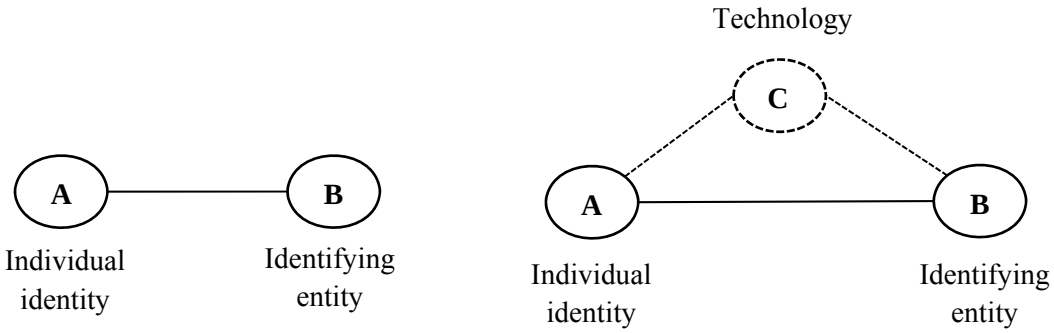


Figure 4: Simple identification system without and including a technical system

Depending on the form of identification, the identity may be represented implicitly by the presence or actions of the individual person, or explicitly by some technological artefact (e.g., an ID card or an online user profile)[9]. This simple illustration highlights the influence of technology on this setting as it transforms the relations and interactions in the system and thus its structure. It thus has an effect on the identification system as a whole. This is a standard-setting in contexts of every-day-life where identification occurs, and it is basically alike in analogue as well as in digital contexts. However, given the peculiarities of ICTs, the consequences of digital identification are different than in analogue (or offline) settings. ICTs enable digital identification partially decoupled from spatial and temporal limits: the additional, technical system (e.g., an online service) involved in the processing of an individuals' identity information may have networking capabilities. Hence, the input to this system is the (digitized) identity information. To some extent, this information then represents the individuals' digital identity. As this technical system is networked, this digital identity or parts of the information can be passed on to further socio-technical systems. Identity is naturally a relational concept, affected by relations and interactions with others. Hence, in this regard, the relations and interactions of the individual identity increase by the use of ICTs. To some extent, the digital identity gains a dynamic on its own as it can be used also decoupled from the individual it represents. A person can be identified by her digital information without being directly involved. Considering that ICTs incorporate manifold socio-technical systems and application contexts to process digital identity information, this network effect on digital identity is likely to boost. Basically, all socio-technical devices a person uses or interacts with, may involve the processing of information about her which implies that some relations between those devices and the identity of this person exists. Although these relations are mostly context-dependent and not persistent, the generated information can be available for other purposes, decoupled from the original processing context. This has manifold societal implications and particularly on the individuals' privacy as her information is being processed, which is of main interest in this research.

A further, related issue concerns the inherent connecting function of identification which enables to establish a link between different entities. The processing of identity-related information is assumed to be an essential process that contributes to the emergence

[9] For more details about the notions and functions of identity and identification see Section 3.1

of connections between socio-technical (sub-)systems. In social contexts, identification is commonly understood as revealing an individuals' identity. This occurs in a formal sense where a person is prompted to show her passport etc. or in less formal contexts where a person introduces herself to others. However, understood in a broader sense, we can speak of identification already when particular information is being processed which refers or relates to an individual entity. In this regard, identification is a form of (re-)cognition. A basic premise here is that in order to establish a connection respectively relation between different entities, some form of (re-)cognition process, i.e., an exchange of information *about* the involved entities is needed. This does not necessarily imply that the factual identity of an individual is completely revealed. For example, already the process of recognizing a by-passing person, so that it is perceived as distinct from another, involves some form of identifiable information (such as facial features or other bodily characteristics). In situations in the physical world, where identification does not happen on purpose, and therefore, no interaction emerges, this is mostly an unconscious process with limited impacts on privacy as this information diminishes and is not recorded (or made available by other means). In the digital world, though, the implications are different as identity information can be gathered more easily. Furthermore, identification is often involved when different digital systems or its elements establish a connection to interact with each other. If one system (sender) interacts with, or transmits information to a particular other, it initially needs some piece of specific information (such as an identifier) about this particular system because otherwise, there is no defined receiver of this information.[10] This is in line with the classical sender-receiver model, as introduced by Shannon (1948), which had strong impact on the development of information systems. As highlighted and explored in-depth by Hofkirchner (2013) in his information theory, the emergence of information as such is, though, far more complex than this classical reductionist notion suggests. When human entities are (directly or indirectly) involved in an interaction between two or more information systems, also digital forms of identification inherent to technology can have privacy consequences.

Also privacy can be explained in systemic terms: (informational) privacy is here seen as a concept with an inherent boundary control function (Section 3.2) which enables the individual to regulate her relations and interactions to others. This implies informational self-determination and individual control over her information (concerning her identity). From a wider perspective, informational self-determination implies autonomy and may be seen as a concept related to the self-organization of an individual (Section 3.2.3). Put simply, the boundary that is determined by privacy is the threshold between the private sphere of an individual identity and its (socio-technical) environment or public sphere. The basic function of privacy is to enable an individual person in regulating the extent to which her information crosses the boundaries of her private sphere. ICTs complicate this function as they enable and reinforce the processing of identity information and thus identification beyond the individual's control. But the transformations of identification and privacy are obviously not reducible to technological progress but result from a complex interplay of

[10] An exception is broadcasting, where information is distributed to a dispersed mass of entities. However, even in this case, some kind of identifiable information related to the sender is involved (such as a unique frequency of a radio station, or a TV channel).

societal factors. Hence, ICTs amplified identification processes, but the usage contexts of these processes emerge from and are driven by the dynamics of societal practices.

Identification is basically a vital process for the functioning of society, serving various social, political and economic purposes. ICTs created new possibilities to implement and employ this process in various domains. ICTs inter alia stimulated a growth in electronic transactions and personalized online services in public as well as in private sector. This boosted the processing of personal information. Consequently, identity management systems as well as the integration of identification mechanisms into online services increased, as prominently exemplified by Web 2.0 and social media. These developments led to a paradigm shift in the handling of personal information with an extension of identification in many respects. The enormous diffusion of ICTs with an entailed broad availability of information about individuals stimulated a broad range of business models based on the commercialization of this information. This affected information self-determination as well as security and surveillance practices. Hence, in brief, the emergence of digital identification results from and is driven by technological development, several political and economic interests located in regimes of the digital economy as well as of security and surveillance actors. Digital identification has an ambiguous function in relation to privacy: it can contribute to improve the security and efficiency of identity verifications, e.g., in electronic transactions or online services (see Section 4). In this regard, it contributes to regain control over digital information. However, identification itself can be used as a mechanism to control the individual for economic as well as security purposes which can lead to limitations of privacy (see Chapter 5).

These developments affect the privacy regime, which is constituted by governance practices, legal frameworks and policies with incorporated values and norms, social, political and economic practices etc., as well as the conglomerate of public and private institutions, organizations and other entities processing personal information. To some extent, there are tensions between the privacy regime and those domains where identification is employed and reinforced. This is particularly the case, when identification is employed as a control mechanism at the cost of individual privacy. In this regard, there is an assumed control dilemma of digital identification as it attempts to regain control over digital information which can lead to a further loss thereof for the individual concerned (see Chapter 5). Or in other words: the socio-technical transition paths (of ICTs and digital identification), alter the requirements for privacy protection and reinforce the pressure on society to adapt to technology. To compensate this loss of control and foster the effectiveness of privacy protection requires additional control mechanisms for the individual as well as for the information processing institutions as part of the privacy regime. This includes enhanced approaches for privacy impact assessment, which can also support the implementation of privacy by design approaches (see Section 6).

2.4 Structure of this research

The issues presented in the previous section build the foundation of this research which is structured as follows: of main interest are the implications of digital identification for the protection of privacy. As starting point, the analysis deals with the interplay of identity, identification and privacy, as investigated in part 3 of this work. This is mainly based on theoretical and empirical literature from privacy and security studies. After a description of the basic concepts of identity and identification, the role and meaning of privacy is explained. The analysis includes central controversies in the privacy discourse such as the assumed trade-off between privacy and security. This trade-off is deconstructed and its close relationship with security policy in the realm of securitization is discussed (which is relevant for the analysis in the main part 5). In line with the notion of privacy as a control function, the interplay of privacy, autonomy and identity is presented and discussed which mainly concerns the concept of informational self-determination. The section ends with a brief discussion on the relationship between privacy and transparency including some issues regarding a lacking of informational control which is a central issue of contemporary privacy protection.

Part 4 of this research sheds light on the emergence and transition paths of digital identification. A starting point is a brief overview on the main development stages of ICTs and the relevant socio-technical systems and practices through the lens of a metasystem transition. The timeframe for these stages is from about the 1990s until today whereas the emergence of the Internet and the WWW serves as a general point of reference. Based in these general transition paths, more emphasis is then put on digital identification. This is done by exploring the major drivers, basic functions and technical approaches of digital identity management (IDM) as well of social media platforms and particularly social networking sites (SNS) as a prominent case study for the networking structure of ICTs and their effects on interactive identity representation as well as its processing. The section finishes with a discussion on the transformation patterns of digital identification in the frame of a metasystem transition.

Part 5 analyses in-depth, to what extent the boundary control function of privacy is affected by ICTs and digital identification. The starting point is an assumed control dilemma of digital identification whereas the basic mechanisms and functions of this dilemma are explored. The analysis involves theories and concepts from surveillance studies to explore the nexus between surveillance and identification as well as its political and economic drivers. The investigation deals with the question how identification practices can reinforce existing mechanisms of power and control, as well as the implications and core challenges for privacy protection. Included in this analysis are empirical findings from the SurPRISE project, where the perceptions of European citizens on the interplay of privacy, security and surveillance were examined. Based on the revealed core issue of information asymmetries between individuals in institutional entities processing their information, the focus is on socio-technical identifiability as a core problem of contemporary privacy protection. This problem is highlighted by the concept of an identity shadow which enables implicit and explicit identification beyond the individuals' control. Then, relevant technological trends and developments are presented

and discussed which contribute to a further expansion of individual identifiability. This is followed by an analysis of the prospects and perils of existing privacy control mechanisms in the realm of privacy by design. The section finishes with a discussion about the limits of individual control against the background of a trend toward a "privatisation" of privacy.

The main part 6 of this research explores concepts and approaches of privacy impact assessment (PIA). This begins with a general overview on PIA, its basic requirements and current limits in relation to the core problem of information asymmetries and identifiability as explored in part 5. Then, existing typologies of privacy types and privacy-affecting activities are examined and discussed. On this basis, a conceptual model is elaborated which refines existing PIA approaches by integrating identifiability as a central privacy-affecting mechanism. This framework allows for a more systemic, process-oriented view on the emergence of a privacy impact. A fundamental part of this proposed identifiability-based PIA framework is a typology of identifiability and the basic different types of identifiable information which may also contribute to improve the theoretical understanding of privacy impacts. The final Section 7 summarizes and discusses the main findings of this research and presents concluding remarks.

CHAPTER 3

The interplay between identity, identification and privacy

Privacy and identity are intrinsically related concepts. In general, this interplay is rather obvious because privacy concerns the private life of the individual. Intrusions into an individual's privacy affect her identity and consequently, protecting privacy embraces the protection of an individual's identity. Therefore, privacy and data protection laws particularly address personal data, mostly understood as information related to the identity of a person. The processing of this information is a form of identification. Identification is commonly known as process of determining who a particular person is. It is an important societal process ranging from personal and professional relationships, service provision in public and private sectors (e.g., citizen-government as well as business-customer relationships) and so forth. But forms of identification are also involved in profiling, security and surveillance practices. Irrespective of the function it fulfils, identification is an information process. This is particularly important because ICTs affect identification in many respects and basically transform the modalities of this information process. Today, ICTs are nearly ubiquitous technical tools which informationally enrich our identities so that they are digitally available. Digital (or electronic) identification became a standard procedure in many domains.[11] This has serious implications for the protection of privacy which are of main interest in this research. In order to explore these implications, it is necessary to shed light on the interplay between identity and privacy first. As a starting point, the basic notions and functions of identity and identification are briefly presented and discussed. Then, the role and meaning of privacy is outlined, including a brief overview on legal issues and protection principles, as well as the crucial function of privacy as a boundary control between the private and the public sphere. This is followed by an exploration of the interrelations between identity, privacy and autonomy including the important role of informational self-determination. Finally, relevant controversies of privacy with other concepts, namely, security and transparency are presented and discussed.

3.1 Basic notions and functions of identity and identification

3.1.1 What is identity? – Overview on concepts and characteristics

Identity and identification are multifaceted phenomena. A number of philosophers have ever since dealt with personal identity as the substance matter of existence, literally ranging from birth to death (cf. Korfmacher 2006). The question "what is identity?" is non-

[11] The terms "digital" and "electronic" identification are used synonymously here.

trivial and an issue of concern for many disciplines (ranging from philosophy, sociology, anthropology, psychology, neurology, history, legal studies, gender studies, computer and information sciences etc.). Hence, obviously, there is no simple, meaningful answer to this question as identity means different things in different disciplines and taking all these issues into consideration is far beyond the scope of this work. Nevertheless, there are some important common features and above all, identity provides information representing a particular individual entity. In this research, the identity of an individual person, respectively personal or individual identity is primarily meant when speaking of identity. Identity is basically understood as the very concept describing and representing the specific characteristics of an individual. It is of less concern here what identity exactly is, but more relevant is how identity is represented by (digital) information in socio-technical contexts. In the following, some basic features of identity are outlined which are relevant in this regard.

At a general level, identity is represented by a number of individual characteristics that refer to a person. In social interactions of every-day-life, the most common identity attribute is the name of a person. Identity is unique in the way that it allows to distinguish one person or entity from another. In this respect, identity is the construct of a set of properties that determines the uniqueness of a particular entity in relation to others.[12] Personal identity may be seen as "sameness of a same person in different moments in time" (Rundle et al. 2008: 7). For Wiener (1954: 96) homeostasis (i.e., the self-regulatory capacity of a biological system) is "the touchstone of our personal identity", which implies a permanent interaction with its environment. In a similar vein, Varela (1997: 76ff.) described the identity of living systems as an "autopoietic unity" with the capability to maintain itself, making it distinct from its environment. At the same time, it maintains its relations with its environment which are vital for its existence. In this regard, identity can be seen as construct that maintains itself, distinct from others, but at the same time, it is shaped by every interaction with its surroundings. Thus identity is continuingly progressing, based on the dynamics of its relations and interactions. This process is also a physical feature of biological systems: For Varela (1997: 73), "living identities are produced by some manner of closure, but what is produced is an emerging interactive level". This dynamic is observed in societal contexts as well: Giddens (1991) argues that a person's identity is coupled with her biography which continuingly proceeds. Hence, "(…) identity is not to be found in behaviour, nor – important though this is – in the reactions of others, but in the capacity to keep a particular narrative going. The individual's biography, if she is to maintain regular interaction with others in the day-to-day world, cannot be wholly fictive. It must continually integrate events which occur in the external world, and sort them into the ongoing 'story' about the self" (Giddens, 1991: 54). Similarly, Paul Ricoeur (1992) highlights aspects of continuity and change whereby he distinguishes between two meanings of identity: idem and ipse. The Latin word "idem" means "same" and refers to the physical characteristics for which "permanence in time constitutes the

[12] For instance: a tree is an entity. There are many different trees in a forest. Without additional information about its properties, one tree is not distinguishable from others. Information about the "identity" of a tree such as shape, location, type of tree (e.g., apple tree), age (number of its annual rings), height etc. allow differing one tree from others. Hence, information about its properties gives identity to an entity.

highest order" (Ricoeur 1992: 2). With ipse or "selfhood", Ricoeur means the part of identity that is dynamic and changeable through one's lifetime. In this regard, identity is dialectical as it comprises both, sameness and selfhood. Similar to Giddens, he also pointed out that identity is narrative (Ricoeur 1992: 147f.), i.e., that identity is partially constructed and determined by the relations to its environments it interacts with. Hence, the different notions of identity share an important commonality: identity is understood as a concept with permanent or stable, as well as dynamic features. Or in other words: identity is a permanent as well as a dynamic concept. Given its dynamic character, identity is thus not reducible to the sum of its (informational) parts. Nevertheless, a certain set of attributes can be sufficient to uniquely identify an individual in a particular context. In this regard, identity entails unique information to allow for recognition of an entity at a certain point in space and time.

Identity features a number of different, intertwined properties (cf. Rundle et al. 2008; Pfitzmann/Hansen 2010; Whitley et al. 2014), which can be briefly summarized as follows: identity is

- Social: humans are social beings and social interactions to some extent need a foundation for recognition of a person, referring to an identity.
- Subjective: the notion of identity differs from person to person as well as the interpretation of the attributes linked to a person is subjective. One (partial) identity can have different (subjective) meanings in different contexts.
- Valuable: identity offers some certainty and confidence between interacting individuals, enables the creation of relationships and can be functional to enable transactions.
- Referential: The items, artefacts respectively the information used for identification links back to an individual (or more generally an entity). But an "identity is not a person" but "a reference to a person" (Rundle et al. 2008: 26).
- Composite: Identity information can consist of many different sources also without the involvement of the individual concerned.
- Consequential: Identity information provides manifold insights into personal details, actions and behaviour. False provision as well as disclosure of this information thus have consequences and can be harmful.
- Dynamic: Identity is not a static concept but changes over time.
- Contextual: Identity is not universal but context-dependent. A person can have different identities in different contexts and separating them contributes to privacy and autonomy.
- Equivocal: "The process of identification is inherently error-prone" (Rundle et al. 2008: 26), because there can be e.g., duplicate identity information, the information can be wrong or incorrect, or this information can be misused in different contexts etc.

The role identity plays in society results from many different but intertwined dimensions (such as social, psychological, economic, cultural, political, technological, organisational, legal etc.). Individual persons have various functions in social, economic and political

domains with roles such as citizens, employees, customers and consumers, partners, members etc. (cf. Raab 2009). Judith Butler (2006) speaks of "identity performance" and deals with the dynamics of identity by understanding it as a performative concept, meaning that identity is socially constructed. For Whitley et al. (2014: 19) personal identity is a "practice-based, relational and dynamic" concept. They further state that "identity is produced and reproduced through ongoing communicative activities that take place within and across people and organisations" (ibid). David Lyon (2009: 10) argues that identity is "a composite and malleable entity, with some elements or derived from the corporeal person and others from categories or collectivities in which they fit". These categories are created and vary from the broad range of usage contexts in which identity is embedded such as social and organisational practices, political system, commercial applications etc. Depending on its usage and implementation of technological identity artefacts, identity may convey particular values and is related to policy and regulation (such as privacy and data protection). Hence, In a broader sense, identity can be seen as a socio-technical co-construct, shaped by a number of interrelated factors such as social, economic, political, technological dimensions and many other issues (e.g., artefacts, knowledge, policy and regulation, infrastructure, cultural meaning, markets) as the Figure 5 below illustrates.

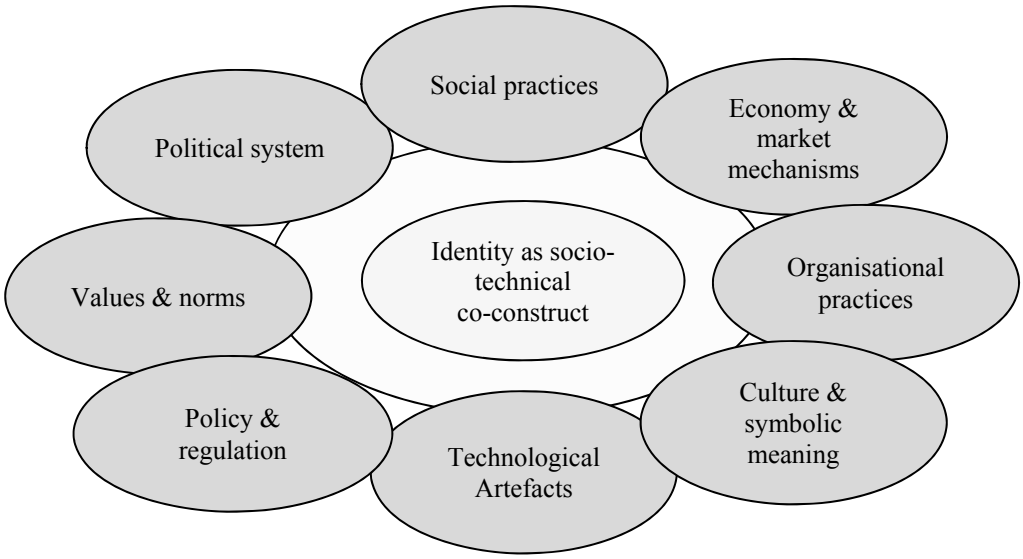


Figure 5: Socio-technical determinants of identity construction

During lifetime, a person is represented by a magnitude of information in multiple different socio-technical contexts in which this information is gathered, collected, processed, stored etc. Although there are basic types of information representing an identity (such as a name), identity not merely composed of a core set of information. It emerges and develops further depending of the interactions with its (e.g., social, economic, political, technical) environment. Referring to Varela, Hildebrandt (2006: 54) states that "the most interesting thing about human nature is its indeterminacy and the vast possibilities this implies: our non-essentialist essence is that we are correlatable humans before being correlated data

subjects". Therefore, "[w]hatever our profile predicts about our future, a radical unpredictability remains that constitutes the core of our identity" (ibid). This core of identity consists of continuity as well as change and is a basic building block of individual freedom. Given its dynamics, "there is no such thing as 'the identity'" (Pfitzmann/Hansen 2010: 30), i.e., an individual is not represented by a universal identity but can have multiple, partial identities (or roles) in different contexts. Identity is thus context-sensitive as it is a "representation of an entity in the form of one or more attributes that allow the entity or entities to be sufficiently distinguished within context" (ITU 2010: 4).

3.1.2 Identification

As outlined, identity is a complex, dynamic concept with multiple features and dimensions. A central function of identity in society is to provide a means to determine or ascertain who a particular person is. Knowledge about a person's identity contributes, e.g., to the building of trust and security among interaction partners in social, political or economic contexts (such as commercial transactions). This process of determining or recognising an individual by her contextual characteristics is identification (ITU 2010: 3). Subject to identification can be an individual human person but basically every object or any kind of entity.

Identification has an essential feature, which determines its significance: it has an inherent connecting function as it enables to establish a link between different entities. Identities thus have the ability to "build and articulate ties to other identities in network-domains" (White 2008: 2). Identification is an integral part of socio-technical contexts, which can be categorised in at least four basic (overlapping) domains: (1) In the social domain, identification is important for interpersonal relationships in general as it fits the societal need to communicate and interact with each other, building relationships etc. (2) In the political domain, identification enables citizenship and the related rights that allow individuals to engage in the political system but it is also used to control the rights and duties of individuals. (3) In the economic domain, identification is related to the provision of public and private services, conduct transactions etc. and to contribute to procedural security. (4) The technological domain serves as vehicle for the other domains. Here, identification is implemented in information systems to enable socio-technical interactions, information exchange etc. and support the other domains with technological means. These domains are basically included when speaking of socio-technical contexts in this research.

Hence, identification of individuals is a practice of everyday life and a basic instrument of governance, serving a variety of functions across all societal domains with a long history and obvious relevance for the functioning of society, polity and economy (cf. Clarke 1994a/1994b; Bennett/Lyon 2008; Lyon 2009; Whitley et al. 2014). Besides ancient forms of identification, the genesis of the modern state in Europe paved the way for the identification of citizens (such as in contact with public administration, at national borders etc.). Among other scholars, David Lyon (2009) provides an overview on historical developments of identification. Early approaches of administrative schemes for purposes of authentication and identification of individuals date back to the mid-1500s. Some people had to wear special badges or insignia to prove legitimacy of their activities,

such as pilgrims, couriers, or diplomats. For example, couriers carried particular signs when submitting a message. Identity badges were also used to mark minority groups such as homeless people, ethnical minorities or immigrants. The emergence of citizenship and citizen registration after the French Revolution contributed to the employment of identity documents such as passports, which became increasingly common in many countries by the twentieth century (Lyon 2009: 20ff.). The identification and registration of persons has a long tradition and is a core function of government and public administration worldwide. In most national states, citizens are initially equipped with a legal identity, beginning with the enrolment of a birth certificate, which then serves as a basic credential to receive further identity documents. Governments have complex administrative infrastructures to provide official identity devices to their citizens (such as birth certificates, passports, driving licences, social security cards, particular citizen ID card schemes etc.). ID documents proof one's legal identity and enable citizenship, including the entailed rights and responsibilities to be involved in the functioning of society. Besides its relevance for administrative procedures, identification is also an important means of criminal investigations, law enforcement, national and international security. Basically, individuals become identified in many socio-technical domains and contexts: identification serves as mechanism to establish a certain amount of trust in social interactions between individuals, groups or institutions, among interacting business partners, customers and vendors in commercial transactions, customer relationship management (CRM) etc. Different forms of identification enable access to a broad scope of services in public and private sector (e.g., in the fields of e-government, e-health, e-commerce and e-business etc.), implemented and amplified by the use of ICTs (cf. Lyon 2003; Raab 2006; Bennett/Lyon 2008; Hildebrandt/Gutwirth 2008; Lyon 2009; Aichholzer/Strauß 2010b; Kubicek/Noack 2010b; Whitley et al. 2014).

In general, a crucial aim of identification in socio-technical contexts is to reduce uncertainty and to improve security, e.g., of individuals, organisations, processes, applications, systems etc. (cf. Clarke 1994a; White 2008; Bennett/Lyon 2008; Raab 2009; Lyon 2009; Kubicek/Noack 2010b; Strauß 2011; Whitley et al. 2014). In this regard, Identity is also linked to control efforts, as White (2008: 1) points out: "An identity emerges for each of us only out of efforts at control amid contingencies and contentions in interaction. These control efforts need not have anything to do with domination over other identities. Before anything else, control is about finding footings among other identities. Such footing is a position that entails a stance, which brings orientation in relation to other identities." With footing White means that identity has a certain foundation which makes it unique, tangible and controllable. In a less abstract sense, this implies a certain piece of (identifiable) information that allows recognising an individual based on its characteristics, distinct from others. This identity (or identifiable) information represents an individual entity in a particular context. Thus a necessary condition of identification is the existence of a certain amount of information serving as (informational) representation of an identity.

As information is a necessary condition for this recognition process, identification can be defined as the processing of information related or referring to the identity of a particular individual. As Figure 6 illustrates, in its broadest sense, identification implies that (at least) two entities interact, whereas identity information is exchanged: The to-be-

identified entity (A) interacts with the identifying entity (B) that processes the identity information representing entity A. In social interactions between two human entities, this process is usually bidirectional, meaning that the entity A and B switch roles as also B becomes identified. In socio-technical systems, though, individuals often become identified in a unidirectional way by e.g., institutional or organisational entities during usage of a technology.

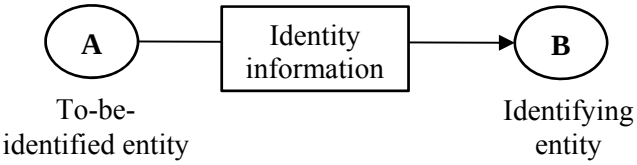


Figure 6: Simple unidirectional identification process

As identification is a mechanism with at least two entities involved, the technical processing of identity information links at least two systems. Consequently, the control over the identification process, and therefore, over the identity information is not merely a matter of a single but of multiple entities. As will be shown, this has impact on the protection of this information and thus on privacy of the identified individual. It also makes a difference, whether a person proactively reveals her identity, or a person is identified without her being directly involved in the identification process. Of particular interest in this research are forms of the latter.

There are several types of identity information. Clarke (1994a) described the following basic categories of identification: (a) names; (b) biometrics, including appearance – how the person looks (e.g., gender, height, weight, colour of eyes, hair colour etc.), bodily and physical characteristics – what the person is (e.g., biometric features such as fingerprint, iris pattern, DNA etc.), or what the person has (such as glasses, tattoos, piercings, or any other additional bodily feature), bio-dynamics and behaviour – what the person does or how she interacts (e.g., pattern of handwritten signature, voice pattern, movement patterns, style of speech etc.); (c) codes and schemes – how the person is represented (e.g., by an identification number such as a passport no., social security no. and so on); (d) knowledge – what the person knows (e.g., a password or a PIN code, i.e., a personal identification number etc.); (e) tokens – what the person has (e.g., an ID document such as a birth certificate, passport, or any other identity device). Gary Marx (2001) suggests a partially different categorisation and speaks of different types of information usable for identification as "identity knowledge". This is, for instance, legal name; information about temporal and spatial location ("locatability") such as address; pseudonyms used in combination with the other two, i.e., an identifier such as phone number, social security number etc.; pseudonyms not relatable to other types and thus providing anonymity; pattern knowledge, i.e., information referring to the distinctive appearance or behavioural patterns of a person (Marx 2001). Identification usually involves social categorisation with categories such as age, gender, nationality, religion, class, employment status, profession, health status, sexual orientation, memberships, activities, relationships etc. Additional categories may comprise financial status, credit scores, individual preferences, lifestyle,

consumer habits, and so on (cf. Marx 2001; Raab 2009). In privacy contexts, identity information is labelled personal data or personally identifiable information (PII) (see Sections 3.2 and 6).

There is no clear-cut answer to the question what attributes refer an identity neither can there be an exhaustive, comprehensive list of such attributes. A reason for this lies in the complex relation between identity and time. On the one hand, identity is permanent and uniquely represents an entity in contrast to others. On the other hand, identity is constructed and context-specific with a highly dynamic momentum. In other words: identity is not static but it evolves over time (cf. Abelson/Lessig 1998; Pfitzmann/Borcea-Pfitzmann 2010). Thus the representation of an "identity as a set of attribute values valid at a particular time can stay the same or grow, but never shrink" (Pfitzmann/Borcea-Pfitzmann 2010: 3). While some attributes might remain steady (such as birth name, date of birth) others may change over time (e.g., weight, hair colour etc.). Therefore, attributes remaining relatively constant over a longer period of time are basically more suitable for identification than rather dynamic ones. Hence, generally speaking, it is neither possible nor necessary for identification to have a large, complete set of identity attributes: depending on the application context, it is sufficient to have a certain amount of information that represents the uniqueness of a person in a certain context. Important is that the uniqueness of one or more attributes is not without limits but context-sensitive and only valid in a certain time frame. Therefore, different identity information is mostly combined, depending on the purpose of identification and its security requirements. A combined set of identity attributes can lower the risk of errors or insufficient identification, e.g., resulting from duplicates. For instance, typical attributes like name or date of birth alone cannot be expected to be unique because there can be more people with the same attributes. But a combination with additional information such as the home address, representing the location of an individual is often sufficient for unique identification. In general, whether a person is uniquely identifiable (without ambiguity) or not depends from the environment of that person. In a crowd of 100 people, for example, the hair colour is likely to be insufficient to identify a person. In combination with eye colour and name, the likelihood for identification increases as ambiguity is likely to decrease with the number of attributes.[13] Or in other words: a combined set of identity attributes enables to draw a distinct pattern related to a particular person. Thus basically, identification implies the recognition of a unique pattern that allows distinguishing one entity from another (cf. Rundle et al 2008; ITU 2010). Once an applicable pattern is found it can be represented by or transformed into a code (e.g., digits, characters of both) – typically an identifier – which then can be used to link to a set of different identity attributes. Identifiers facilitate identification (as it refers to a collection of attributes) and they also allow cross-linking data over different contexts. In practice, code schemes are commonly used as an implicit component of most identification forms.

As regards the modality of identification, there are two basic approaches: knowledge-based identification makes use of information that a person is expected to know for identity

[13] This is an issue of mathematics and statistics, which is not discussed more in-depth in this work.

verification. This type is wide-spread and most common as combination of username and password, or PIN codes. Token-based identification is based on an ID device as identity proof and is well-known in everyday life in the form of ATM[14] cards, social security cards, passports, debit cards etc. as well as of mobile phones. In general, the combination of knowledge and possession of a token usually provides a higher level of security than approaches without a physical device (though obviously depending from the quality of the implementation). These general approaches are also known as multi-factor authentication, which is often used as a method for (computerized) access control or security measure for transactions (such as in e-banking systems). Another common distinction is between the categories knowledge (what one knows; e.g., a password), possession (what one has; e.g., an identity document), and inherence (what one is; e.g., biometric information such as a fingerprint) (cf. Clarke 1994a; De Cristofaro et al. 2014).

Irrespective of these categories, identification requires the representation of identity information, either directly by the person or by some kind of object. In social, interpersonal relationships, identification is often an implicit process. For instance, a person that is already known does not need to be explicitly identified because her sheer physical presence may be sufficient as well as the knowledge of her name, frequently used as a typical identifier. But besides that, identification mostly requires some kind of artefact (such as an identity token, device, scheme, or credential), which determines the form of identity representation. An ID artefact serves as means to recognize a person and/or entity by providing information to ascertain who one is, i.e., as tool of identification. Depending on the application context, an ID device can be formal or informal. Typical formal identity artefacts (or schemes) are for example identity documents like a passport, driver's license, social security card etc.; less formal are typical user profiles based on credentials (e.g., username and password) such as in computer applications or information systems. Therefore, technical definitions describe identity often as "collection of attributes", "collection of claims" (cf. Cameron 2005; Rundle et al. 2008) or "a set of attribute values related to one and the same data subject" (e.g., Pfitzmann/Borcea-Pfitzmann 2010: 3). Identity attributes respectively identity information can serve as identifiers. In information systems, identity information is mostly associated with one or more (unique) identifiers. An identifier is information referring to a certain identity's collection of attributes or claims (cf. Rundle et al. 2008; ITU 2010). A unique identifier allows distinguishing a set of identity attributes from another, at least in a particular context (such as a telephone number, passport number, user ID no. etc.). If identity information is represented by digital means, we can speak of digital identity. Thus identity can be understood as a concept of the "real world" as well as a (digital) artefact representing an individual entity. The field of identity management (IDM) deals with peculiarities of digital identities and identification, which is described in Section 4.2. Technical definitions see it as "a digital representation of a set of claims made by one party about itself or another data subject" (Rundle et al. 2008: 7). Similar also Cameron (2005) who notes that claims are expressions of identity attributes. Claims typically include information referring to the identity of an individual (e.g., "my name is John Doe"), but it can be sufficient to have information that qualifies a

[14] Automated Teller Machine

person for a specific action such as a certain age ("I am over 18 years old") may allow to cast a vote in official elections, or to drive a car etc. In practice, identity verification and authentication is often combined. For example, border control verifies the identity of person as well as the (implicit) claim to have a certain nationality etc. This process of verifying a claim, i.e., that particular information referring to an entity is valid in a particular context is authentication (cf. Rundle et al. 2008; ITU 2010).

Identification can have different qualities and a basic distinction can be made between "hard" and "soft" identification: the former means the explicit processing of identity information to prove the identity of a particular person and know exactly who that person is, (e.g., by providing an ID document, ID card etc.). The latter, soft identification means that a person provides some information related to her but he or she is not requested to proof her real identity. Hard identification is usually requested for transactions with a legal relevance where a public or private authority (e.g., a company, a government institution) is involved. This form is the main issue, e.g., in official procedures of government and public administration as well as commercial transactions (such as requesting for a public service or conducting a financial transaction). Soft identification occurs, for example, in online services, social media and similar applications that typically require a registered user account or profile but not necessarily proof of one's real identity. In practice, the boundary between these types is mostly fluid. Given the growing amount of identifiable information being processed in digital environments and socio-technical systems the room of possibilities for hard identification is likely to expand (the emergence of digital identification is explored in Section 4). In the same application context, many different forms of identifiable information can be processed (e.g., in social media platforms). Furthermore, soft identification can smoothly become hard identification due to the aggregation or linkage of different information sets related to a person (e.g., by linking a person's username with her date of birth and address, or further information). A further aspect concerns the modality of identification. In general, an individual can be confronted with identification in different ways, e.g.: 1) as voluntary act based on an informed decision or as an act of an accepted social norm, i.e., the person reveals her identity or provides personal information (e.g., in a social interaction) because she wants to or finds it appropriate; 2) as a mandatory, intentional act based on a legal requirement (e.g., providing a passport at a national border, requesting a public service, making a contract etc.); 3) as a mandatory act without choice or with negative consequences, e.g., a person is enforced to reveal her identity or becomes identified by law enforcement or other authorities; 4) as non-mandatory act but with limited choice and thus a quasi-obligation, e.g., the provision of personal information is requested to access a service; 5) as an involuntary and unknown process that happens without the consent and knowledge of the concerned individual, such as an illegal action or the further processing of identity information by third parties. Each mode can be problematic in terms of privacy protection but the latter, i.e., being identified without even noticing it, is particularly critical as it precludes the concerned individual from taking action against unwanted identification. Moreover, technology can lead to various overlaps between these modes. Indeed, there are numerous legitimate purposes of identification where the individual is aware of being identified in a particular context for a particular purpose (such as a traveller crossing a border proofs his identity with a passport,

an applicant provides personal information to gain entitlement to benefits, or a driver's licence is requested when buying a car). However, identification is neither adequate nor necessary in every context. As revealing one's identity can be a sensitive process, anonymity used to be the standard case in most contexts of everyday life. Technology usage challenges anonymity in many respects.

Identification is commonly expected to require particular, aware action of the individual concerned (such as showing an ID or providing identity information). However, as outlined, this is not necessarily the case as a person can also be identified without her direct involvement. An example for direct identification is where the person is requested to enter her username and password; an example for indirect identification is where identity information is gathered and processed automatically (e.g., by being observed through a video camera, or being identified based on technology usage, such as via an IP address in a computer network). Thus, as regards the knowledge or awareness of the identified individual, a further distinction can be made between explicit and implicit identification: in cases of the former, a person is, e.g., prompted to provide identifiable information and is aware of that; in cases of the latter, a person is not aware of being identified or identifiable by information concerning her. Profiling and surveillance activities are particular examples for the latter which are dealt with in Chapter 5.

3.1.3 Anonymity, pseudonymity and identifiability

Today, identification procedures are widespread in many different contexts; and, as will be shown, there are several tendencies for a further growth in different forms of identification. Reasons for this growth are manifold. For instance, Clarke (1994a) claims that institutions growing in size and structure, decreasing trust between individuals and organisations, as well as increasing forms of long-term economic relationships (i.e., customer loyalty respectively customer-relationship management) contribute to increasing identification. Clarke argues that these developments stimulated the presumption among many organisations, that identifying a person is necessary in most cases to conduct transactions (ibid). From a wider perspective, considering the connecting function of identification, globalisation and networking structures (cf. Castells 2000) are likely to foster identification as an increasing amount of entities interact with each other. Particularly in distant communications and interactions, a certain demand to identify individuals as well as institutions is plausible, which is one reason for the increase in digital identification (as explored further in Chapter 4).

Nevertheless, there are a many contexts in everyday life, where identification and knowledge about an individuals' real identity is not needed at all. Therefore, being anonymous (and thus not identified) used to be a standard mode in many societal contexts. Anonymity is particularly crucial for the protection of privacy and other fundamental human rights such as freedom of expression. Put simply, anonymity means to be not identifiable. Or in other words: anonymity is the absence of identifiability. It is a common part of every-day-life, and in many cases, individuals usually remain anonymous in pursuing their activities in their private sphere as well as in the public sphere. Anonymity is an essential concept for the functioning of democracy: secrecy of the ballot and

anonymous elections ensure that citizens can decide freely and anonymously who to give their vote; sensitive professional groups such as journalists, medical doctors, lawyers, researchers, diplomats, politicians, police officers, security agents etc. all have certain situations where anonymity is crucial for exercising their professions; insiders, informants and whistle blowers need some degree of anonymity to inform the public about eventual scandals, human rights abuses, illegal actions etc.; and without anonymity, individuals living in authoritarian regimes are in permanent danger to be persecuted. In case of criminal activity, anonymity can also be problematic, which is one reason why law enforcement and security authorities aim at extending security and surveillance measures. Indeed, there are many plausible reasons for hard or formal identification as well. But leaving these issues aside, from a privacy perspective, anonymity is fundamental. Moreover, anonymity does not necessarily imply the absence of any form of authentication. In fact, authentication is basically possible without identification, also in digital environments (cf. Chaum 1985; Clarke 1994a; Pfitzmann/Hansen 2010). In many transactions it is sufficient to (implicitly or explicitly) have plausibility about the individual's qualification or ability to, e.g., conduct a transaction, request a service etc. without need to reveal one's real identity. For instance, simple commercial transactions such as the buying of a good or service may not require any knowledge about an individual as long as a good is delivered and paid in exchange. Therefore, anonymity in typical commercial transactions with cash payment is usually unproblematic. Some transactions may require age verification (the acquisition of alcohol), but also in these cases it is not per se necessary to know exactly, who a person is. Depending on the context, it may be sufficient to know, e.g., that a person is not underage.

Basically, the real identity of a person is often less relevant in a transaction or interaction than a specific attribute. However, in practice, identification and authentication are mostly combined and often difficult to distinguish. As a consequence, more identity information than may be necessary in a particular context is being processed (cf. Chaum 1985; Nauman/Hobgen 2009; Pfitzmann/Hansen 2010; Strauß 2011; De Andrade et al. 2013; Danezis et al. 2014). This can have many different (intended as well as unintended) reasons. Leaving these reasons aside here, an important factor why face-to-face transactions may be easier to conduct anonymously than digital transactions concerns uncertainty. As outlined, identification and authentication are mostly used to reduce a certain degree of uncertainty. In economic terms this is particularly relevant to reduce eventual risks when, e.g., a transaction does not succeed as intended (for example when a client does not pay etc.). This is also a matter of trust and confidence among the interacting parties. Not knowing a transaction partner can trigger uncertainty, e.g., about who is responsible in case of failure, fraud etc. Transactions, where the exchange of goods or services and payment are conducted instantly (e.g., paying cash in a store), may not require identification as there is usually no uncertainty in this regard. Put simply, without payment, the good is not handed over. The instant character of a transaction can be supportive to reduce uncertainty about its successful operation. Hence, space and time of a transaction (or more generally of an interaction) may affect the need for identification among the interacting parties. For instance, in face-to-face transactions, identification happens implicitly (e.g., by recognizing a person's appearance) and (if not recorded) identifiable

information diminishes at the end of the transaction. A further aspect is that this setting is usually less complex in an analogue environment where no technology is involved, compared to a digital environment, where one or more additional entities are included.

In digital environments, the achievement of anonymity is non-trivial. In a technical sense, anonymity can be seen as a "situation where an entity cannot be identified within a set of entities" ITU (2010: 2). To enable anonymity of an individual entity compared to others requires that the identity attributes of this entity are non-distinct from others (Pfitzmann/Hansen 2010). Information is anonymous when all items are removed which could identify the person concerned and none of the remaining information is sufficient to re-identity that person (cf. Pfitzmann/Hansen 2010; FRA 2014). Anonymous information is not equivalent to pseudonymous information which merely means that an identifiable part of an information set is replaced by another type of information (e.g., a name being replaced by a synonym). For the provision of anonymity, the concepts of unlinkability and pseudonymity are highly relevant. Unlinkability means that "the exchange of identity information should not allow a linking of information from different transactions, unless a user specifically wishes to do so" (Rundle et al. 2008: 34). It is a crucial requirement to avoid that identity information is linked and aggregated across different contexts (cf. Strauß 2011). Pseudonymity means that a piece of information that does not directly link to one's identity is used for identification (e.g., an alias or a number that changes in every usage context) (cf. Chaum 1985; Pfitzmann/Hansen 2010). Pfitzmann and Hansen (2010) distinguish five forms of pseudonymity: transaction pseudonyms enable the highest level of unlinkability and thus strong anonymity. Each transaction uses a new pseudonym, which is only applied for a specific context[15]. A person pseudonym, i.e., a substitute for the identity of the holder (e.g., a unique number of an ID card, phone number or nickname) provides the lowest anonymity level. Moderate linkability is given by role and relationship pseudonyms, which are either limited to specific roles (e.g., client) or differ for each communication partner. Figure 7 illustrates these different types:

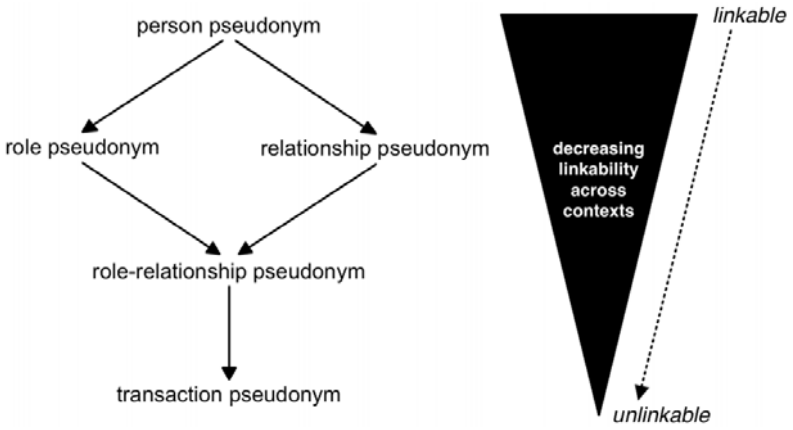


Figure 7: Levels of pseudonyms and unlinkability
(adapted from Pfitzmann/Hansen 2010: 27)

[15] For example, transaction authentication number (TAN) method for online banking.

The use of pseudonyms in different contexts is a means to establish an intended degree of (un)linkability. This also implies the avoidance of (global) unique identifiers that are valid in multiple contexts and then can be used to link and aggregate data across different contexts and thus used for privacy infringements such as data mining and profiling (cf. Hildebrandt/Gutwirth 2008; Hildebrandt/Rouvroy 2011).

Hence, unlinkability is an important principle to protect identity information and thus privacy (for an overview on technical examples in the field of privacy by design see Section 5.4). More precisely, it contributes to the boundary control function of privacy and informational self-determination, as described in Section 3.2. Unlinkability is supportive to avoid a concentration of informational power by avoiding information aggregation across separated domains. The mechanism is partially comparable to the separation of powers in democratic states, i.e., keeping different administrative domains detached from each other in order to inhibit a concentration of power and prevent from totalitarianism. For instance, domains such as health, education, tax, social security etc. have different data repositories so that information about the population is not stored in a centralized control unit. These domains are mostly not allowed to link their records with those of other domains in order to reduce the risk of mass surveillance and population control. However, unlinkability is difficult to achieve and undermined by the widespread and further proceeding pervasion of society with ICTs. They enabled a multitude of possibilities to collect and store information about individuals in public and private contexts which entails expanding, networked representations of digital identities, and increasing digital identification practices (see Section 4). There is a general increase in identifiability observable resulting from the extensive amounts of identity information available. Digitally networked environments entail multiple application contexts which foster the aggregation and cross-linkage of identity information, collected and processed by various institutions in the public as well as the private sector. This has serious implications for privacy (as explored more thoroughly in Chapter 5). The next section deals with the basic meaning and function of privacy which builds the fundament for the further analysis.

3.2 The role and meaning of privacy

Privacy is a more than a fundamental human right. It is a sine qua non for individual well-being as well as for the functioning of society. Although regulations on privacy and data protection are an achievement of modernity, its basic role and meaning can be found in ancient societies and in nature. Seclusion and differentiation from the individual in and between communities is important among humans and even observable among animals. Cultural and anthropological studies show that different aspects of privacy are and were part of different societies worldwide (cf. Westin 1967; Moore 1984). Eventual conflicts and balancing acts between one's private sphere and the public sphere can be described by Arthur Schopenhauer's metaphoric porcupine's (or hedgehog's) dilemma: porcupines (as well as hedgehogs) are reliant on their fellows spending warmth and affection. However, as they have spikes on their backs, they can hurt each other if they do not keep an adequate

distance from each other. This metaphor fits well to human society and the complex interrelations between social proximity and distance[16], privacy is involved in. Indeed, several scholars assume that privacy also has a biological meaning: even in the life of animals, e.g., in the form of seclusion or intimacy in small groups, privacy seems to play an important role (Westin 1967; Klopfer/Rubenstein 1977). Thus in short terms privacy is not "just" an invention of modernity but can be seen as a cultural universal (Westin 1967). As a multidimensional concept, privacy comprises different types and functions. Westin (1967: 31f.) identified four basic states of individual privacy: (1) solitude, meaning the physical separation of an individual from others; (2) intimacy, i.e., private relationships between two or more persons or distinct groups, (3) anonymity, understood as state of being unknown, and (4) reserve, i.e., the individuals' psychological barrier that protects her personality from unwanted intrusion. Clarke (2006) distinguishes between privacy of the person, of personal behaviour, of social communications and of personal data. Finn et al. (2013) suggest seven types of privacy based on Clarke's typology, adding privacy of thoughts and feelings, location and space as well as association.[17] Basically, privacy encompasses all domains of an individuals' life (e.g., personal, social, political, economic, cultural, psychological etc.) including her characteristics, racial or ethnical origins, health, preferences, desires, thoughts, religion, beliefs, opinions, behaviour, actions, communications, associations and relationships with others etc. (cf. Westin 1967; Clarke 2006; Solove 2006; Hildebrandt 2006; Rovroy/Poullet 2009; Nissenbaum 2010; Finn et al. 2013). Or in other words: privacy enfolds identity.

Protecting privacy involves safeguarding corresponding types of information from being processed without a legal basis or beyond the intention of the individuals concerned. The relationship between identity and privacy is also part of legal data protection frameworks. The forthcoming new EU data protection regulation defines personal data as "any information relating to an identified or identifiable natural person"[18], i.e., an individual "who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person" (Art. 4 (1) GDPR).

Another term for personal data is so-called "personally identifiable information" (PII), i.e., any related, directly or indirectly linked to a natural person, information usable to identify that person (ISO 2011: 2). PII is a common term in technical privacy frameworks and basically meant in this research when personal data/information, identity (or identifiable) information is mentioned.[19] Hence, in short, personal data is information related to an identified or identifiable person. This implies that direct as well as indirect identification affects privacy, in a legal but also in a broader, socio-technical sense.

[16] This is a common issue of psychoanalysis, also known as porcupine problem; see, e.g., Luepnitz(2003).
[17] For a more detailed discussion on the different types of privacy with respect to information processing see Section 6.2.
[18] This definition is widely similar to the Data Protection Directive 95/46/EC Art. 2 (DPD 1995).
[19] The preferred term in this research is identity or identifiable information because it is broader and allows considering technical information related to a personal identity as well. This makes an important difference as will be shown Sections 5 and 6.

3.2.1 Overview on legal issues and basic privacy protection principles

As privacy affects all domains of an individuals' life, its significance as a fundamental human right is obvious. Thus, most countries worldwide have specific laws for privacy and data protection. Irrespective of the national peculiarities, privacy is a fundamental human right since 1948: Article 12[20] of the Universal Declaration of Human Rights (UDHR) states that "[n]o one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks." Similarly, in the European Convention on Human Rights (ECHR) Article 8 is dedicated to privacy:

1. Everyone has the right to respect for his private and family life, his home and his correspondence. 2. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

In a similar vein, the Charter of Fundamental Rights of the European Union (EUCFR)[21] includes the right to privacy and the protection of personal data (in Articles 7 and 8). Democracies worldwide have directly or indirectly incorporated similar regulations in their rules of law. In Europe, the ECHR and the EUCFR represent general reference points which are incorporated in the legal frameworks at EU as well as national level. In European legal contexts, the protection of privacy is mostly described as data protection. However, from an individual human rights perspective, essentially, both strands address the same issues.[22] Therefore, privacy is the preferred term in this research as it is broader, includes data protection and implies that protection is not limited to data. The prevalence of the term data protection has some historical reasons: as the necessity to create particular legal frameworks emerged from the growing importance of electronic data processing since the 1960s and 1970s. Many privacy frameworks emerged during the 1970s and 1980s. The first countries worldwide to create privacy laws were Germany (initially in the state of Hessen), Sweden and France during in the early 1970s. In the United States, the first legal framework was the Privacy Act of 1974. Germany is particularly relevant as its privacy regime enjoys a good international reputation and had impact on the emergence of the European privacy framework as well (González-Fuster 2014). In 1980, the Organisation for Economic Cooperation and Development (OECD) published guidelines "on the Protection of Privacy and Transborder Flows of Personal Data" (last updated in 2013) which are still today of international relevance (OECD 2013b). One year later, with the so-called Convention 108 of the Council of Europe, the first international privacy treaty, currently ratified by over 40 countries worldwide, was created. Convention 108, entitled "Convention for the protection of individuals with regard to automatic processing of personal data", contains general principles for the proper processing of personal data with respect for the rights of individuals. Being a treaty with multiple nations involved, it

[20] Similar is Art. 17 of the UN Covenant of Civil and Political Rights, see http://www.hrweb.org/legal/cpr.html
[21] CFREU (2000)
[22] The differences in particular are more relevant in a legal sense which is not the main focus of this research.

is yet the only international data protection instrument with a legally binding effect (FRA 2014). In 1983, the right to informational self-determination was created in Germany which represents a landmark ruling for the European privacy regime (for more details about its relevance see Section 3.2.3). The central regulation of the European Union yet is the Data Protection Directive 95/46/EC (DPD 1995), being the first legal instrument of the European community. It was adopted in 1995 as the first privacy regulation valid for all member states aiming at creating a common legal baseline for privacy laws in Europe. It also substantiated and expanded some of the norms of Convention 108 (FRA 2014; González-Fuster 2014). The DPD is currently the main instrument of the EU data protection framework, until 2018, when the new general regulation becomes effective (see below). It contains mandatory minimum standards for all member states, which had to implement the regulation into their national laws. In addition to the DPD, there are some further laws such as the ePrivacy Directive 2002/58/EC complementing the DPD. It inter alia regulates data processing in the context of publicly available electronic communications networks (FRA 2014). The ePrivacy Directive was controversial from its beginnings among privacy experts due to its relatively narrow scope and the insufficient regulation of transborder data flows. Given the widespread use of global communications networks, developments such as pervasive computing, the Internet of Things etc. its applicability is limited. Therefore significant updates were highly recommended (e.g., recently by the European Data Protection Supervisor, EDPS 2016a). This law is currently under revision with respect to the EU privacy reform.[23] In May 2018, the recently enacted General Data Protection Regulation (GDPR) becomes effective which replaces the DPD (GDPR 2016). The implementation of the GDPR is a crucial milestone of the privacy reform. It aims at further harmonizing and strengthening the privacy regime in Europe to cope with the technological challenges of globally available personal information. The GDPR is partially based on the DPD. However there are several changes and novelties such as high penalties for illegal data processing for enterprises, a stimulation of privacy by design, as well as the introduction of obligatory privacy impact assessments under certain conditions (for more details see Section 6, which is dedicated to privacy impact assessment). The creation of the GDPR is a crucial part of the still ongoing privacy reform in the European Union which has an international impact, particularly on the political and economic affairs between the EU and the US. For several years, there are attempts to create an international standard for privacy and data protection on the basis of Convention 108 (FRA 2014). Since 2016, there is an according revision in progress (cf. Greenleaf 2017). A further regulation regarding the transborder flow of personal information between the EU and the US is the so-called "Privacy Shield" adopted in 2016. This agreement is based on self-commitment and replaced its predecessor the "Safe Harbour decision" which was declared as invalid by the European Court of Justice in 2015. However, privacy experts criticized this regulation for being insufficient (WP29 2016). Moreover, according the media reports, current US president Donald Trump could revoke the privacy-shield, which makes its efficacy yet generally uncertain (McCarthy 2017a).

[23] European Commission: Proposal for an ePrivacy Regulation, January 2017, https://ec.europa.eu/digital-single-market/en/proposal-eprivacy-regulation

Besides particular legal regulations, there are some commonly accepted principles (also referred to as fair information practices – FIPS, particularly in the US) which represent the backbone of most privacy and data protection laws worldwide. Key principles are especially (cf. De Hert 2012; OECD 2013b; Danezis et al. 2014; FRA 2014):

(a) *Legality and lawful processing*: i.e., personal data is processed in accordance with the law, for a legitimate purpose, and the processing is necessary in a democratic society;
(b) *Purpose specification and purpose binding/limitation*: i.e., the purpose needs to be specified in advance of the processing and the use of personal data for other purposes requires a legal basis as well. Another purpose is particularly given in case of data transfer to third parties;
(c) *Data minimisation*: only those data necessary for a particular purpose should be processed and deleted when the purpose of processing ceases to exist;
(d) *Data quality*: data needs to be adequate, relevant and not excessively collected in relation to the purpose of processing, data should be accurate and kept up to date, limited use/retention so that data is not kept longer than needed;
(e) *Transparency and fair processing*: so that the individuals concerned can comprehend how their data is being processed, by whom and for what purposes;
(f) *Accountability*: i.e., the data processing entity have to act in compliance with the law and safeguard personal data in their activities.

These principles build a basic fundament of privacy regulation and are inter alia part of the EU Data Protection Directive 95/46/EC as well as the forthcoming GDPR. For instance, according to Article 5 GDPR, personal data must be: "(a) processed lawfully, fairly and in a transparent manner; (b) collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes[24] (…); (c) adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation'); (d) accurate and, where necessary, kept up to date; (…); (e) kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed (…)"

For technical implementation of safeguards, particularly as regards (b) and (c), the concept of unlinkability (as outlined in Section 3.1.3) is important. The ISO/IEC 29100:2011 privacy framework provides relevant specifications and guidelines, particularly as regards the handling of personal information (for more details on technical privacy protection and impact assessment, see Sections 5 and 6).

[24] The processing of data "for historical, statistical or scientific purposes shall not be considered as incompatible provided that Member States provide appropriate safeguards;" (Article 5 (1) (b) GDPR).

3.2.2 Private versus public spheres? – the boundary control function of privacy

Basically, privacy is located at the boundary between the public and the private sphere. Privacy constitutes the private sphere of an individual. However, it is not to be misinterpreted as a means where the individual protects her privacy by decoupling from the public. The framing of privacy as a means of decoupling and a concept of seclusion is a general issue which complicates to understand what the claim that privacy needs to be protected precisely means (cf. Solove 2006). For instance, confidentiality and protection of information from being unintendedly disclosed are indeed relevant issues of privacy. However, this does not imply that privacy is equivalent to secrecy or hiding something, although often misinterpreted in this regard (cf. Agre/Rotenberg 1998; Solove 2004/2006; Strauß 2017a). The framing as concept of seclusion is particularly given in the common, classical notion of privacy (or data protection) as "the right to be let alone" (Warren/Brandeis 1890). As will be shown, this notion is too narrow and misleadingly reduces privacy to a concept of seclusion and separation from others. In fact, privacy serves as a substantial enabler of both, personal development of an individual as well as individual involvement in community and society. Metaphorically speaking, the previously mentioned distance hedgehogs keep to each other, does not imply that they are unsocial and lack in community. In contrast to its common classical notion, privacy is not to be misunderstood as an individuals' claim and right to decouple from society. To the contrary, it is a societal achievement that relieves from different kinds of social frictions and "enables people to engage in worthwhile activities that they would otherwise find difficult or impossible" (Solove 2006: 484). Privacy is not an isolated concept of seclusion but it empowers other fundamental rights, values and freedoms such as expression, thought, movement, association etc. Hence, it is a vital prerequisite for democracy and political participation (cf. Westin 1967; ibid; Hildebrandt 2006; Nissenbaum 2010). Therefore privacy can be located "in its functional relationship to valued ends, including human well-being and development, creativity, autonomy, mental health, and liberty" (Nissenbaum 2010: 74f.). Privacy is a sine qua non for human well-being and development: to freely develop thoughts, ideas, intellectual capacity, creativity, artistic expressions, ethical and political judgements etc., individuals need spheres free from permanent scrutiny and "leeway to experiment without the distraction of being watched (…); free from pressure to conform to popular conventional standards" (ibid). Protecting privacy involves protection of the process of "becoming, being and remaining a person" (Reiman 1976: 44). Thus, privacy essentially preserves human dignity and individual identity development (Kahn 2003; Rouvroy/Poullet 2009; Nissenbaum 2010). In this regard, privacy is "(…) a unifying principle that is itself the means for creating the notions of dignity and identity (…)" (Kahn 2003: 410). Harms to privacy thus have negative impact on individual dignity and identity-building. This also has wider societal effects as privacy impairments "impede individual activities that contribute to the greater social good" (Solove 2006: 488). To protect this fundamental right is thus essential for the well-being of the individual as well as of society.

Hence, although occasionally presented as contradictory, the private sphere is not the opposite of the public sphere. To the contrary, both are intrinsically linked and a

prerequisite for each other. Several scholars underline that privacy is not merely a private but also a public value, vital for democratic processes (e.g. Habermas 1989; Kahn 2003; Hildebrandt 2006; Rouvroy/Poullet 2009; Nissenbaum 2010; Cohen 2012). Privacy directly serves the individual, but its functions and benefits are wider and translate into societal good (Nissenbaum 2010). Hence, there is a relation to the public sphere. The public sphere is an essential element of deliberative democracy that serves the function to intermediate between citizens and political decision makers (Habermas 1989). With its inherent deliberative quality it also provides a domain where public communication transforms into public opinion (ibid; Frazer 2007; Trenz 2008). But the public sphere is not to be (mis-)understood as a single, universal space of public deliberation and discourse. It is a "communicative network where different publics partially overlap" (Nanz 2007: 19). The formation of the public sphere and its deliberative quality are closely linked to private sphere(s), i.e., those domains and spaces where individuals enjoy their privacy and have the ability to be and act freely without interference from others. Thus, to some extent, the relationship between the private and the public sphere is complementary. In this regard, privacy also contributes to societal diversity as individuals are enabled in their specific personal development which can benefit the well-being of society. In the private sphere, individuals build their very opinions; as these opinions are communicated, shared, discussed, shaped etc. with others, the public sphere emerges and takes shape (cf. Habermas 1989). Therefore, both spheres are intrinsically linked and vital for democratic will-formation. Both need physical and virtual space to emerge and develop where individuals can interact, meet, exchange ideas and thoughts etc. freely without fear of intrusion, impairment or repression. Thus, limitations to privacy can also put the deliberative quality of the public sphere at stake.

Briefly speaking, "privacy defines a state where an individual is free from interference" (Strauß 2017a: 260). This means that privacy enables the individuals' being, actions, behaviour, relationships etc. to be free and without haphazard, uncontrollable intrusion or observation from external entities such as the state, public and private institutions or other individuals. But as outlined, enjoying privacy does imply that individuals seclude from others, or decouple from society and public engagement. Essentially, privacy has an inherent boundary control function. This function allows the individual to open as well as to close herself from others (Altman 1975; Hildebrandt 2006). For Westin (1967: 7), privacy is the result of a recurring process of negotiation and enforcement as "each individual is continually engaged in a personal adjustment process in which he balances the desire for privacy with the desire of disclosure and communication of himself to others, in light of the environmental and social norms set by the society in which he lives". As societal contexts are dynamic, this boundary control is a continuous process, where the individual shifts her boundaries in relation to her environment and the entities which may access information about her (cf. Hildebrandt 2006). In this regard, "privacy can best be understood as the virtual and actual space needed to continuously reconstruct the self in the face of ever changing contexts" (ibid: 44). Hence, the boundary control function of privacy is naturally dynamic as the individual is in permanent interaction with her environment. The subject of control mainly concerns information about the individual whereas (in an ideal case), she selectively controls whom to provide

personal insights. This is in accordance with Westin's (1967: 7) definition of privacy as "the claim of individuals, groups or institutions to determine for themselves, when, how, and to what extent information is communicated to others." Hence, information plays as essential role for privacy protection and is the major determinant of the boundary control function of privacy. Accordingly, Agre and Rotenberg (1998: 7) highlight the relationship of privacy with identity and state that "control over personal information is control over an aspect of the identity one projects to the world, and the right to privacy is the freedom from unreasonable constraints on the construction of one's own identity." Comprehending the interrelations between privacy and identity is also important to overcome the misleading reductionist view on privacy as means of seclusion or secrecy (discussed in Section 3.3.3). Correspondingly, Hildebrandt (2006: 50) points out "the core of privacy is to be found in the idea of identity (...) because the process of identity-building is what is at stake in privacy". Floridi (2013: 243) even speaks of privacy breaches as "a form of aggression towards one's personal identity". Kahn (2003) points out that privacy invasion can cause harm to dignity and destabilize the integrity of an individuals' identity.

3.2.3 The interplay between privacy, identity and autonomy

The presented boundary control function involves the complex interplay of privacy, identity and autonomy. As shown, privacy is an intermediate value for other human rights and an essential means for identity-building which fosters liberty, autonomy and self-determination (cf. Rovroy/Poullet 2009; Nissenbaum 2010; Cohen 2012). Privacy enables the formation of a particular space or sphere (different from the public sphere), i.e., the individual's private sphere, in which identity development, autonomy and self-determined living are enabled free from interference. In this respect, privacy, identity and autonomy build a triad of intrinsically interrelated concepts, as illustrated in Figure 8 autonomy presupposes identity, because without identity there is no entity that can act autonomously, and privacy can be seen as a constitutive framework for autonomy and identity development.

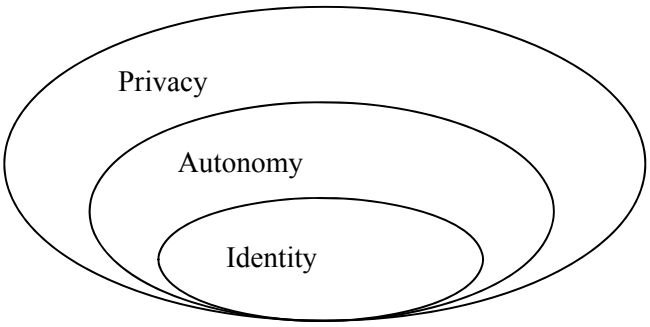


Figure 8: Privacy as room for autonomy and identity development

As outlined in Section 3.1, the identity of an individual person represents herself distinct from others. At the same time, identity is continuously shaped by its interactions with other entities and society. In this regard, identity can be seen as a construct which enables linkage between the private and the public sphere. In line with Ricoeur's notion of

enduring as well as dynamic features of identity (idem and ipse), privacy "allows a person to hold together while changing; it presumes some measure of autonomy, (…) intimacy and some space to rebuild the self in accordance with one's past while anticipating one's future" (Hildebrandt 2006: 52). Hence, identity-building is dynamic and does follow a particular order. Consequently, the "construction of the identity of the self implies indeterminacy, and privacy therefore implies the recognition of this indeterminacy" (ibid). This recognition refers to autonomy and self-determination which are essential for privacy and vice versa. Autonomy requires some domain in which it is factual, and where individuals can act self-determined and free from interference. This domain is represented by one's private sphere, in a virtual as well as in a spatial sense. Hence, the protection of privacy implies to protect autonomy and its function to act on free will. Consequently, the intrusion into the privacy of an individual also affects her ability to exercise autonomy (e.g. Treiger-Bar-Am 2008; Rovroy/Poullet 2009; Nissenbaum 2010). As a form of autonomy, Nissenbaum (2010: 81) describes privacy as "self-determination with respect to information about oneself". Put simply, autonomy means self-rule. The term consists of the Greek words for "self" (autos) and "rule" or law (nomos) and thus literally means to have or make one's own (moral) laws. Autonomy may be seen in relation to processes of self-organization inherent to autopoietic systems as human beings represent. Being autonomous implies to take free decisions. Autonomy is an essential concept in moral and political philosophy that plays a crucial role in law and culture, strongly coined by Immanuel Kant or John Stuart Mill (cf. Treiger-Bar-Am 2008; Christman 2015). According to Kant, autonomy is the ability to freely and self-determinedly establish moral laws and act upon them, closely linked to the freedom of the will to choose (Kant 1785/1997). In this regard, autonomy is the capacity to self-governance with a state of self-determination that empowers the individual to act freely without external intervention. Hence, it is a core element of personal well-being. But it is also the source of (moral and non-moral) obligations as it is "the capacity to impose upon ourselves, by virtue of our practical identities (…)" (Christman 2015). Hence, autonomy is also a condition for reason, a condition for humans to act reasonable "on rational principles and freely to exercise the moral reasoning will, through the freedom of choice" (Treiger-Bar-Am 2008: 555). Being autonomous means to be a self-determined person, governed by her intrinsic values, beliefs, desires, reflections, conditions, characteristics etc. which are not purely imposed upon her from external sources and thus can be seen as part of her authentic self (Christman 2015). Certainly, autonomy is an ideal and as such is not always completely obtainable in every societal context as regulated by the rule of law. Citizens enjoy rights but are also obliged with certain duties. Hence, in socio-political contexts, autonomy is often provided only to a certain degree under specific conditions. However, irrespective of eventual limiting conditions, the provision of liberty and autonomy is a sine qua non for a democratic society. Democratic processes require autonomy, may it be to freely express one's opinion, cast a vote without being enforced or constrained, or counteract against oppression and eventual injustices in society (e.g., by the rights of civil disobedience, demonstrations, freedom of expression etc.). Though closely related, autonomy is not equal to but a consequence of freedom (and vice versa). Liberty, i.e., political freedom, is "a prerequisite to the exercise of autonomy" (Treiger-Bar-Am 2008: 557). The idea of

autonomy articulates freedom and obligation at one haul: through the form of self-legislated laws (Khurana 2013). Autonomy overlaps with positive freedom, i.e., the capability to act on free will (Christman 2015). An intrinsic part of autonomy is also the capability to self-reflection, i.e., (re-)assessing one's desires, reasons and actions etc. in relation to universal duties and obligation. Briefly speaking, autonomy concerns an individuals' capability to rethink and reflect upon her moral identity. Thus, autonomy also plays a crucial role for justice: "it serves as the model of the person whose perspective is used to formulate and justify political principles, as in social contract models of principles of justice. (…) [C]orrespondingly, it serves as the model of the citizen whose basic interests are reflected in those principles, such as in the claim that basic liberties, opportunities, and other primary goods are fundamental to flourishing lives no matter what moral commitments, life plans, or other particulars of the person might obtain" (Christman 2015). Autonomy implies freedom from interference, but it is also relational as being autonomous requires a point of reference. This implies recognition or acceptance of autonomous actions by others, which is a necessary but not a sufficient condition of autonomy as well as of personal identity. Both, autonomy and identity are based on mutual recognition and acceptance between different individuals: as freedom implies to be free from something, a person cannot be free in her actions, i.e., act autonomously merely by herself without a referent object; the same is given for a personal identity being distinct from others, who are its referent objects. Personal identity is based on autonomy and its capacity to constitute for oneself a type of action that is recognizable to others. (Treiger-Bar-Am 2008; Christman 2015).

As shown in the previous section, privacy does not imply that the individual is completely decoupled from its environment in her private sphere. Private and public spheres complement each other and privacy has an inherent boundary control function to regulate the interactions between individual identities and society (perceivable as multiagent system). Identity here serves multiple functions: as its fundamental core, it constitutes the private sphere of a particular individual, because individual privacy has no meaning without her identity as prerequisite determinant. In this regard, it also predetermines the boundary between the private and the public sphere, even though this boundary is dynamic and changeable. At the same time, identity enables interactions through these spheres as individuals (represented by their identities) interact with other entities in its societal environment. The boundary privacy enables to regulate can be understood as an informational one. According to Floridi (2013: 232), privacy fulfils "the function of the ontological friction in the infosphere". This means that privacy provides an environment for the free flow of personal information, whereas this information flow is (ideally) not disclosed or accessible to other entities unless intended to be by the individual concerned. Thus privacy supports the individual in avoiding that her information flows seamlessly from one context to another without her decision to information disclosure. Hence, privacy facilitates autonomy, which here refers to informational self-determination. This includes determining what kind of information about one's private life and parts of one's identity are disclosed or shared with others. In this regard, privacy allows the individual to partially regulate social interactions and the level of transparency in relation

to its societal environment. Without privacy, this ability to self-regulation and self-determination would be significantly constrained if not repealed.

The form of autonomy intrinsic to privacy is informational self-determination (ISD). ISD is thus a core concept of privacy protection. It involves "connections between privacy and self-representation, and (…) identity formation" (Nissenbaum 2010: 81). Or in other words: privacy also concerns the individuals' self-determined maintenance and performance of her personal identity. ISD defines a state in which the individual that is affected by the processing of her information is aware of this process and enabled in controlling how and what personal information is involved. There are two major aspects of ISD: *knowledge* about the context of information processing and *control* over that context, i.e., over the personal information flows (Strauß/Nentwich 2013). Thus, in order to make an informed decision about the disclosure and use of her personal information, the individual needs to know, e.g., what personal information is collected, stored and processed for what purpose, and by whom. This adds the issue of *transparency* of the information processing as intrinsic part for ISD.

ISD is an integrative part of the European privacy and data protection regime. The legal discourse in Germany on ISD is an important case in order to understand the relevance of this concept. In German law, informational self-determination is explicitly defined as a legal right since the 1983. In his famous decision on the population census, the German Federal Constitutional Court[25] ("Bundesverfassungsgericht") highlighted that the individual needs to "be protected from unlimited collection, storage, use, and transmission of personal data as a condition of development of his or her free personality under the modern conditions of data processing". (BVerfGE 65 E 40, cited from De Hert 2008: 74; see also De Hert 2008; Rouvroy/Poullet 2009). The major reason for the Court's fundamental decision[26] was a census of the whole German population planned for 1983 which triggered a number of protests and constitutional complaints which enabled the Court to examine the law regulating the census ("Volkszählungsgesetz"). The Court found that the law is not in accordance with the German constitution and created the right to ISD (cf. Hornung/Schnabel 2009). This right to ISD was articulated by the Court as "the authority of the individual to decide himself, on the basis of the idea of self-determination, when and within what limits information about his private life should be communicated to others" (Rouvroy/Poullet 2009: 45).[27] The intention of the right to ISD is to "enable citizens to freely develop their personality" (Hornung/Schnabel 2009: 86). Following De Hert's (2008: 75) interpretation, the Court explained with precision "the shift of power that takes places whenever the state or private actors interact with an individual through ICTs". The reasoning of the Court highlighted that "a person's knowledge that his or her actions are being watched inevitably curtails his or her freedom to act" (ibid). Indeed, the decision points out that against the background of technical means that allow unlimited collection, storage and processing of information concerning individual persons, ISD needs particular protection; especially in the light of advances in autonomic data processing (cf. Rouvroy/Poullet 2009). Furthermore, the Court also explained that with the possibility to

[25] Judgment of the First Senate from 15 December 1983, 1 BvR 209/83 et al. – Population Census, BVerfGE 65, 1.
[26] In Germany known as the so-called "Volkszählungsurteil".
[27] This is similar to Westin's (1967) definition of privacy as outlined in the previous Section.

create "a partial or virtually complete personality profile" by aggregating or linking data from different sources, the persons controlled have "no sufficient means of controlling its truth and application" (ibid: 53). The Court argued further, that technological means reinforce the possibilities of influencing individuals' behaviour to an unknown degree so that public interests can exert psychological pressure on individuals. Irrespective of the "certain conditions of modern information processing technology", an essential precondition of informational self-determination is thus "that the individuals left with the freedom of decision about actions to be taken or omitted, including the possibility to follow that decision in practice. (…) If someone is uncertain whether deviant behaviour is noted down and stored permanent as information, or is applied or passed, he will try not to attract attention by such behaviour. If he reckons that participation in an assembly or a citizens' initiative will be registered officially and that personal risks might result from it, he may possibly renounce the exercise of his respective rights. This would not only impact his chances of development but would also impact the common good ("Gemeinwohl"), because self-determination is an elementary functional condition of a free democratic society based on its citizen's capacity to act and to cooperate" (excerpt from the Court's decision; cited from Rouvroy/Poullet 2009: 53). In order to reduce the risk of profiling and compilation of all personal information of an individual, the Court's ruling also included the prohibition of "the introduction of a unique personal identifier for every citizen" (Hornung/Schnabel 2009: 87). This is particularly crucial as the creation of "a complete personality profile would violate the guarantee to have one's dignity recognised" (ibid).

The Court's decision and the introduction of the right to ISD was a milestone for privacy protection. It had high impact on data protection legislation in Europe and represents a cornerstone of contemporary privacy protection frameworks (cf. De Hert 2008; Hornung/Schnabel 2009; Rouvroy/Poullet 2009). According to (Hornung/Schnabel 2009: 85), the reasoning of the Constitutional Court is partially based on social systems theory, whereas fundamental rights "have the function of guarding the differentiation of society into sub-systems". Privacy and ISD play the essential role "to protect the consistency of the individuality of the individual, and consistent self-expressions rely heavily on the separation of societal sub-systems" (ibid). This separation means that individuals can enjoy a free and self-determined life which is essential for the functioning of a free and democratic society. In this regard, ISD relates to the boundary control function of privacy and its double effect: "[t]he individual is shielded from interference in personal matters, thus creating a sphere in which he or she can feel safe from any interference. At the same time, data protection is also a precondition for citizens' unbiased participation" in political processes. (ibid: 86). A further legal issue in German privacy law, related to ISD is the basic right to confidentiality and integrity of information systems (basic IT right "IT-Grundrecht") that complements the right to privacy and ISD. It aims at specifying the "Volkszählungsurteil" as it explicitly addresses also personal use of ICTs that are to be protected from privacy infringement. This new right was created in 2008 because of a broader societal debate in Germany during that time about the use of online surveillance tools by security authorities to monitor personal computers (so-called "Online-Durchsuchungen"). The court argued that information systems of whatever kind (laptops, mobile phones etc.) enable to gather insight into significant parts of the life of a person or

even to draw a detailed picture of one's personality (De Hert 2008; Wehage 2013). For De Hert (2008: 75) this new fundamental law represents a "landmark ruling, that recognises a citizen's right to the integrity of their information-technology systems and introduces elements of user-centric identity management". Interestingly, this new regulation seems to play a marginal role so far in socio-technical practices. Reasons for this may lie in the relative novelty of the regulation and lack of reference points about its practical applicability (Baum et al. 2013; Wehage 2013). Irrespective of its legal meaning, the pervasiveness of ICTs and their entailed socio-technical transformations essentially strain and undermine ISD: in digital environments, personal information is often processed without the knowledge and control of the individual concerned. While the processing of her information is opaque to the individual, her identity being digitally represented can be highly transparent beyond her control. To some extent, ISD is exposed to controversial issues which are seemingly contradictory to privacy protection, as discussed in the next sections.

3.3 General privacy controversies

As outlined, the emergence of national and international privacy regimes is strongly influenced by technological progress. Basically, novel technologies and their usage practices frequently put privacy protection under pressure. The growing importance of electronic data processing during the 1960s and 1970s had some ignition effect for the creation of privacy regulation on global level. Since then, the rapid diffusion of ICTs (particularly between the 1990s until today) significantly boosted privacy challenges. To cope with the changed socio-technical requirements for privacy protection, law and policy makers try to adopt regulation accordingly. The European privacy reform is an important development in this regard. Nevertheless, several privacy problems aggravate with altering socio-technical practices beyond the scope of the law. The enormous complexity of international policies and regulations hampers the development and enforcement of international standards. A general crux of privacy and data protection is the highly complex legal situation. Besides issues resulting from privacy intrusions serving various indefinite security purposes, a complicated aspect is, for instance, the so-called "informed consent" of the person concerned of data processing: when a person accepted the processing by giving consent, processing mostly receives a legal basis. Article 7 of the expiring EU Data Protection Directive (DPD) determines that "personal data may be processed only if (a) the data subject has unambiguously given his consent or (b) processing is necessary for the performance of a contract to which the data subject is party (…)". The forthcoming GDPR regulates informed consent basically in the same manner, whereas processing is lawful if "the data subject has given consent to the processing of his or her personal data" (Art. 6 GDPR). Consent is an important and necessary condition for personal data processing with the aim to protect from abuse. Hence, the idea is that information should not be processed against the will of the individual concerned. However, in practice, there are several difficulties. Giving consent can be especially problematic

when an individual wants or needs to use a particular service, has no other option than to accept the processing of their data and cannot control its usage (cf. Nissenbaum 2010; De Andrade et al. 2013; EGE 2014). Informed consent is mostly a necessary condition for using an online service, application, or a technology, e.g., by accepting the terms of use. This is in most cases sufficient to allow for, e.g., commercial exploitation of personal information, including third party access to the data. As digital data processing mostly includes data flows between multiple entities beyond national borders and legal frameworks, appropriate regulation and privacy protection is increasingly challenging. Consequently, individuals lack control over their information and thus ISD, as will be shown in the further Sections of this research (issues related to informed consent are discussed in Section 5.4.3).

Given the strong linkage between technology, personal identity and identification (see Section 4), privacy protection and ISD are increasingly hampered. Digital identities are often free-floating in socio-technical contexts and their holders exposed to risks of misuse without knowledge or control thereof. This is aggravated by the fact that threats to privacy do not necessarily result from abuse of legal rights. For instance, profiling activities of various kinds can gather information related to individual identities from various sources. This may be acceptable in a strict legal sense as individual users have accepted the terms of use, e.g., of an online service where third party access to data, such as for further commercial purposes is included (these issues are discussed more in-depth in Sections 5). Although this is no privacy abuse by law, profiling activities intrude into privacy and affect identity-building (cf. Hildebrandt 2006; Hildebrandt/Gutwirth 2008). A further issue concerns security and surveillance practices based on the exploitation of personal information for purposes of law enforcement and national security. Interfering with privacy is foreseen by the law under certain conditions: i.e., to fulfil public interest in accordance with the law and the protection of the foundations of a democratic society (cf. FRA 2014; EGE 2014; Strauß 2017a). As security is also a basic human right and the state aims to protect is integrity, privacy intrusions and surveillance activities are frequently justified by security purposes. Hence, also in these cases, privacy may not be violated in a legal sense. Nevertheless, profiling and surveillance mostly are privacy-intrusive activities entailing serious risks. Or as De Hert (2008: 73) put it: "Lawful collection and processing of personal data does not prevent per se unethical practices deployed in the name of security, or unjust decisions made on them." Hence, while to some extent, ICTs enable new forms of privacy intrusion, they particularly reinforce pre-existing practices, which apply technology in a privacy-intrusive manner (Chapter 5 explores these issues more in-depth). It is thus important to shed light on controversies of privacy resulting from concepts that are in conflict with privacy. In the following, two important issues in this regard are discussed: the relation between privacy and security as well as between privacy and transparency. The former is frequently framed as the main reason to justify privacy intrusion and surveillance practices. The latter is linked to notions of post-privacy.

3.3.1 Privacy and security – a contradiction in terms?[28]

In the public discourse, privacy is often presented as a sort of counterpart to security, especially, when political interests strive for an extension of security measures. Indeed, privacy intrusions are foreseen by the law under certain conditions, i.e., for law enforcement to protect the interest of the public under the premise of protecting issues vital for a democratic society. However, this does not imply that there is a state of permanent conflict between privacy and security. Above all, these limitations are the exception from the rule while human rights catalogues and legal frameworks suggest that privacy protection is the standard mode (Strauß 2017a). However, this quasi-standard setting is strained by a number of issues, reinforced by the assumption of a permanent trade-off with an inherent conflict between privacy and security.

In general, the term security stems from the Latin word "securus" which consists of "sine" (without) and "cura" (concern, worry, or problem). Hence, generally speaking, security addresses as a state without a need to worry or be cautious. This already indicates that security has a subjective meaning as the perception of what is a security threat may differ from individual to individual. Thus, security is subjective and relative. Consequently, objective and subjective notions of security are most often not equal. The former rather "refers to the low probability of damage" and the latter is "the feeling of security, or the absence of fear that acquired values are threatened" (Chandler 2009: 123). However, both are closely related and can influence each other. Security is not an absolute concept that can be permanently assured but it is depended from its environmental conditions that cannot be fully controlled. In other words: there is no state of absolute, one hundred per cent security achievable as security is naturally limited and dependent from the conditions of its referent object. Many security scholars therefore point out that security is an ambiguous and contested concept (Buzan et al. 1998; Balzacq 2005; Guild et al. 2008; Bigo 2008; Watson 2011). The relationship between the individual and the state plays an important role in the security discourse. Since the 18th century, there is a long tradition inter alia based on Hobbes' contribution to state philosophy[29], where security is seen as a responsibility of the sovereign (UN 1994; Owen 2004). Security is also part of human rights catalogues. The Universal Declaration of Human Rights (UDHR) (resolved in 1948) states in Article 3 that "Everyone has the right to life, liberty and security of person." Social security is addressed in Article 22 UDHR: "Everyone, as a member of society, has the right to social security and is entitled to realisation, through national effort and international co-operation and in accordance with the organisation and resources of each state, of the economic, social and cultural rights indispensable for his dignity and the free development of his personality." These legal terms already indicate that liberty, security and personality development or identity are interwoven and at the same side of the coin.

While the provision of public or national security was in the main responsibility of the state, the traditional framing of security also put the main focus on protecting the integrity

[28] Parts of this section refer to Strauß (2017a).
[29] Referring to Thomas Hobbes' major work "The Leviathan" first published in 1651.

of the state from different kinds of threats (such as conflicts, wars, nuclear proliferation). International tensions contributed to strengthen this focus also as regards military force. Traditional state-centred security was the dominating concept that reached a peak during the Cold War, as Owen (2004: 16) points out: "This type of security relied primarily on an anarchistic balance of power (power as the sole controlling mechanism), the military build-up of two superpowers, and on the absolute sovereignty of the nation-state. (…) Security was seen as protection from invading armies; protection was provided by technical and military capabilities; and wars were never to be fought on home soil – rather, proxy wars were used if direct combat were necessary." This traditional, state-centred security concept was predominating for many decades. However, after the end of the Cold War during the 1990s, a new approach emerged which put more emphasis on the protection of the individual rather than on the national state. In 1994 the UNDP introduced the new concept of human security with two main aspects, the freedom from chronic threats such as hunger, disease and repression, linked with the protection from sudden calamities (UN 1994; Owen 2004; Jolly/Ray 2006). In a speech in 2000, former general-secretary of the UN Kofi Annan highlighted human security as a concept that "in its broadest sense, embraces far more than the absence of violent conflict. It encompasses human rights, good governance, access to education and health care (…). Every step in this direction is also a step towards reducing poverty, achieving economic growth and preventing conflict. Freedom from want, freedom from fear, and the freedom of future generations to inherit a healthy natural environment – these are the interrelated building blocks of human – and therefore national – security" (Annan 2000). This human-centred security concept initially aimed at reducing insecurities in order to ensure that human development is in accordance with freedom and health. Owen (2004: 19) defines human security as "the protection of the vital core of all human lives from critical and pervasive environmental, economic, food, health, personal and political threats." However, increasingly complex security challenges reinforced surveillance and security measures on a global scale and also led to an extended conceptualization of security. These developments already emerged before the terrorist attacks on September 11 2001 in the US. However, the tragedy of 9/11 amplified the shift in security policy at global level. Many governments worldwide significantly intensified their security and surveillance efforts (cf. Ball/Webster 2003; Haggerty and Samatas 2010). Increasing claims for a holistic security approach play an important role for the transformation of security policy which also complicated the concept of human security. This development includes a tendency to frame security threats as form of physical attack or violence from external sources (such as terrorism, organised crime etc.). While there are without doubt many external security threats, there are also internal ones that do not necessarily involve physical violations. Threats to economic stability or human rights are also security threats. This circumstance was initially addressed by the human security concept. Originally, human security focused on protecting the integrity of the individual from different kinds of threats such as hunger, diseases, natural disasters, conflicts or human rights violations (Owen 2004). However, to some extent, it created a life of its own (cf. Jolly/Ray 2006; Watson 2011).The extension of security measures with an inherent logic of increasing pre-emptive and preventive approaches (to early detect and combat potential threats) also fostered surveillance. In this regard, individuals are not merely

objects of protection but are, to some extent, also framed as potential threats. Thus, there is a partial inversion of the human security concept observable. This changed framing in security policy also stimulated surveillance activities and further complicates the relationship between privacy and security (Strauß 2017a).

From a theoretical stance, this paradigm shift and the transformations in security policy refer to so-called "securitization", which frames security as a permanent process and seduces with a seemingly predicting view on threats to foster the effectivity of security measures (cf. Buzan et al. 1998; Bigo 2000; Balzacq 2005; Bigo 2008; Guild et al. 2008; Watson 2011). In the sense of Foucault, securitization is a technique of government, "a mode of governmentality, drawing the lines of fear and unease at both the individual and the collective level" (CASE 2006: 457). It includes a rhetorical technique that makes strategic use of the term "security" to foster political objectives by rephrasing problems into existential security threats (Buzan et al. 1998). In this regard, securitization benefits self-fulfilling prophecy in the security discourse, and is thus also "a capacity to manage (and create) insecurity" (Bigo 2000: 174). Consequently, this can create a security dilemma where "the more one tries to securitize social phenomena (...) to ensure 'security' (...) the more one creates (intentionally or non-intentionally) a feeling of insecurity" (CASE 2006: 461). Hence, security here becomes an indeterminate, continuing process that is "marked by the intersubjective establishment of an existential threat with sufficient saliency to have political effects" (Watson 2011: 3). In this process, security is not framed as an objective condition but is linked and exposed to the political discourse (Balzacq 2005). Herein entailed is the risk that securitization seeks and creates legitimacy for exceptional security measures that are "outside the normal bounds of political procedure" (Buzan et al. 1998: 28f.). Due to its own particular dynamics, the process of securitization can lead to a "security continuum" in a problematic sense where the designation of "certain persons and practices as 'threats'" happens in a rather arbitrary manner (Guild et al. 2008: 2). Several scholars highlight that the linking of security and (im)migration is a prominent example for the dangerous effects of securitization (e.g. CASE 2006; ibid; Watson 2011). Securitization becomes particularly problematic if security is presented as a dominant issue of societal concern deserving higher priority than other state functions, and the protection of fundamental rights such as the right to privacy. In this regard, it justifies privacy intrusions and surveillance as necessity for a security purpose.

There is obviously no state of permanent security in the sense of an effective protection from all possible threats. However, security policy affected by the logic of securitization gives the impression that permanent security would be achievable. In fact, an overwhelming focus on security framed as predominating issue can undermine the effectiveness of security measures and goes at the cost of fundamental human rights, above all of the right to privacy (cf. Schneier 2003; Schneier 2006a; Chandler 2009; Nissenbaum 2010; Solove 2011; EGE 2014). More precisely, a central issue that complicates the relationship between privacy and security is its framing as contradictory in terms of a trade-off in the public discourse. Securitization reinforces this framing. The employment of security and surveillance measures is predominantly based on this trade-off that assumes a necessity to trade privacy for security. Privacy intrusions are then simply presented as a basic requirement in order to improve security. Two basic arguments that are often brought

into the debate are that personal information needs to be gathered to improve security and that citizens accept this trade-off as they require more security. However, a number of scholars criticized this trade-off model as being too simplistic as it reduces the complex interplay between privacy and security to a permanent contradiction in terms (e.g. Schneier 2006a; Nissenbaum 2010; Solove 2011; Pavone/Degli Esposti 2012; Friedewald et al. 2015; Valkenburg 2015; Strauß 2017a). Amongst others, Friedewald et al. (2015) explored the validity of this trade-off with a survey about individual attitudes on privacy and security. Following the logic of a trade-off suggests that those people with high security concerns care less about privacy. However, research results refute this as no correlation between security concerns and privacy in the sense of a trade-off was found. Hence, this statistical evidence against the validity of this trade-off model at the individual level shows that security is not weighed more important than privacy. Similar findings result from the SurPRISE project, the author was involved in (for empirical results concerning the perceptions of citizens see Section 5.2). Altogether, these research results point out that the relationship between privacy and security is not as simple as the trade-off model suggests.
But also besides this empirical evidence against a simple trade-off model, there are conceptual counter-arguments against it. Beginning with the term as such, already the wording "trade-off" implies that there are contradictory items where one wins at the cost of the other and vice versa. This trade-off operates at a political as well as at an individual level: at the political level, privacy is framed as a burden to effective security measures. Hence, privacy intrusions here are quasi-justified as precondition to improve security. At the individual level, the trade-off conveys that more security is possible but only if individuals accept privacy intrusions (EGE 2014). Thus, the assumption of a trade-off is an "all-or-nothing fallacy" (Solove 2011) where privacy and security are misleadingly presented as concepts with a "natural" inherent conflict. This assumption seduces to neglect privacy and data protection because they are framed as obstacle to security. Consequently, it inhibits to comprehend that the relationship between privacy and security is to some extent complementary and not per se contradictory. This complementarity is undermined by the assumed permanent demand to make an excluding choice between these values. Furthermore, also the view on security is reduced because measures that do not self-evidently entail privacy intrusion are neglected in this logic. Alternative and less intrusive security options are then hardly conceivable (Strauß 2017a).

The trade-off model with privacy as antagonist to security is ethically problematic as it finally jeopardizes liberty which essentially is the defining value for both concepts: privacy and security. Guild et al. (2008: 9) thus highlight that "democracy, the rule of law and fundamental rights are designed to protect the liberty of the individual within the society. (…) The precedence of liberty as a value that must be protected by all states (…) is key to ensuring that security in its coercive form is used only as a tool to support freedom and is subject to its priority. Hence, the individual is entitled to freedom and any interference with that freedom must be justified by the state on limited grounds and be subject to the important procedural requirements set out in (...) human rights instruments" (ibid). This means that security is not a prior value as often presented in political discourse. To the contrary, liberty is the superior value that connects privacy and security (cf. Lever 2006; Hildebrandt 2006; Nissenbaum 2010; EGE 2014; Klitou 2014). The essential role of

liberty and freedom in fundamental rights frameworks highlights this fact. Privacy as well as security represents a human right being part of these legal frameworks. Hence, neither privacy nor security is an absolute right but each is ever legally linked to the broader public interest and the well-being of the general public. It is the task of jurisdiction to clarify eventual conflicts. Here, the principle of proportionality is crucial which aims at coming to a fair balance between the general interests of the public and fundamental rights of individuals and their basic needs (De Hert 2012). The negotiation of this fair balance to ensure proportionality is the role of jurisdiction in the case of conflicts. However, this does not imply the existence of a permanent conflict. The right to privacy in short terms means that an individual has the general right to live free without interference into her private life. Thus, in legal norms (such as Art. 8 ECHR or Art. 12 UDHR)[30] privacy intrusions are only exceptionally allowed under certain conditions. Such conditions have to be generally in compliance with the public interest and have to be in accordance with the law and the protection of the foundations of democracy. In short terms, although the interference with privacy is foreseen by the law, this mode is always the exception to the rule but by no means a permanent choice.[31] As described in the previous sections, privacy is an enabler of other fundamental rights and represents a form of liberty, namely autonomy. Hence, the trade-off model that frames privacy intrusions as sine qua non for security neglects this and declines the function of privacy as enabler for other fundamental rights such as freedom of expression, information, assembly and movement (Lever 2006; EGE 2014; Klitou 2014). Finally, such a setting of inevitable privacy intrusions would reverse this as the exception and falsely present privacy intrusions as the norm. Furthermore, a security approach entailing a permanent conflict with privacy also undermines its very aim, to improve security: because intrusions into privacy can also create more insecurity by reducing the subjective, individual perception of security. For instance, a citizen may feel rather insecure and endangered in a political system that employs mass surveillance and exercises extensive power over public deliberation (for more details on this issue, see Chapter 5). Considering the intrinsic relationship between privacy, autonomy and identity, protecting privacy implies the protection of identity which is an important condition for individual security. Surely, balancing privacy and security is necessary as neither the former, nor the latter is an absolute right. However, as shown, this does not imply that there is a permanent conflict in the sense of a trade-off given. Nevertheless, this trade-off yet rather dominates security and surveillance practices, entailing a self-reinforcing dynamic as regards the security continuum of securitization.

3.3.2 Notions of post-privacy and transparency

Besides the security discourse, another privacy controversy related to ICTs concerns the tensions relationship between privacy and transparency. On the one hand, ICT-related societal transformations underline the need for re-conceptualizations of privacy. Enhancing

[30] Such as Article 8 of the European Convention on Human Rights (ECHR 1953); Article 12 of the Universal Declaration of Human Rights (UHDR 1948), or Article the European Fundamental Rights Charter (CFREU 2000).
[31] The European Court of Human Rights declared that already the storage of information about individuals interferes with privacy (EU-CHR 2017). See also Section 6.1.

transparency of information processing and of the processing (institutional) entities is a crucial part in this regard. On the other hand, due to continuing, extensive growth of personal data being processed and society becoming more transparent, the role of privacy and its protection is also increasingly questioned. Proponents of a so-called "post-privacy" notion occasionally proclaim the end of privacy due to lacking individual control over information flows. Though, as networked data and free information flows would entail a more open society without discrimination, there would also be no need for privacy anymore (e.g. Brin 1998; Heller 2011). This post-privacy view received some attention after science fiction author David Brin (1998) published his book called the "transparent society" with the main argument that societal transparency would increase and privacy increasingly erode due to informatisation and electronic surveillance. Similar arguments were brought after the Snowden revelations (e.g. by Spivack 2013). Brin further argued that surveillance practices could benefit society bottom-up if citizens employ these practices to observe the observers. This line of argumentation partially corresponds with a statement of former CEO of Sun Microsystems Scott McNealy, who claimed in 1999 "you have zero privacy anyway. (…) Get over it". Interestingly, McNealy revised his statement a few years later and referred to the vast importance of appropriate privacy protection: "It's going to get scarier if we don't come up with technology and rules to protect appropriately privacy and secure the data, and the most important asset we have is obviously the data on people(…)" (Lemos 2006). Several years later, similar pro-privacy statements come from other tech-companies such as Google. In 2013, executive chairman Eric Schmidt stated in an interview: "You have to fight for your privacy, or you will lose it" (Colvile 2013). However, at the same time, Schmidt denied any responsibility of Google or other tech-companies for privacy tensions as they would just provide innovative services. In Schmidt's view, privacy protection is mainly the task of individuals. In fact, Google, Facebook and other big technology companies contribute in many respects to contemporary privacy problems. Hence, such statements can be rather seen as part of their PR-strategy (such as Google's motto "don't be evil"). Moreover, the entrepreneurship spirit of internet companies seems to be relatively similar to a post-privacy notion. Schmidt e.g., also pointed out that "Google is a belief system. And we believe passionately in the open internet model. So all of the answers to the questions that we give are, at the core, about the benefits of a free and open internet" (ibid). Put simply, the main entrepreneurial attitude here seems to be that "privacy is important but not part of our business model". Furthermore, being asked about his opinion on the surveillance programs revealed by Edward Snowden, Schmidt[32] also said that state surveillance is "the nature of our society" (ibid; Holpbuch 2013). Similar statements can be found from Facebook founder Mark Zuckerberg, who among other things, said that privacy would be no longer a social norm as people today were more open and comfortable with sharing information (Johnson 2010). Being confronted with Facebook's privacy issues he stated: "What people want isn't complete privacy. It isn't that they want secrecy. It's that they want control over what they share and what they don't" (Zimmer 2014). These views suggest that privacy is not a main

[32] Google also cultivates some contacts with security authorities. CEO Schmidt, for instance, is also the chairman of a think tank of the US pentagon, the Defense Innovation Advisory Board. See, e.g., Alba (2016).

concern of major players in the online business but rather seen as an obstacle to a flourishing post-privacy age. In this logic, privacy protection is then framed as the responsibility of the individual alone (for a discussion about this aspect see Section 5.4). Indeed, companies like the mentioned provide many innovative technologies and services which contribute to an open society in many respects, ranging from fostering free access to information, distant communications, interactions etc. However, it is also a fact that there are privacy-intrusive practices related to these services. A notion of post-privacy with the complete disregard of corporate social responsibility undermines the societal function of privacy and, as a consequence, also the vision of a more open society supported by ICTs.
Considering the complexity of institutional power structures, agency problems, information asymmetries, imbalanced control over personal information etc.; the elusiveness of the utopian post-privacy vision is nearly self-explanatory (these issues are discussed in Chapter 5). For instance, the Snowden revelations (Greenwald 2014; Lyon 2014) prominently highlight that also in our highly networked society, surveillance practices are still the main result of the exercise of institutional power. Furthermore, surveillance mostly leads to a certain disparity and thus creates or reinforces power imbalances at the cost of the observed (these issues are discussed in Chapter 5). Hence, although bottom-up surveillance can contribute to relativize institutional power exercised top-down, it cannot abandon it and more importantly, it cannot emerge effectively without privacy: institutional power has the "starting advantage" of having already established control structures and organisations. In a fictive scenario without privacy being protected, individuals would not be able to organise themselves in an undetected and uncontrolled manner beyond this institutional power then. Consequently, bottom-up surveillance would be hampered from its very beginning. While the post-privacy notion itself is delusive, the debate is interesting as it mirrors the interplay of transparency and privacy which alters with ICTs. Transparency has multiple though interrelated meanings: In common usage, transparency relates to comprehensibility. This common notion includes that information is transparent and broadly available (although with different implications top-down and bottom-up). Technology challenges privacy protection, increases transparency of information so that individuals are increasingly observable. At the same time, the function of transparency transforms bottom-up with ICTs, enabling individuals and civil society to scrutinize organizations and thus fostering institutional accountability. This form of transparency is inter alia enforced by novel forms of activism and civil disobedience (e.g., visible in contemporary phenomena such as Wikileaks, the Anonymous collective, the Occupy movement, online activism etc.) as (self-declared) counterweight to institutional power regimes in the public as well as in the private domain. Also whistleblower Edward Snowden argued that reinforcing transparency and accountability of security agencies was a basic motivation for him to reveal secret surveillance programs (Gellman/Makron 2013). Leaving a discussion about eventual controversies of these examples aside, they indicate that with extensive ICT diffusion, informatisation and digitization of society seems to boost the societal demand for more transparency and accountability. In this regard, transparency is not a counterpart to privacy, but rather a complementary concept as regards the aim to foster accountability of (public and private) power regimes. In contrast to the common notion of comprehensibility, there is another meaning of transparency in the sense

of hiding or masking information. This form of transparency is inter alia relevant in human-computer-interaction and the development of distributed computer systems, where the design of transparent user interfaces helps to avoid information overload of users. Information, or parts of it, which are not necessary for a particular usage context are hidden from the user in order to reduce complexity and improve usability (e.g., of a technology, or application etc.).[33] At the first glance, this meaning of transparency is contradictory to a notion of comprehensibility. However, the aim is similar, i.e., to foster users in comprehending, e.g., information processes relevant for an interaction. The crux is that maximum information is difficult if not impossible to scrutinize. Hence, there is need for approaches to reduce complexity of information processing in order to support scrutiny. This rather technical notion of transparency includes this by restructuring information in layers so that information is easier to understand in different usage contexts for an individual user. A similar demand for restructuring information is given for the broader, socio-political dimension of transparency because otherwise, individuals and civil society in general, may be overburdened with masses of information which are neither comprehensible nor controllable. Therefore, intermediary entities are required, such as interest groups, the media, data protection authorities or similar, who e.g., scrutinize information processing of technologies and applications on behalf of individual citizens. Generally speaking, individuals alone can hardly comprehend the full complexity of societal processes. The same is given for all types of processes and systems in which their personal information is involved, which are hardly controllable by a single individual. The line of argumentation of the post-privacy proponents is thus misleading as it neglects these aspects. This meaning of transparency is particularly relevant for visions of ubiquitous or pervasive computing with the aim, to deeply integrate technology in society so that it is permanently available in a transparent, i.e., invisible way (cf. Weiser 1991; Zhang/Zhou 2006). Indeed, this vision can be in conflict with privacy, when information flows seamlessly from one socio-technical context to another so that there is no (informational) boundary anymore between the individual and technology (this aspect is discussed more in-depth in Chapter 4 and the further).

The different shades and meanings of transparency also mirror in the privacy discourse. Traditional conceptualizations of privacy frame it as a form of secrecy (which partially overlaps with the latter form of transparency in the sense of hiding information). Solove (2004: 42) called this the "secrecy paradigm", which "is so embedded in our privacy discourse that privacy is often represented visually by a roving eye, an open keyhole, or a person peeking through Venetian blinds." The problem Solove addresses is that a reductionist view on privacy as a form of secrecy neglects its vital functions for democracy and further complicates effective protection. Indeed, confidentiality and the way information is disclosed is an important privacy issue. However, privacy problems not merely concern breaches of secrecy or confidentiality of information and privacy is not just about hiding information or avoiding disclosure. Every individual has some things kept private and some selectively disclosed to others. However, this does not imply that

[33] For instance, a standard e-mail program would be less usable with a text-based command line and without a graphical user interface which provides simple mechanisms to send and receive e-mails.

"avoiding disclosure is the sum and substance of our interest in privacy" (Solove 2004: 43). For example, a person (call her Alice) with cancer may tell her closest friends and some colleagues about her disease; but this does not mean that she wants everyone in her workplace to know about her disease. Hence, in this case, a privacy breach would occur when e.g., a trusted colleague breaks confidentiality and tells another colleague. In the logic of privacy as secrecy, Alice would be better off not telling anyone and mistrust all of her colleagues. However, as privacy is more than secrecy, i.e., "a right to context-appropriate flows [of information] (…) there is no paradox in caring deeply about privacy, and, at the same time, eagerly sharing information as long as the sharing and withholding conforms with the principled conditions prescribed by governing contextual norms" (Nissenbaum 2010: 189). Thus, equally important than confidentiality is that the processing entities, as well as the purposes and contexts personal information is used for, are trustworthy and transparent, i.e., comprehensible and accountable. Similar to the logic of a false trade-off between privacy and security (as discussed in the previous Section), a notion of post-privacy implicitly embraces this secrecy paradigm, where privacy is reduced to a concept of hiding information. This view falsely frames privacy as concept contradictory to transparency. At the same time, privacy is seen as rather irrelevant as keeping information secret is hardly applicable as digital environments boost the availability of information.

This framing of privacy as a form of secrecy misses the public value of privacy (as highlighted in Section 3.2.2). Most individuals normally and legitimately expect their privacy to be respected not just in their private homes but also in the public. Privacy is thus not merely about avoiding information disclosure but also about ensuring that information is only available to selected persons as well as used for particular purposes (cf. Solove 2004; Nissenbaum 2010). This issue is observable in social media and ICT usage in general. These technological means offer a large variety of options to share, link and exchange personal and non-personal information. The fact that these options are enjoyed by a vast range of users does not necessarily imply that these persons care less about their privacy or accept permanent disclosure and use of their personal information. Several studies (e.g., Hazari/Brown 2013; Leimbach et al. 2014; EB 2015) challenge the assumption of lacking privacy awareness among social media or internet users. As will be shown in Section 5.2, similar is given as regards the perceptions of citizens on surveillance technology. Chapter 5 takes a closer look at the various privacy issues related to ICTs and the associated usage practices in security and surveillance contexts.

While unintended disclosure of information is only one aspect, another essential issue is the purpose and use of information beyond the control of the individual concerned. Transparency of information processing is also an important criterion for ISD, though in the sense of the individual concerned being enabled to comprehend and control her personal information flows, i.e., the contexts in which her information is being processed. The issues concerning ISD (presented in Section 3.2.3) are clear indications for how technology fundamentally altered the notion of privacy. The German Constitutional Court's line of argumentation impressively highlights several of today's core problems of privacy protection which have significantly reinforced since then. Especially as regards the creation of personality profiles undermining privacy. This refers to the problems of data

mining and profiling, i.e., the use of technological means and algorithms to explore patterns in large data sets for the gathering and storage of information about individual persons (Hildebrandt/Gutwirth 2008). ICTs provide manifold options for sophisticated profiling techniques, facilitated by the transformed relationship between public and private spheres, accompanied by increasing transparency of digital identities. This is prominently exemplified by social media platforms such as social networking site Facebook and others, which at first glance appear as sort of technology-mediated public spaces (cf. Boyd/Ellison 2007). However, in fact, social media platforms significantly differ from traditional public spaces as their accessibility and availability has other implications compared to the physical world: as a result of their technical design, in social media environments, user profiles including personal details, interactions, social relationships to personal entities (contacts, friends etc.) and non-personal entities (content used, produced, linked, shared, liked etc.) as well as the content of communications are commonly disclosed and thus explicitly observable (Strauß/Nentwich 2013). Hence, there is an explicit transparency of individual identity representations, their preferences, relationships and interactions etc. given, which did not exist in a similar form in physical public domains. In the analogue world, there is usually no systematic monitoring of this kind. Similar issues are given in case of other ICTs which foster connectivity, interactivity and transparency of their individual users. Substantially, ICTs enable numerous means to represent individual identities online or in other digital environments and thus facilitate digital identification (these developments are explored in Section 4). Entailed to this development are additional options for privacy intrusion, profiling and surveillance (as analysed in Chapter 5). These issues highlight that ICTs transformed the relationship between the public and the private, with a certain process of renegotiation. With their informational connectivity and interactivity, digital environments represent a blurry hybrid between private and public spheres with diminishing boundaries in-between. This has consequences for the representation and availability of individual identities and thus for privacy. Nissenbaum (2010) argues that privacy intrusions in general affect what she calls contextual integrity, as information concerning an individual is misused beyond the original context of information processing the individual may has agreed upon. In her view, informational norms determine the integrity of a particular context in which personal information is being processed. Otherwise, if these norms are breached then privacy is violated (ibid). With context she means "structured social settings characterized by canonical activities, roles, relationships, power structures, norms (or rules), and internal values (goals, ends, purposes)" (ibid: 132). Hence, for Nissenbaum, social contexts and context-relative informational norms are the central building blocks of contextual integrity which she proposes as a benchmark for privacy. She argues that the transgression of context-relative norms is among the core problems of privacy protection: "The norms, which prescribe the flow of personal information in a given context, are a function of the types of information in question (…) When these norms are contravened, we experience this as a violation of privacy (…)" (Nissenbaum 2010: 127). Hence, put simply, the context of information processing is of utmost importance for privacy protection. Contextual integrity embraces ISD and the privacy principle of purpose binding. Each is heavily strained as ICT usage entails multiple contexts of information processing fostered by the blurry boundaries

between private and public spheres. The crux is that socio-technical systems are increasingly networked so that information processing can span across multiple contexts and domains (these issues are discussed in Section 5.3). The multitude of socio-technical systems and practices may intendedly or unintendedly breach contextual integrity. Thus, individuals can hardly comprehend and control the processing of their information as these contexts often remain unknown or hidden to them. Therefore, a lack of accountability and transparency of information processing in socio-technical systems is a critical privacy problem. Or in other words: privacy protection is hampered when information asymmetries exist between the individual and the information processing entities, at the cost of the individuals' ISD (a more detailed discussion about this can be found in Section 5.1). Furthermore, even in known contexts, personal information may not be controllable for the individual. Consequently, a sheer fostering of individual control over her personal information may not be enough to ease this problem (for a further discussion, see e.g., Section 5.4).

Therefore, the boundary control function of privacy is not merely a matter of the individual concerned but also of the entities that process her information. Although privacy protection apparently involves individual responsibility, privacy also needs to be a predetermined, institutionalized concept inherent to societal values, norms and practices etc. provided and protected by all societal actors. Otherwise, the individual is literally left alone in protecting her privacy. The fact that privacy is intrinsically linked to autonomy and ISD does not imply that the individual is solely responsible for its protection. Self-controlled handling and provision of personal information is an important aspect of privacy, but it is by no means the only relevant one. As outlined in the previous sections, privacy is not merely a private but also a public value. Basically, the institutionalisation of privacy and data protection (at least in Europe) is an expression of this public value, including legal frameworks based on common privacy principles and their consideration in socio-technical practices. However, ICTs have significantly challenged the applicability of regulation and the privacy regime is considerably hampered in its effectiveness. One aspect is the lack of effective legal protection as technology usage partially involves legal grey areas (such as issues regarding consent). A general problem is that the law to some extent lags behind technological development. However, this is less a matter of insufficient laws but rather of their implementation and enforcement. To some extent, the effectiveness of legal regulations and international privacy protection standards also suffers from insufficient knowledge about the emergence of privacy impacts. Ambiguity about the mechanisms and forms of information processing that affect privacy also complicates the implementation of appropriate protection mechanisms. This also mirrors in socio-technical practices with digital identification playing an important role in this regard, as will be shown in the further parts of this research. As discussed in this chapter, intrusions into privacy and insufficient protection thereof can also have negative impact on individual identity-building. But also conversely, the way identity information is represented, accessible and processed can significantly affect individual privacy. Before privacy impacts and issues of lacking control over identity information are explored (in Chapter 5), the next Section sheds light on the emergence of digital identification resulting from ICTs. This is particularly important because of the strong interplay between privacy and identity:

both concepts are mutually dependent and consequently, transformations of socio-technical identification practices affect privacy and vice versa. As will be shown, ICTs entail and foster novel forms of digital identity representation and identification. As a consequence, identity information becomes more networked, available and transparent.

CHAPTER 4

The emergence of digital identification

The previous Sections dealt with the basic functions of identification as well as the complex interplay of privacy, identity and autonomy. It was shown that privacy provides vital space for free and autonomous identity-building, with informational self-determination as a crucial requirement. There is thus a naturally close relationship between identity and privacy given. As discussed previously, controversies with security and transparency affect the quality of this relationship. These controversies mirror in and are reinforced by socio-technical transformations in the realm of ICTs, which continuously pervade and transform society in numerous ways. Socio-technical practices are accompanied by increasing networking structures, connectivity and interactivity of digital information. Among other things, these developments enabled novel forms of identity representation and of digital identification which are explored in this chapter.

The emergence of ICTs can be generally understood as a socio-technical transition, where technological change entails the occurrence of new functionalities as well as incremental transformations of existing socio-technical functions and practices. Involved in these transformations is a certain pressure to adapt existing practices to technology. In our case, identification is such a practice which alters with ICTs. Digital identification practices can be found in a variety of contexts and online applications. These different developments are subject to the field of identity management (IDM) which deals with the digital processing and handling of identity information. For electronic services to conduct transactions, or online platforms to provide user profiles, certain concepts for authentication and identification are required. The emergence of IDM is thus closely connected to e-government and e-commerce as well as the rise of personalized online services and social media platforms. IDM approaches generally aim at enhancing control over digital information flows that represent personal identities. In this respect, IDM alters (digital) identification writ large. Similar profound are the transformations related to social media platforms which affect the way individuals interact with each other, present and perform their identities online. These phenomena are a showcase for the networking structure of ICTs which enable and foster networked identity representations. The emergence of digital identification results from various technical and societal developments which are presented and discussed in the following. This starts with a brief overview on the relevant technological developments of ICTs which influenced the increasing societal role of digital identification. Especially relevant in this regard is the field of IDM, which is subsequently explored including their major drivers. The next Subsections then deal with the emergence of social media platforms and their core functions as regards networked representations of individual identities. As will be shown, there are several overlaps between IDM and social media also as regards their societal

functions and usage purposes. Finally, major transition paths of digital identification are explored, which indicate a further extension of identification practices.

4.1 Overview on main development stages of ICTs

Before taking a closer look at the emergence of digital identification, the following paragraphs provide a brief overview on some of the main developments of the information society with ICTs becoming increasingly integrative parts thereof. The different development stages of ICTs are explored through the lens of a (metasystem) transition to grasp the functional transformations with increasing socio-technical connectivity and interactivity. Since the occurrence of the Internet, the information society evolves quickly driven by technological progress, which induced several changes in the processing of information, as Figure 9 shows based on a rough distinction between four development stages. These stages are derived from the three stages of a metasystem transition (individuation, interaction, integration) as outlined in Section 2.2.1. The individuation phase is divided into information and communication for two reasons: firstly, both are inherent to ICTs and secondly, to point out how ICTs framed as socio-technical metasystem turned from a relatively individual unidirectional medium towards a multidirectional, interactive tool.

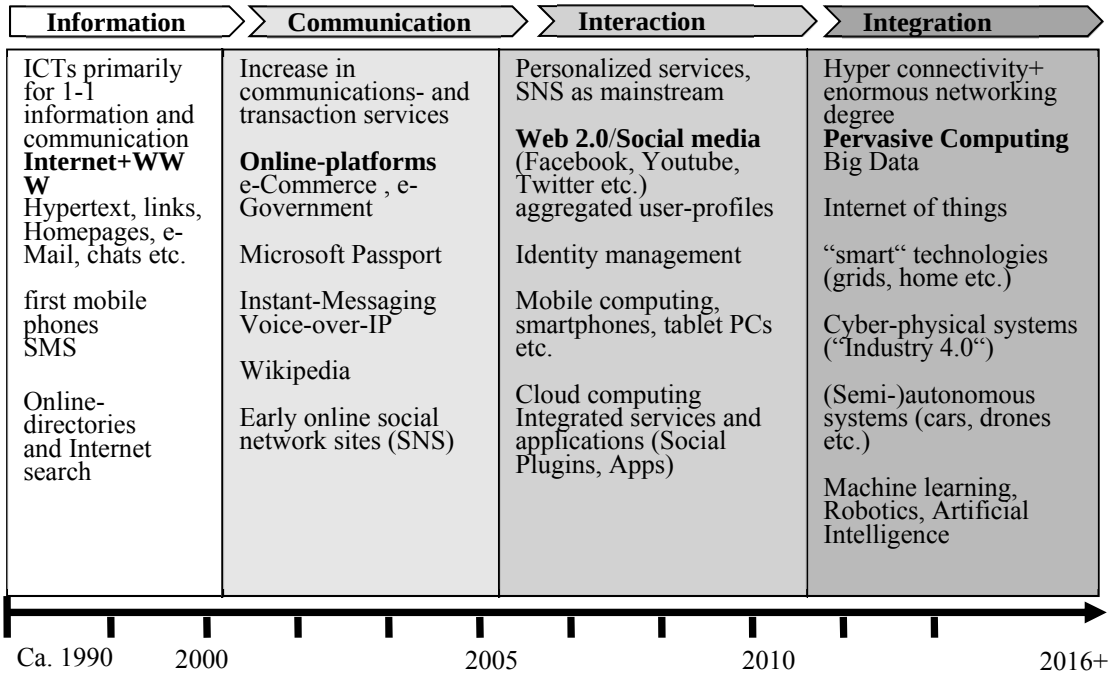


Figure 9: Development stages of the information society

These different stages highlight the transformative capacity of ICTs and how they became an integrated part of society in a relatively short period of time. The focus is on the changing modes of interaction exemplified by some key developments which intensified usage patterns of ICTs. While in the early years of internet usage, the focus was on one-to-one interaction, usage and exchange of information between two entities, this has significantly changed towards sophisticated modes of interaction and increasing integration of systems and applications, as observable today in many respects.

The first stage sketches the beginnings of ICTs, where new technological means to send, receive, access and distribute information occurred. The Internet is obviously one of the central technological innovations of this evolutionary development that significantly altered information processing in many respects. Emerging from its forerunner the ARPANET[34] during the 1980s, the Internet quickly turned into a large-scale system of interconnected computer networks on a global level. The development of the World Wide Web (WWW) in the early 1990s (cf. Castells 1996/2003; Leiner et al. 2003; Lessig 2006) paved the way for a myriad of innovations based on the digital processing of information. With the WWW and hypertext, the Internet became widely popular and transformed towards a new mass medium and a crossover technology carrying a broad range of applications, services, platforms etc. During the 1990s, the Web grew quickly with homepages, websites, online directories and so forth. The central online communication medium in that time was e-mail. In parallel to the expansion of the Web, mobile phones gained in popularity and increasingly entered the market. During that time, they were decoupled from the Internet. In 1992, the first SMS (short message service) (O'Mahony 2012) was sent; the first mobile phone with Internet capabilities, the Nokia Communicator, appeared in 1996 (Pew 2014). Also one of the first online instant messaging services, ICQ, appeared in that time which used to be widely popular (Patrizio 2016). In the same period, the search engine market took up. Today's global key player Google was found in 1998 and quickly challenged the former incumbent search engine providers Altavista and Yahoo. Also the first e-commerce sites appeared: e.g., Amazon is online since 1995, initially as online book store; in the same year, the online auction service eBay occurred (ibid). In this early information stage, ICTs were primarily used for one-to-one interaction and (compared with today) simple information communication services.

In the second stage, starting about 2000, it came to an increase in communication and transaction services. The first online platforms were created which made online services more accessible with the one stop shop principle, i.e., a single point of access enabling the usage of different services (e.g., in the fields of e-commerce and e-government). Commercial platforms such as Amazon significantly extended their scope during this stage and shifted from a book store to a global supplier of various products. Dominant vendors of operating systems such as Microsoft and Apple intensified their efforts to provide software tailored to online applications (e.g., Microsoft Windows 2000, or Apple's Mac OSX). In 1999, Microsoft released an early identity management approach, called "Passport" (see Section 4.2.1). In 2001, the largest interactive online encyclopaedia Wikipedia was launched. With Skype occurring in 2003, phone communication via the

[34] Advanced Research Projects Agency Network of the US defense department

Internet (voice-over-IP) as well as instant messaging became popular online applications. In the same year, myspace.com and linkedin.com were launched as well as other early online social network sites (SNS) occurred in that period – Facebook appeared in 2004 (Pew 2014).

From about 2005, the third stage is characterized by the emergence of so-called Web 2.0 (O'Reilly 2005) which led to novel forms of interactions, content-specific platforms and a further growth in the amount of digital information. For instance, the photo-sharing site Flickr or the video portal Youtube appeared in that period. Social media became mainstream platforms with today's major social networking platform Facebook accompanied by various specialized SNS (e.g., LinkedIn, Yammer, Xing, Researchgate etc.), various services from Google, microblogging site Twitter, news aggregation sites (e.g., Reddit), and many more. Web 2.0 and social media also boosted personalized services, aggregated user profiles and the commercialization of personal information. In about the same period, identity management became increasingly important (see next Section). During this time, also mobile and cloud computing became widely popular mainly boosted by smartphones and other portable devices. For instance, Amazon launched its first Cloud service in 2006.[35] The first iPhone appearing in 2007 boosted the smartphone market and mobile computing gained in importance (Pew 2014). These developments paved the way for integrated services such as apps, i.e., micro programs to extent functionality of applications on smartphones or other devices. They inter alia contribute to a closer linkage between online applications and mobile telecommunications.
Today, the fourth and current stage initiated about 2010 is present, where the integration of technology (and the generated digital information) into society has reached a new quality. This stage is accompanied by a high and rapidly growing degree of networking and hyper connectivity in a wide range of domains. Hence, ICTs today do not just play a dominant role in most daily practices may it be work, private communications, social relationships etc. They are also increasingly embedded into socio-technological domains that used to function basically without ICTs and the Internet. Examples are "smart" technologies entering private households with networked home appliances, smart TV sets, smart grids and -metering where ICTs are embedded into energy supply systems or other "traditional" domains that used to be offline. In parallel, there are developments towards autonomic computing, an umbrella term for information systems operating at multiple levels to provide, e.g., "hidden complexity, pervasive computing, context awareness, real time monitoring, proactive computing, and smart human-machine-interfacing" (Hildebrandt 2011: 5). Examples can be found in recent trends in the field of cyber-physical systems and "Industry 4.0" addressing novel forms of further automation. But also technologies such as remote-controlled aerial vehicles (drones) or self-driving cars can be seen as sort of (semi-)autonomous systems. These developments involve a convergence of different technologies; and the extensive arrays of data produced by these technologies feed into the big data paradigm aiming at exploiting large data sets (cf. Strauß 2015a). Related is also an increasing relevance and progress in the field of machine learning (cf. LeCun et al. 2015).

[35] Announcing Amazon Elastic Compute Cloud (Amazon EC2) - beta. August 4 2006, https://aws.amazon.com/de/about-aws/whats-new/2006/08/24/announcing-amazon-elastic-compute-cloud-amazon-ec2---beta/

Recently revitalized trends of artificial intelligence and the progress in robotics entail further increase in digital information processing. In total, classical visions of pervasive (or ubiquitous) computing (cf. Weiser 1991; ITU 2005), ambient intelligence, the Internet of Things, "smart" networks etc. with hyper-connected systems become more tangible.

Hence, ICTs became deeply integrated into society based on a variety of different but interwoven developments fostering networking structures and connectivity. In general, there is an increasing convergence between analogue/physical and digital environments observable stimulated by ICTs and digitisation (cf. Floridi 2010/2013; Hofkirchner 2013; Gillings et al. 2016). A major reason for the increasing integration of ICTs in society lies in their core capacity, i.e., to process digital information widely decoupled from space and time. This capacity significantly changed society and economy on a global scale in many respects. On the one hand, the expansion of a globally acting economy supported the emergence of novel networking domains and increasing need for mobility and connectivity of organizations and individuals. On the other hand, the peculiar characteristics of ICTs fit well to these increasing needs for networking, mobility and connectivity and amplify them. In this regard, ICTs entail a self-dynamic that reinforces further connectivity and integration. In his work on the information age, Castells (1996/2003) dealt with these issues and observed an increasing convergence of specific technologies that may lead to a highly integrative (information) system. Connectivity can be defined as "the mechanisms, processes, systems and relationships that link individuals and collectives (e.g., groups, organizations, cultures, societies) by facilitating material, informational and/or social exchange. It includes geophysical (e.g., space, time and location), technological (e.g., information technologies and their applications) as well as social interactions and artefacts" (Kolb 2008: 128). Indeed, as outlined, there are many different but interrelated technologies today that point towards a further integration of ICTs into society. As mentioned in Section 1, Floridi (2010: 6) uses the term "infosphere" to subsume the developments in the context of informatisation and digitisation. He describes ICTs as "re-ontologizing technologies" which represent the emergence and expansion of the infosphere, i.e., the "transition from analogue to digital data and the ever-increasing growth of our informational space" (ibid: 6). With re-ontologization he means "a very radical form of re-engineering, one that not only designs, constructs or structures a system (…) anew, but one that also fundamentally transforms its intrinsic nature, that is, its ontology or essence" (ibid: 6). Consequently, with the further digitisation of analogue objects or entities, "the digital deals effortlessly and seamlessly with the digital" (ibid: 7). Information is the essence of this infosphere that enables and fosters connectivity. Floridi (2010: 9) argues "what we still experience as the world offline is bound to become a fully interactive and responsive environment of wireless, pervasive, distributed, a2a (anything to anything) information processes, that works a4a (anywhere anytime), in real time".

Although Floridi's notion of an infosphere is useful to highlight the rapid progress of ICTs, it is debatable whether it appropriately grasps the manifold challenges of the information society; not least as it suggests the emergence of a hyper-connected society without alternatives (for a critical discussion see, e.g., Hofkirchner 2010). Nevertheless, there is little doubt that ICTs deeply affect and transform society in an extraordinary fashion. But it is not certain that an interconnectedness or hyper connectivity of all things

in a sense of anything2anything emerges, as Floridi assumes. However, a seamless flow of information between completely digitized entities does not exist yet. In fact, there are various natural frictions in information processing observable in many domains, and daily routines involve completely non-digitized interactions. However, it is true that as soon as ICTs are involved in an interaction, information can be digitized which may reduce informational frictions (as e.g., prominently exemplified by smart phones as quasi-pervasive devices), so that it can be easily processed further. Ontological frictions diminish in the sense of a reduced "amount of work and effort required to generate, obtain, process and transmit information" as Floridi (2010: 7) argues. As informatisation and digitalization continuingly proceed (including trends such as the Internet of Things, pervasive computing etc.), societal transformations reducing frictions are likely. In this regard, ICTs significantly affect society with the increasing potential to seamlessly process information in digital environments from one system to another.

Already today, ICT-induced connectivity is visible in various every-day contexts (such as the standard setting of personal computers to automatically establish an online connection, smart phones being constantly connected via telecommunications networks and increasingly also via the Internet, heavy use of social media platforms etc.). Permanent connectivity of individuals is not least envisioned and constantly promoted by technology vendors and other actors of the digital economy. The already extensive amount of networking devices indicates that these visions incrementally take shape. The OECD (2013a) estimates that by 2022, about 14 billion "smart" devices will be used in households on a global scale. Network provider Cisco (2016) even expects over 26 billion networked devices worldwide by 2021 which is more than three times the world population. A study on big data (sponsored by data storage enterprise EMC), predicted the amount of digital data produced globally will exceed 40 zettabytes then, which equals 40 trillion gigabytes. This corresponds to more than 5200 gigabyte per person (Mearian 2012). Although these predictions are naturally linked to the business models of their sponsors, a further extensive growth in the amount of digital data can be expected by all means. The expansion of networked devices implies a growth in the relations between different informational entities which produces further digital data. This expansion makes an exorbitant growth in the processing of personal information likely as well. Especially when individuals are surrounded by networked devices gathering and processing their information; as suggested by promoted visions in the realm of pervasive computing (e.g. Weiser 1991; ITU 2005) and the according socio-technical developments.

4.2 (Digital) Identity management (IDM) – overview and basic concepts

The presented technological transformations also affect the societal role of digital identities and identification. ICTs basically enhanced connectivity, interactivity and availability of information. This also brought novel means to digitally represent, gather and process identity related information, and thus altered identification practices for various social,

political and economic purposes. Corresponding transformations mirror in the increasing importance of these issues in science and society in parallel to the technological developments. Several scholars dealt with digital identities in different respects: for instance, Clarke (1994a/1994b) observed the changing role of identity already at the dawn of digital technology, when organizations began to use technological means for identification in e-transactions. He described it as emergence of a "digital persona" which he defined as "a model of an individual's public personality based on data and maintained by transactions, and intended for use as a proxy for the individual." (Clarke 1994b). Clarke here early foresaw the emergence of what today is addressed by the field of identity management (IDM) which several years later became an important domain in research and technology development (see below). Abelson and Lessig (1998) dealt with the unbundling role of identity in cyberspace and argued for technical design to improve users in controlling "the strength of the link between their real world and cyber-identities". Solove (2004) pointed out that information technology enabled the creation of digital dossiers that refer to digital representations of a person. These digital dossiers "resulted in an elaborate lattice of information networking, where information is being stored, analysed, and used in ways that have profound implications for society" (ibid: 3). Issues of identity in the information society were the main focus of the EU funded research project FIDIS[36] which created a network of excellence to support identification and the management of identities with technology (cf. Halperin/Backhouse 2008; Rannenberg et al. 2009). Several scholars explored the effects of ICTs on identity-building and the right to privacy (e.g. Hildebrandt 2006; Hildebrandt/Gutwirth 2008; Rouvroy/Poullet 2009; Nissenbaum 2010). Hildebrandt (2008: 56) argued that society has entered "the age of identification" as governments and enterprises create systems for identification and authentication. Similar also Lyon (2009) who observed a global growth in the creation of identity cards pushed by governments and the security industry. A number of studies found that Web 2.0 and social media changed self-representation and identity performances online (e.g. Boyd/Ellison 2007; Leenes 2010; Nentwich/König 2012; Ellison 2013). The broad availability of social media profiles is also found to be supportive for surveillance and thus challenges privacy protection (e.g. Acquisti/Gross 2006; Strauß/Nentwich 2013; Fuchs 2015). The implementation of electronic identification systems created new means to foster e-commerce and e-government services which affects the identification of consumers and citizens (e.g. Kubicek/Noack 2010b; Aichholzer/Strauß 2010a/2010b; Strauß 2011; De Andrade et al. 2013). These systems are also instruments of governance, serve political and economic objectives as well as security and surveillance practices (e.g. Bennett/Lyon 2008; Lyon 2009; Glässer/Vajihollahi 2010; Whitley/Hosein 2010).

Hence, ICTs entail various issues related to digital identities and identification. The field of IDM encompasses different approaches to deal with these issues and serves as an umbrella term. IDM is thus an emerging field of research in the information society that gains increasing relevance due to the further pervasion of ICTs (e.g. Halperin/Backhouse 2008; Rundle et al. 2008; Rannenberg et al. 2009; Kubicek/Noack 2010a; ITU 2010; Pfitzmann/Hansen 2010; OECD 2011; Strauß 2011; De Andrade et al. 2013; Whitley et al.

[36] Future of Identity in the Information Society, www.fidis.net

2014; EU-C 2016a; Grassi et al. 2017). The International Telecommunication Union (ITU) defines IDM as "set of functions and capabilities (e.g., administration, management and maintenance, discovery, communication exchanges, correlation and binding, policy enforcement, authentication and assertions) used for assurance of identity information (e.g., identifiers, credentials, attributes); assurance of the identity of an entity and supporting business and security applications" (ITU 2010: 4).

The emergence of IDM is part of the outlined societal transformations related to the deep pervasion of society with ICTs. IDM gained in importance at about the same time as ICTs became increasingly interactive and personalized services began to spread. This indicates a certain societal demand for concepts to manage digital information flows related to personal identities. Digital identification gained in importance with online services, e-transactions and social media becoming mainstream activities. This reinforced trends of service personalization and led to a significant growth in the processing of personal information. Against this background, IDM can be seen as an attempt to improve the controllability of digital information flows related to personal identities (Strauß 2011). Besides their many benefits, these developments entail several tensions and challenges as regards privacy, security and control over personal information (as will be explored in Chapter 5).

IDM approaches occur in many different shades and applications in the public as well as in the private domain. This may also include areas where practically no strict (hard) identification (disclosure of one's real identity) is required such as in online services or social media applications. Particular forms of IDM exist in the government and business sector, as these domains have higher requirements for secure means to ascertain one's identity (such as citizens requesting public services, or customers conducting an online transaction). IDM in these contexts aims at improving identification and authentication procedures between citizens and public administration respectively customers or other parties involved in commercial transactions. A digital identity device here usually fulfils two basic functions: the identification of its holder based on a unique identifier and authentication by providing a digital signature, e.g., to enable legally binding online transactions (Aichholzer/Strauß 2010a; Kubicek/Noack 2010a). But besides this special forms, in a broader sense, IDM is involved in every kind of personalized online service or application which processes identity information. Basically, IDM concepts are employed to provide electronic (or digital) identity devices (eID) such as user profiles and the according login procedures including the relevant architectural components. The extent to which a user is explicitly identified naturally depends on the concrete application.

4.2.1 Basic building blocks of digital identity management

In brief, an identity management system comprises all applications, infrastructures and procedures where identity information is involved. There is a broad scope of technical IDM concepts ranging from different forms of single-sign-on (SSO) solutions enabling access to multiple applications, token-based approaches based on smart cards, mobile phones or other technical devices, equipped with identity information up to biometrics (cf. Halperin/Backhouse 2008; Lips et al. 2009; Nauman/Hobgen 2009; Glässer/Vajihollahi

2010; Aichholzer/Strauß 2010a; ITU 2010; Kubicek/Noack 2010b). There are four general IDM architecture models: siloed (or isolated), centralized, federated and user-centric (cf. Jøsang et al. 2007; Bhargav-Spantzel et al. 2007; Rundle et al. 2008; Strauß/Aichholzer 2010). In a siloed system, identification is isolated from other applications, and therefore, it does not provide SSO. Users can choose different credentials (username, password) for each application. Hence, the very same ID is not used for multiple purposes. A centralized approach provides SSO whereas also the identity information is processed centrally by a single authority. Hence, the same ID is in multiple use and accessible to different service providers. Widespread are federated models which include siloed and centralized components. This federation involves a central identity provider (IdP) who serves as intermediary between the individual user and the different service providers who accept the IdP as trusted entity. This allows SSO with the same ID in multiple applications and ideally, identity information is not controlled by a single authority as the IdP provides identity credentials (e.g., a unique identifier) but does not process all personal information. The user-centric model follows a federated approach but aims at providing more user control. This can be e.g., by providing a particular technical ID device (e.g., a smart card or other hardware token) to the user and/or options to choose between different IdPs independent from applications and services.

In practice, all types can be found in various contexts. One of the first centralized digital IDM approaches was Microsoft's ".NET Passport", created in 1999. It provided a SSO solution to access different Microsoft applications with one central user account. Further attempts to establish Passport as central IDM system also for non-Microsoft applications failed as the service was neglected due to its proprietary design and a number of serious privacy and security concerns. In 2001, several privacy and consumer protection groups together with the US Federal Trade Commission (FTC) brought in legal complaints against Microsoft's Passport system. Similar actions were undertaken by the European Commission (EPIC 2003; Dmytrenko/Nardali 2005). Microsoft abandoned the system in 2006 and re-launched the concept under the label Microsoft Live ID (Dmytrenko/Nardali 2005; Jøsang et al. 2007). The current IDM system is called "Microsoft Account" which can be used as SSO to all Microsoft-related applications and platforms[37]. Similar IDM approaches exist from Internet and social media enterprises such as Facebook Connect or Google Sign In, which are labeled as "social logins" (see Section 4.3.1). A related approach is the OpenID project[38] formed in 2007, a decentralized, more user-centric IDM concept for web-based authentication, in which several Internet companies are involved (among others also Facebook and Google). While OpenID basically provides a low threshold to act as IdP, there are certain risks regarding privacy and trust as well as its vulnerability to phishing attacks (Bonneau et al. 2012). Many of the current governmental eID systems follow a federated approach with some user-centric features including smart card technology which corresponds to token-based (the smart card) and knowledge-based (a PIN code) identification. A reason for this is that smart card technology is widespread in many contexts such as ATM or credit cards but it can also be used in combination with

[37] See, e.g., https://account.microsoft.com/about
[38] http://openid.net/developers/specs/ https://de.wikipedia.org/wiki/OpenID

smartphones which enjoy significant growth in usage rates and thus become increasingly attractive as eID devices (cf. Kubicek/Noack 2010b; Strauß/Aichholzer 2010). Each approach has advantages and disadvantages as regards privacy and security, which are discussed in Section 5.4.2.

Besides the different system design options, there are some basic components relevant in each IDM system. In general, there are at least two basic processes involved (cf. Jøsang et al. 2007): in the initial (registration) process, a user is requested to provide personal information and is then issued with a particular identity device (e.g., a user profile or account, an ID document or a similar artefact containing a unique identifier and possibly additional identity information). This ID device is then used to authenticate users and to control their access to one or more services in every usage process or user session. The figure below sketches a simplified model of a typical user session with the main building blocks of an IDM system:

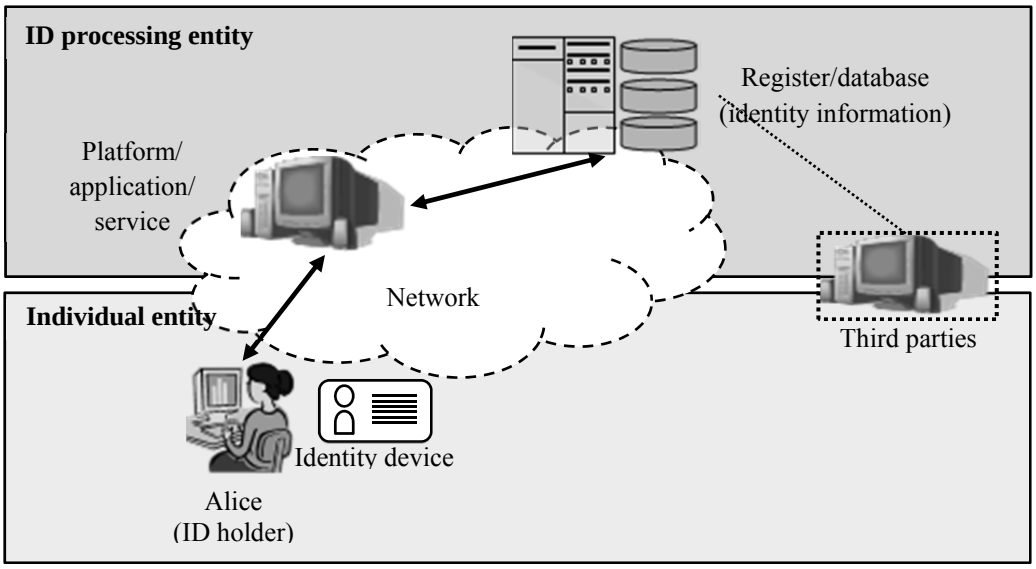


Figure 10: Basic building blocks of digital IDM

This model highlights the different components relevant for digital identification. There are at least two entities involved: the individual entity (Alice), i.e., the identity holder using a service (e.g., of an online platform), who is being identified by an (institutional) entity. The latter can be a single person but mostly is an organization or institution proving e.g., the service. The interaction begins with Alice visiting the platform of the provider by using a technical system (e.g., a computer) and becomes identified based on her (pre-registered) identity device. This device can be merely virtual (such as typical credentials username and password to access her account or profile) or an additional physical token (e.g., a smart card, mobile phone etc.). In each case, the ID device carries an identifier or additional ID information. Depending on the service, different types of information related to the user are processed and partially stored in some kind of repository or database which makes it available for further processing (e.g., for eventual third party entities). In case of an online

or mobile service, the whole interaction involves some kind of network medium (e.g., Internet, cellular network etc.) where the information is processed.

From a systemic perspective, this simple example includes at least six different elements: the (social or human) system A (individual Alice), socio-technical system B (the institution), the information processing system C (the service or application, or front office), the information processing system D (the database, or back office), the network system E, and the ID device F. Hence, compared to face-to-face identification, there are more systems involved (at least the network but mostly also the users machine). This seems trivial but it is important to understand how many entities are basically involved in the digital processing of identity information. The more entities are involved, the more relations exist to the individual identity. This has consequences for informational self-determination as identity information may be processed in multiple domains beyond knowledge and control of the individual. The crucial peculiarity of digital identification, though, is that the whole interaction of an individual with a socio-technical system is reproducible and traceable. The technology and each sub-system involved may create additional information referring to the individual's identity which can be theoretically gathered and processed further over the involved network. This issue has several privacy implications which are elaborated more in-depth in Sections 5 and 6.

4.2.2 Major drivers of digital identification

(Digital) identification serves as a socio-technical "connector" in many respects and has multiple social, political and economic functions. It became increasingly relevant with the possibility of electronic transactions. Organizations began to integrate IDM in business processes in order to provide and conduct e-commerce services where identity verification is involved. Central aims of IDM here are to improve efficiency and security of electronic services as well as to harmonize identification and authentication procedures (cf. Strauß 2011). An expanded distance between individuals and organizations contributed to the relevance of digital identification. This distance increased with the growing size of organizations and administrative bureaucracy in modernity, further reinforced by globalization (cf. Giddens 1984; Clarke 1994a; Castells 1996/2003). As ICTs enable mobility, distant communications and information processing, partially decoupled from space and time, they allow overcoming this distance in a virtual sense. With its inherent connecting function (Section 3.1.2) identification is an essential process in this regard as it creates ties between different entities. In the analogue world, anonymity is a standard setting and many transactions are possible without identification as there is no necessity given. In the digital world, this is more complicated. If a transaction is completed in a single step (e.g., in the case of instant payment and exchange of goods[39]), usually no identification is needed. With a greater distance between individuals and organizations (e.g., customer and company, citizen and government, user and service etc.), as given in

[39] If a product is bought by direct cash payment, usually no identity information is digitally processed. This is already different if a debit card, credit card etc. is involved. Then, the card as a technological device serves as identity token. In this case, there is a distance between the buyer and the seller as a technical system handles the payment transaction. The same principle is given in online services where technology is the intermediary of an interaction.

online services, reliable authentication and identification becomes more important (e.g., for secure transactions). Consequently, this distance contributed to organizations increasingly gathering personal information based on the assumption of its necessity to do so (cf. Clarke 1994a). Hence, digital identification in this regard serves as means to compensate this distance by providing an informational link between different entities.

Against this background, it is reasonable that digital identification and IDM are particularly important for a variety of public and private services online (e-business/e-commerce, e-government, e-procurement, e-health etc.). But this increasing relevance is not merely the result of technological progress. There are also certain political and economic considerations that reinforce the extension of digital identification mechanisms. The digital economy plays an important role, which is observable in national and international policies and strategy documents concerning the information society. For instance, for the OECD as well as the European Union, digital IDM is seen as a fundamental means for the further development of the digital economy (OECD 2011; EU-C 2016a). A central aim is to stimulate digital markets and to improve the efficiency and effectiveness of administrative procedures. In its recent e-government action plan, the European Commission announced to strengthen the efforts "to accelerate the take up of electronic identification and trust services for electronic transactions in the internal market"; as well as further "actions to accelerate the cross-border and cross-sector use of electronic identification (eID), including mobile ID, and trust services (in particular eSignature, website authentication and online registered delivery service)" (EU-C 2016: 4). Accordingly, the introduction of digital (electronic) identity management systems (eIDMS) is seen as a key driver for online services in public and private sector. IDM plays an important role for the digital agenda of the European Union for many years. Specific directives at EU level were created to regulate the use of electronic signatures[40] and identification.[41] Consequently, most countries adapted their legal frameworks to EU regulation and built the techno-organizational infrastructures to implement their eIDMS. During the last decade, many approaches have occurred and a number of governments world-wide (including Asian, Arab, African, North American and European countries) and particularly in Europe have already implemented eIDMS or started according initiatives in this regard (cf. CEN 2004; Arora 2008; Naumann/Hobgen 2009; Aichholzer/Strauß 2010a; Kubicek/Noack 2010a; Whitley/Hosein 2010; WH 2011; Gemalto 2014; Whitley et al. 2014). In the long run, the European Union aims at establishing a Pan-European eID system. According large scale pilot projects STORK and STORK 2.0[42] were already set-up to foster interoperability of different eIDMS at national and EU level (cf. EU-C 2010a; De Andrade et al. 2013; Brugger et al. 2014). In the same vein, the US strategy on identity management aims at establishing an operational identity ecosystem including individuals, businesses, non-profits, advocacy groups, associations, and governments at all levels (WH

[40] The e-signature Directive 1999/93/EC created a community framework for the use of electronic signatures in the EU that invited member states to create according national laws; see http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31999L0093&from=DE

[41] The so-called eIDAS act: Regulation (EU) No. 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32014R0910

[42] STORK is the acronym for secure identity across borders linked https://www.eid-stork2.eu/

2011). Since 2012, a special identity ecosystem steering group (IDESG[43]) is entrusted with the funding of pilot projects in the field; as part of an initiative of the US National Institute of Standards and Technology (NIST).[44] The NIST is elaborating guidelines for digital identities (Grassi et al. 2017). Hence, there are various activities concerning digital identities on a global level. The so far largest national electronic identification program was implemented in India with over one billion identification numbers issued by the end of 2016.[45]

These developments highlight that the employment of digital identification is related to political and economic interests. Besides these governmental initiatives which mainly focus on formal identification, social media contributed significantly to a general expansion of digital identification mechanisms. Social networking sites (SNS) such as Facebook and other social media platforms represent the most common technological phenomenon related to digital identities. They are showcases for personalization in online services. In general, social media stimulate the societal needs to communicate, interact and exchange (personal) information with others as well as to self-presentation. A mapping of all these activities can provide many details of individual identities. Social media platforms use IDM to handle user registration and create a user profile which is required to access the platform. This profile equals the digital identity of an individual user. The handling of identity information is obviously important in social media as they naturally enable and foster personalization, social interactions, user-generated content etc. where massive amounts of personal information are created and processed. Hence, social media represents a further example for an IDMS though without the need for a legally binding identification and disclosure of one's real name. In this regard, SNS employ "softer" forms of IDM compared to governmental eIDMS. However, at the same time, identification in social media is more extensive: it is a precondition to access social media platforms and enjoy the broad range of interaction possibilities. Moreover, most SNS providers try to engage users in providing their real names and it is challenging to effectively use SNS anonymously (see Section 4.3.3). Before the occurrence of Web 2.0 and social media, explicit identification and provision of identity-related information in online services used to be rather limited to transactions where identification is legally relevant. Social media changed this significantly as it fostered the use of digital identities also for other purposes than transactions. The previously rather occasional processing of personal information bound to particular applications and domains (e.g., for personal communication or transactions online) became widely boosted and "normalized" in social media environments which are easily accessible. The extensive amounts of information that directly and indirectly refer to an individual provide deep insights into one's interests, relationships, activities, behaviour etc. and thus one's identity (for more details see Chapter 5). While individuals basically use social media for social reasons, there is a strong economic rationale behind the provision of these services. The identity profile thus feeds into numerous business models.

[43] https://www.idesg.org/
[44] Five Pilot Projects Receive Grants to Promote Online Security and Privacy. *NIST News Post*, September 20 2012, https://www.nist.gov/news-events/news/2012/09/five-pilot-projects-receive-grants-promote-online-security-and-privacy
[45] Unique identification authority of India https://uidai.gov.in ; for an overview on the program see also https://en.wikipedia.org/wiki/Aadhaar

The basic business model of global internet companies like Facebook and Google is service-for-profile (cf. Elmer 2004; Rogers 2009), i.e., the usage is for free but users indirectly pay the services with their personal information.

Hence, identification in the government as well as in the private sector is closely connected to the digital economy. A major reason for its relevance in economic contexts is that identification enables the personalization of services for customer relationship management (CRM), targeted advertising, profiling activities of Internet and social media users etc. Thus digital identification is seen as promising means to exploit new digital markets aiming at monetizing personal information[46]. The high economic value of this information and the profitability of the according business models can be derived from the massive annual revenues of global Internet companies such as Facebook or Google/Alphabet.[47] According to Forbes, Facebook's market cap in May 2017 was about 407 billion USD; the corresponding the market value of Google's mother company Alphabet is approx. 580 billion USD.[48]

Besides these economic aspects, another strong driver of digital identification is security. This includes several objectives ranging from information security to provide secure user accounts and digital identity devices, secure digital signatures to allow for legally valid online transactions, up to cyber security, fighting identity fraud and terrorism and thus national security (cf. CEN 2004; EU-C 2006; EU-C 2010a; Glässer/Vajihollahi 2010; Strauß 2011; OECD 2011; Whitley et al. 2014). As regards information security, multiple user accounts are often seen as problematic as users tend to reuse the same passwords for different accounts. This makes security breaches more likely (Metz 2012). The assumed number of accounts per person ranges from 8.5 (Riley 2006) to 25 or more (Egelman 2013). Against this background and the expected further growth in services requiring user accounts, IDM concepts that aim at harmonizing login procedures to reduce eventual security risks of multiple user accounts seems plausible. The basic consideration is that a unified IDM approach with an according ID device usable for multiple applications improves efficiency as well as security of online services. A rationale is to develop approaches that ease the problem of the currently dominating "bricolage of isolated, incompatible, partial solutions" of online identification (Leenes et al. 2008: 1). However, at the same time, a "pervasive IDM layer" creates additional security and privacy risks (Rundle et al. 2008: 19). Although privacy and data protection are particularly essential issues for governmental eID systems, their role in the IDM discourse is rather an implicit one, compared to the emphasis put on service efficiency and security (Strauß 2011).

---

[46] For example, consulting companies like the Gartner group give advice to companies on how to monetize their customer data: Gartner (2015): How to Monetize Your Customer Data. December 10, http://www.gartner.com/smarterwithgartner/how-to-monetize-your-customer-data/

[47] Google's revenue in the first quarter 2016 is over 20 billion USD (see e.g. CNBC 2016a).
Facebook's revenue in the first quarter 2016 is over 5 billion USD (see, e.g. CNBC 2016b).
The estimated global advertisement revenue of Facebook in 2015 is over 17 billion USD, (e.g., Statista 2017a).
Google's total revenue over 74 billion USD in 2015 (e.g. Statista 2017b).

[48] Forbes Ranking of the world's biggest public companies #119 Facebook as of May 2017, https://www.forbes.com/companies/facebook/
Forbes Ranking of the world's biggest public companies #24 as of May 2017, https://www.forbes.com/companies/alphabet/

Besides the central aim to improve security of online services, digital identification is also framed as means of national security in political contexts. For instance, in some EU policies, the employment of national eIDMS (and in a long term perspective also of interoperable systems across Europe) is also mentioned as a tool to combat crime and terrorism (cf. CEN 2004; EU-C 2006; Bennett/Lyon 2008; Glässer/Vajihollahi 2010; Whitley et al. 2014). Bennett and Lyon (2008) provide a critical overview on eID card schemes in different countries in this regard. Several scholars (e.g. De Hert 2008; Lyon 2009; Ajana 2013) observed a global growth in personal data collections and centralized databases for law enforcement and national security including trends to gather biometric information. According to De Hert, the general increase in identification is part of a "global tendency towards ambient intelligence security enforcement scenarios, relying on the massive collection and processing of (personal and non-personal) data in combination with data mining and profiling techniques" (De Hert 2008: 73). Correspondingly, in the past ideas and claims to establish digital IDs at a global level to improve information security and/or national security were occasionally expressed: such as from Eugene Kaspersky, the CEO of a prominent anti-virus company, who raised the idea for a governmental ID for every computer user (Long 2010); or claims of Interpol to introduce a global digital ID for migration control (HNS 2011). As social media provide extensive arrays of identity information, their value for law enforcement is relatively obvious. Social media activity and thus also its users are increasingly monitored by security agencies (Belbey 2016) and since the Snowden revelations in 2013, there is hard evidence for global mass surveillance programs of Internet activity (Greenwald 2014).

There is a general trend of increasing identification and processing of identity information, which is not limited to IDMS for e-government and e-business, where a need to provide more sophisticated online identification and authentication to improve service quality is plausible. In general, IDM involves (explicit) attempts to unify the handling of identity information. IDM mostly focuses on easing identification processes and controlling identity information from the perspective of institutions or organizations. This does not necessarily comply with the needs of individuals to be identified and having their information protected. Especially as an individual may not always be aware of the extent to which her identity information is being processed. An e-commerce provider, for instance, is naturally interested in providing secure transactions to minimize his risks of economic loss. Therefore, gathering personal information from his customers is an obvious approach from his point of view, which may be accepted by the customer as a standard procedure. However, depending on the particular business model, this information may be used also for further purposes (such as targeted advertising). These practices may lead to tensions and conflicts with the needs and rights of consumers. The use of governmental IDM systems for public services is mostly regulated and citizens are usually explicitly requested to identify for a particular, limited purpose. However, this is not exactly the case in social media or similar applications where identity related information is commercialised. Moreover, as the Snowden case highlighted, digital identity information serves a variety of security and surveillance purposes. Regardless of the variety and multiple functions of IDM approaches, two predominating rationales become apparent which stimulate the increase in digital identification: economic and security aspects. As outlined, economic

growth and security objectives play an important role in the implementation and extension of digital identification. Furthermore, due to the extensive availability of identity information in digital environments (such as social media), profiling or what could be called "identity mining" is a lucrative activity for several public and private actors as will be discussed more in-depth in chapter 5. The next Section presents and discusses the role of social media for the representation and processing of identity information, which is a showcase for networking dynamics of digital identities.

4.3 Social media and networked online identities

Social media and particularly social networking sites (SNS) are specific phenomena having significantly contributed to the trend of online identity representation and expanded identification. Firstly, because the foundation for its basic functionality is the processing of massive amounts of personal information related to individual identities. Secondly, given their enormous networking degree of SNS and similar platforms, they demonstrate how profoundly interconnected online identities are already today. In this regard, social media can be seen as a blueprint for our identities becoming highly interactive, dynamic but also transparent and controllable. Considering a further expansion of digitally networked environments as indicated by several developments (e.g., internet of things, ambient intelligence, pervasive computing etc. as outlined previously), these issues can be expected to intensify in other socio-technical contexts as well. The following sections provide an overview on the emergence and core functions of social media platforms. Thereafter it is shown that the scope of social media expands, including tendencies to integrate online user profiles into formal identification procedures.

Social media and SNS as prominent application occurred in a relatively short period of time and quickly evolved from its beginnings as niche applications towards a global mainstream phenomenon. Ellison and Boyd (2013: 158) define SNS as "a networked communication platform in which participants 1) have uniquely identifiable profiles that consist of user-supplied content, content provided by other users, and/or system-provided data; 2) can publicly articulate connections that can be viewed and traversed by others; and 3) can consume, produce, and/or interact with streams of user-generated content provided by their connections on the site." The origins of SNS can be traced back to the 1990s where early community pages (e.g., Geocities), online contact groups and the first instant messaging services (e.g., ICQ) emerged. Initially, these applications where separated from each other and merely part of small niches in the WWW, such as smaller interest groups or grassroots organisations used community pages to share thoughts and ideas. Instant messaging services provided new options to connect and communicate online. As these different services became accessible via a single webpage and linked to a user profile, the first SNS appeared in 1997 with sixdegrees.com. Though this site shut down in 2000 as its business model failed, its general setting including a single user profile as standard access feature became state-of-the-art and stimulated the further development. Since 2000 the amount of SNS grew that covered a variety of different, larger communities or particular

interests (ranging from ethnical community building such as sports, music, travelling, business contacts, to online dating). In the early 2000s the music-centred network MySpace was the most popular SNS (cf. Boyd/Ellison 2007; Strauß/Nentwich 2013; Leimbach et al. 2014). This rapidly changed in 2004 when Facebook came up which today is the most popular social media service followed by Google+.[49] Besides these dominant players a number of specialized SNS exist, i.e., for friend finding, dating, job seeking, education, professional business contacts as well as for science and research (Nentwich/König 2012; Leimbach et al. 2014). Furthermore, with their rapid evolution and expansion of Web 2.0, other social media services such as video portals (e.g., Youtube), micro-blogging services (Twitter), news aggregation sites (e.g., Reddit), photo sharing (e.g., Instagram), mobile messengers (e.g., WhatsApp) become increasingly linked or embedded to SNS environments. In this regard, social media technology transforms towards a (quasi-centralized) platform with a number of integrated services and applications (cf. Gillespie 2010).

Today, SNS, and particularly Facebook as dominant platform play an important role in social mainstream with many effects on the interplay between society and ICTs. Several studies observe positive impacts of social media on collaboration, social learning and collective action, community and social capital building, or on participation and activism, stimulating public discourse (e.g. Steinfield et al. 2008; Wimmer 2009; Baringhorst 2009; Gillespie 2010; Redecker et al. 2010; Benkirane 2012; Dahlgren 2013; Boulianne 2015). Critical studies highlight inter alia issues of privacy and trust, manipulation, power and surveillance, the political economy and digital labour, or the blurry boundaries between public and private spheres (e.g. Acquisti/Gross 2006; Debatin/Lovejoy 2009; Leenes 2010; Pariser 2011; Strauß/Nentwich 2013; Fuchs 2015; Fuchs/Sandoval 2015; Milan 2015). Recent studies also deal with the relation between social media and the "fake news" debate amplified by the US elections in 2016 (cf. Allkott/Gentzkow 2017). As regards the widespread use of social media, a number of studies show that the main motivations of SNS users are to connect to others, maintain contact and relations with friends, relatives and acquaintances, socialise and participate in community building (cf. Brandtdaeg/Heim 2009; Wimmer 2009; Smith 2011; Singh et al. 2012; Leimbach et al. 2014). Social media thus fit perfectly to the societal need to communicate and exchange with others. A further aspect concerns the entertainment value of social media as motivational factor (cf. Kim et al. 2011; Matikainen 2015). In this regard, SNS usage is partially comparable to traditional mass media consumption such as watching television: users browse through social media content in a similar way than zapping TV channels for pastime and entertainment. This is, though, only one usage factor amongst many others. In social media services primarily covering professional or special interests (such as for business networking, science and research) other usage rationales are likely. Furthermore, the different motivations are interrelated and thus usage patterns cannot be reduced to merely one issue. According to mass communication theory (cf. McQuail 2010), there are four main motivational factors of media usage, i.e., gathering information, entertainment, social interaction, and personal

[49] According to rough estimations as of the first quarter of 2017, Facebook had over 1.9 billion users worldwide; in 2010, it were about 600 million (Statista 2017c). As of September 2016, Google+ was assumed to have about 375 million users (StatB 2016).

identity representation. These factors and their interplay are visible in social media as well. Furthermore, with its enormous outreach, social media gains increasing attention in public discourses and affects the role of traditional mass communication channels (such as television, radio or newspapers) (cf. Stefanone et al. 2010). In this regard, social media may be seen as a relatively novel, quasi-individualized form of mass media although there are many differences in particular: traditional mass media provide unidirectional communication to a mass audience of individuals. Social media offers a variety of options for many-to-many communication and interaction that can be targeted at specific persons as well as groups to foster involvement in discourses, campaigns, issues etc. Moreover, being a sort of semi-public space (Strauß/Nentwich 2013), social media allow users to present themselves and perform their identities in a variety of ways. In this respect, "social media function as performative 'front stage' for the self" (Milan 2015: 7). While individuals basically use social media for social reasons, there is a strong economic rationale behind the provision of these services. The relatively low threshold to use social media and its high degree of interactivity facilitate easy (re-)distribution and co-creation of content as well as its re-contextualization, i.e., to embed existing content into new contexts. These additional values contribute to the popularity of SNS and other social media. Individual users (directly and indirectly) provide vast amounts of information about their personal details, preferences, contacts, networks and modes of interactions within these networks. Global players such as Facebook or Google employ a number of tools such as the social graph (see next sections) to analyse and integrate this information into other applications which are also provide to external entities. This information is highly valuable for business and makes social media a very attractive platform for many economic actors to promote their products and services.[50] Hence, a number of businesses, software providers, marketing companies etc. have a strong focus on social media channels to exploit this information for targeted, personalized marketing, service-oriented business models, nudging etc. This includes viral marketing which is well-supported by social media "as the community element embedded in them makes it convenient to transmit the marketing message to a large group of people" (Kaplan/Haenlein 2011: 255). In particular the many young people being active social media users are a very lucrative target group of viral marketing campaigns (e.g., for the music industry, to name just one of many examples). Consequently, celebrities are well-represented in a broad scope of social media channels to create and maintain their brand images (cf. Kaplan/Haenlein 2012). The commercialisation or economization of social media has impact on its users as the information presented to them is filtered and preselected for direct and indirect advertisement and marketing in order to stimulate consumption. As a result, individuals mostly receive information based on their (assumed) interests, while other information which is assumed to be irrelevant is filtered out. Consequently, SNS users are in a "filter bubble" (Pariser 2011). Hence from a business perspective, social media is a highly valuable personalized marketing tool. Platform providers argue that tailoring information to individuals improves user experience. However, from an individual's perspective, this

[50] Facebook, for instance, encourages advertisers to develop and integrate apps into social media for monetization of user data: https://developers.facebook.com/products/app-monetization/

may be perceived as form of surveillance, censorship and manipulation. Naturally, there is thin line from one to the other (for a further, more thorough discussion on these issues see Chapter 5).

4.3.1 Basic structure and functionality of SNS

McLuhan's (1964) claim that "the media is the message" is certainly valid for social media: already the sheer presence of a user profile represents information that is disclosed to others within the social network. Similar it is with interactivity, as every interaction of a user produces further information that can be linked to her profile. Users can, e.g., create and share information that refers to their interests, opinions, feelings, activities etc. Network effects facilitate the easy spreading of messages and informational content within the network at high potential to reach vast amounts of users. There is a broad scope of possible activities in SNS ranging from posting comments, opinions or appraisals, sharing hyperlinks, files, pictures, videos or similar, finding friends and managing contacts, synchronous or asynchronous communication (e.g., chatting, instant messaging, group discussions), promoting events and so on. The basic features and characteristics of SNS thus include novel forms of self-representation, contact management, community building, bottom-up activities and the visualization and externalization of data (e.g., by automatically displaying connections, contacts, content etc. while interacting) (cf. Cachia 2008; Nentwich/König 2012; Boyd/Ellison 2013). Thus social media bear various pieces of information that refer to individual identities, represented by the SNS user profile. This profile also has a unique identification number (e.g., the Facebook user-ID), e.g., used in internal and external applications of the SNS. Consequently, individual users are identifiable by this ID as it links to their user profile. In this regard, in the sense of McLuhan (1964), the user profile is "the message" to a variety of entities accessing social media platforms to process personal information.

Although SNS differ in design and functionality, there are certain similarities as regards their main structure. The very design of SNS enables and stimulates various new modes of interaction (one-to-one, one-to-many, few-to-few) among users as well as software agents (e.g., web sites, software applications, algorithms). Figure 11 highlights some of the main building blocks of a typical SNS structure. SNS are specific, quasi-exclusive online spaces requiring user registration. Hence, they represent centralized environments which are (at least partially) separated from other domains in the Web. What contributed significantly to the success of SNS is the central user profile serving as single access point to a variety of formerly separated services, made available in one single digital environment (the SNS). The provision and maintenance of user profiles represents a form of IDM. As SNS naturally process vast arrays of identity information, they can also be seen as centralized identity platforms (cf. Gillespie 2010; De Andrade et al. 2013). From an individual user's perspective, her profile builds the core of an SNS. The user profile represents the digital identity of a user and is perceivable as a sort of multidimensional identity card that provides access to the whole set of SNS features and applications. As social media usage is usually based on the gathering and processing of vast arrays of personal information, the user profile reveals a multitude of identity

information ranging from personal details (e.g., age, gender, date and place of birth), contact information (e.g., address, e-mail, phone), record of personal contacts, relationships, group affiliations, pictures and photographs of the user, status messages (e.g., information on personal and professional relationship status, education etc.) interests, preferences (likes and dislikes), opinions, attitudes and beliefs, behaviour, activities and interactions, favourite places, the actual location and movements, participation in events, time of being online etc. Put simply, social media can provide extensive information about the broad spectrum of an individual's identity, in the present as well as in the past. Many pieces of this information are per default disclosed and thus generally available within the SNS environment. Although users have some options to regulate the degree of visibility by adjusting their profile settings, these are mostly insufficient to effectively protect privacy.[51]

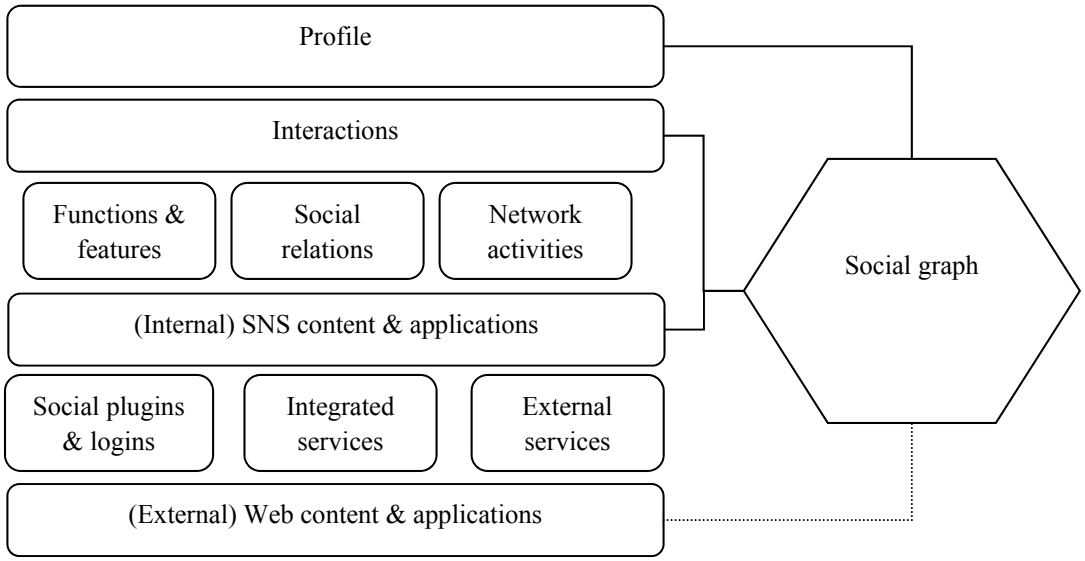


Figure 11: Main building blocks of a typical SNS structure
(own representation, adapted from Leimbach et al. 2014: 100)

Above all, interactivity is a core feature of SNS with the basic idea to enable and foster new modes of (dynamic) interactions between different entities. These entities are not merely individual users but can also be groups or organizations as well as applications or content; or in other words: informational agents. The representation and mapping of social relations and its dynamics build a central component of the manifold functions and activities available via the network. Automatic tools inform users about activities in related domains or promote new ones to stimulate further interactions. This interactivity thus entails the production and distribution of user-generated content. Information consumption became "prosumption" as users can both consume and produce content via social media (cf. Beyreuther et al. 2013; Buzzetto-More 2013; Fuchs 2014). The production of content is not necessarily bound to proactive user involvement. Every interaction generates new content that becomes associated to one's user profile. Content emerges e.g., by posting

[51] Insufficient privacy control is a general problem also beyond social media, which is discussed more thoroughly in Section 5.4.

comments, sharing hyperlinks, photos, videos etc. with other users; but also without others involved e.g., by using applications, games, watching videos, rating content (e.g., with the "like" button) and so on. Hence, there are multiple options to share and generate content over an increasing number of integrated services, features, apps etc. These integrated services can be seen as further (socio-technical) entities (or subsystems) that are related to the user.

Two forms of content production can be distinguished: internal content resulting from resources within the SNS (interactions, applications etc.) and external content produced by other sources (e.g., web sites, online services, games, apps and so forth) from the Web outside the SNS environment. While initially, SNS were relatively closed spaces, separated from the "outer" Web, new features enabled the integration of non-SNS domains. External content is increasingly integrated and pulled into the SNS by linking to external services. For developers, most SNS provide application programming interfaces (API) to integrate external applications ("apps"). On Facebook, these apps are often entertainment-focused (e.g., games such as Farmville, Angry Birds, Pokémon, or quizzes, puzzles, music applications, shopping or travelling apps etc.) and (directly or indirectly) driven by commercial interests.[52] Furthermore, there is a special form of technology for the integration of external context: the so-called social plugins[53]. A social plugin is a standardized micro-program that enables to establish a connection between a social media site and other Web content. Prominent examples are the well-known "like", "share", "follow" or "send" buttons of Facebook (as well as of Google), which today are integrated in many Web sites. These features allow users, for example, to express their opinion or share content via a simple click. In the background, they gather and process additional user information also from external sources. Every user interaction with a social plugin such as clicking a like button is traced. If the user has a SNS profile, this information becomes integrated into the user profile. Otherwise, the information is collected by other means (such as by tracking cookies). The social media platform thus gathers information from members and non-members alike.[54] Hence, social plugins enable to track user activity also beyond the systemic boundaries of a social media platform.

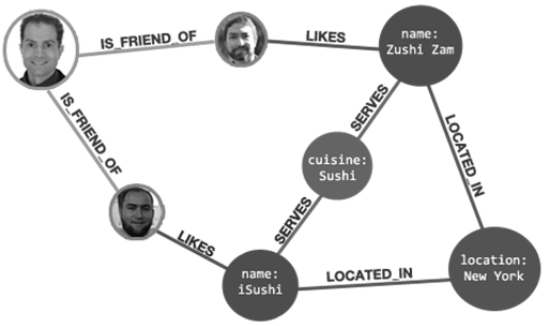
Particular forms of social plugins enable individual users to access other web services with their social media profiles (prominent examples are Facebook Connect, Google, LinkedIn or Twitter Sign In[55]). These forms are also known as "social login" or (cf. Robinson/Bonneau 2014; Gafni/Nissim 2014). Social logins represent a form of IDM whereas social media platforms act as central identity providers to external sites. External sources applying a social login can access user profile information of the SNS platform which basically contains name, profile photo, gender, user ID, associated networks and list of friends. Depending on the privacy settings of a user, further information such as status

[52] A variety of apps is, e.g., available via Facebook's App center (https://www.facebook.com/games/), Google's play store (https://play.google.com/), or Twitter's site for app management (https://apps.twitter.com/)
[53] E.g., from Facebook https://developers.facebook.com/docs/plugins/, or Google https://developers.google.com/+/web/
[54] For instance, in 2016, Facebook announced to intensify tracking of non-members for targeted advertising. See, e.g. (Toor 2016).
[55] Facebook Connect/Login: https://developers.facebook.com/docs/facebook-login, Google Sign In: https://developers.google.com/+/web/signin/, Sign In LinkedIn: https://developer.linkedin.com/docs/signin-with-linkedin, Twitter SignIn: https://dev.twitter.com/web/sign-in

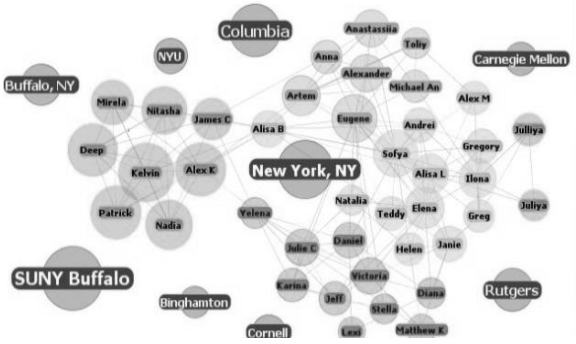
updates, content posted, comments, likes etc. can be accessed by external sources (cf. Egelman 2013). Altogether, the SNS environment stimulates interactions between personal and non-personal entities within the system but also with external systems.

4.3.2 Social graphs and the mapping of social relations

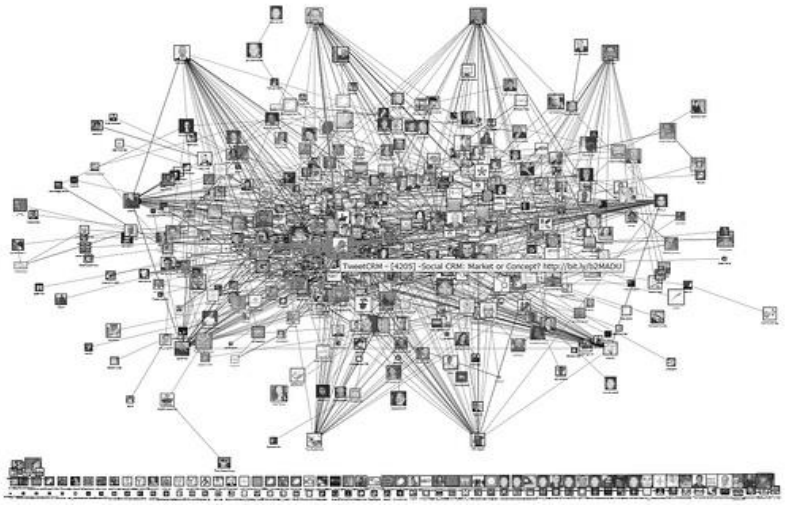
The interplay of the outlined main building blocks yields enormous amounts of information about users' identities being processed by social media. These extensive arrays of information feed into social graphs which are applied to analyse and visualize this information. Social graphs make use of network and mathematical graph theory. A social graph aims at identifying and mapping the number of entities (users, content, applications etc.) and their connections among each other in the network. The social graph is thus a dynamic form of modelling and visualizing the emergence of network relations and interactions. For instance, the number of connections of an entity to others affects its relevance. This can be visualized by nodes or hubs with different size or structure. Figure 12 below shows some simple examples of social graph visualization. For instance, Jin et al. (2013) differ between four general types of social graphs: (1) friendship graph to show to whom and how a user is related to, (2) interaction graph, i.e., to model user interactions with persons, content, applications etc., (3) latent graph to reveal latent forms of interactions e.g., site or profile visiting, (4) following graph to visualize the share of followers and followees such as in micro-blogging services. The exploitation of social graphs is a core instrument in the toolbox of data mining (see, e.g., Leskovec et al. 2014).



Friends, their location and favorite food. (Source: https://neo4j.com/blog/why-the-most-important-part-of-facebook-graph-search-is-



Graph showing connections of persons, their related social cliques and universities. (Source: http://www.touchgraph.com/facebook)



Graph showing twitter users who posted the term" social CRM" (Source: http://www.smrfoundation.org/2010/05/18/mapping-connections-among-twitter-users-who-tweet-crm-scrm-and-social-graph/)

Figure 12: Simple examples of social graph visualisations

Given the extensive information available in social media, there are many further types of social graphs that traverse this information and allow gaining deep insights into user relations, interactions, activities, behaviour etc. While social graphs are primarily employed by providers to analyse their networks, some features are also publicly available or provided to third parties with commercial interests. Facebook, for instance, provides a particular search function[56] that includes many options to conduct particular searches: e.g., for persons with particular interests, events visited, content liked etc. There is also a number of additional tools to explore and visualise social graph data.[57] Via APIs, software developers have many options to use these features (e.g., the open graph protocol[58]) to link other web content outside an SNS with a social graph. The manifold search functions of

[56] http://search.fb.com/
[57] See, e.g.: 6 Facebook Search Engine & Data Visualization Tools, http://www.toprankblog.com/2010/08/6-facebook-search-engine-data-visualization-tools/
[58] https://developers.facebook.com/docs/sharing/opengraph https://developers.facebook.com/docs/graph-api/

the graph can be integrated into other applications, e.g., for profiling or customized marketing. Besides other things, the graph protocol also allows setting predefined actions based on usage behaviour: for example, when a user uploads a photo, a particular action can be set such as the presentation of a tailored advertisement suitable to the photo. There is a variety of similar options to exploit user information for customized profiling, marketing or other purposes which makes the social graph a very powerful tool. Hence, Social graphs bear manifold information about the relationships of an individual. Having such information also enables to explore relationship patterns of non-members of a social media platform (as e.g., Horvát et al. 2013 demonstrate). This enables to map individuals' social relations, interactions, interests, activities, behaviour, movements etc. on a global level. Given the networking structure and the various ties referring from and to a user profile (e.g. contacts or other related entities), information about a particular person can be gathered via these ties. This can undermine the privacy settings of the person concerned (Bonneau et al. 2009). Currently, due to its enormous popularity and accordingly high user rates, Facebook's social graph contains extensive datasets about social interactions and networking structures of individuals worldwide. But Facebook is only one prominent example among others. Several other technology companies have similar approaches, such as Google's knowledge graph[59] or the identity graph of database provider Oracle (Oracle 2015; see also Section 5.1.2).

From a theoretical stance, social media sheds new light on classical theories about networks and social interactions. Besides graph theory, the social graph also grounds on other theoretical concepts such as Milgram's "small world problem" (Milgram 1967); also known as the "six degrees of separation" which claims that worldwide, each person is related to each other over six contacts.[60] It is thus no coincidence but rather reminiscence that one of the first SNS was labelled sixdegrees.com. Milgram's small world theory was inter alia criticized for its bias and lack of sound empirical evidence. While there may be many small world phenomena as part of a large complex world, these small worlds are not necessarily connected (cf. Kleinfeld 2002). Internet communication and social media alleviate the exploration of the small world hypothesis. Some evidence was found e.g., by Leskovec and Horvitz (2008) who studied the relations of 240 Mio. instant messenger accounts. According to their results, users are related to each other over approx. 6,6 knots. However, even though other studies may found proof for fewer degrees, social reality differs from these mathematical approaches. Hence, "we may live in a world where everyone is connected by a short chain of acquaintances, but it is hard for most people to find these connections" (Kleinfeld 2002: 66). Nevertheless, against the background of social graphs and the entailed options to analyse and map social network relations, it may be less relevant today whether the chain of connections is more or less than six degrees of separation. Another classical theory dealing with the quality of connections in social networks is the theory of "the strength of weak ties" by Granovetter (1973). According to

[59] https://www.google.com/intl/es419/insidesearch/features/search/knowledge.html

[60] An entertaining approach to address the small world phenomenon is the so-called Bacon number, referring to the US actor Kevin Bacon. The number indicates the distance to the actor: the higher the number the greater the distance. For instance, an actor who did not occur in a movie with Bacon but played with an actor who did has the number 2. In the sense of "I know somebody that knows somebody … that knows Kevin Bacon". For an online tool exploring the Bacon number see, e.g., https://bacon.mybluemix.net/

this theory, the strength of a tie is determined by several factors, i.e., time, emotional intensity, intimacy, and reciprocity. Close relationships (e.g., between friends and relatives) rather represent strong ties that are usually trustworthy, long lasting and stable. Strong ties thus contribute to the stability and consistency of a network. In contrast to that, weak ties are rather loose connections with a higher dynamic. Thus, the growth of a social network strongly depends on weak ties as they function as bridges across different network domains or nodes. Compared to strong ties, weak ties allow distributing information more widely across greater social distance. Information shared by strong ties "is much more likely to be limited to a few cliques than that going via weak ones; bridges will not be crossed" (ibid: 1366). However, this information may also be more stable and reliable. In this regard, strong ties contribute to a fluid flow of information in a network. Evidence for the strength of weak ties can be found, for instance, in the micro-blogging service Twitter. This service provides easy forms of information distribution via online media. The relevance of information (the tweet) depends on the number of followers of a twitter account: the more followers, the more likely it is that other network domains or communities are reached. The number of followers is highly dynamic and may change quickly. Similar it is with contacts in SNS. A high number of weak ties (e.g., the number of Facebook contacts, twitter followers etc.) contributes to the size of a social network. However, it may have negative impact on the quality of the network (e.g., its reliability and trustworthiness) if ties are too weak. Consequently, connections may turn into "absent ties", i.e., "ties without substantial significance" or become irrelevant (ibid: 1361). Thus, the stability of a network is strongly depending on the existence of strong ties and its interplay with weaker ones. The strength of a tie can provide e.g., information about how an individual is embedded in a social network.

In total, these issues highlight that social media and similar applications are powerful tools, providing an enormous depth of potential insights into individual identities, including their relationships and interactions even beyond the systemic boundaries of the primary platform. Social graphs and the like demonstrate how this information can be exploited to gain deeply intrusive identity profiles. They basically make use of the connecting function of identity (see Section 3.1) and its ability to create different ties.

4.3.3 Expanding social media identities

All in all, social media have significantly affected social interactions and entail many societal impacts. They have similarities with Floridi's notion of the infosphere (see Section 4.1) as everything in their realm is interactively networked. Moreover, even content outside the social media environment becomes increasingly absorbed by social plugins, logins and the like. This results in extensive information collections providing deep insights into individual users' identities. As shown, by their very design, SNS and other social platforms not only stimulate social interactions online but significantly foster digital representation of individual users, groups and usage of identity information. Some scholars argue that social media enhance individuals in their self-representation with options to choose which personal details to reveal and which to conceal. Users can construct their online identities or present themselves in ways they would like to be perceived by others

(cf. Salimkhan et al. 2010; Ellison 2013). However, in fact there are very limited to options for individuals to keep their (real) identities private. Users who provide little or no personal details can hardly benefit from social media features. Furthermore, every movement or activity in a social media environment creates information that relates to the user. Consequently, the more a user interacts, the more information about her is collected. Moreover, social media encourages and seduces users to present themselves, their interests etc. and disclose parts of their identity.[61] This already starts during registration as most social media platforms have a real name policy to prompt users to enter their real names (cf. Boyd 2012; De Andrade et al. 2013; Edwards/McAuley 2013; Whitley et al. 2014). For example, Facebook's name policy states that: "Facebook is a community where everyone uses the name they go by in everyday life. This makes it so that you always know who you're connecting with and helps keep our community safe. (…) The name on your profile should be the name that your friends call you in everyday life. This name should also appear on an ID or document from our ID list."[62] Google has a similar policy for its social network: "Google+ profiles are meant for individual people. That's why we recommend using your first and last name on your profile. It will help friends and family find you online, and help you connect with people you know."[63] Although the providers failed in enforcing a strict real name policy (cf. Edwards/McAuley 2013), it is widely implemented as the majority of social media accounts reveal the real names of their holders. Several studies show that online profiles mostly provide a relatively detailed presentation of their users (cf. Weisbuch et al. 2009; Ellison 2013; Madden et al. 2013). As shown in the previous sections, user profiles build the foundation of social media, usually containing various forms of identity information, often including personal images and photographs, date and place of birth, area of living, education, profession, relationships, friends and acquaintances etc. Being semi-public spaces, social media entail many ways of direct and indirect disclosure of this information (Strauß/Nentwich 2013). As every interaction of a user with his profile is persistent, replicable, scalable and searchable in the social media environment (Boyd 2010), one's identity representation is constantly extended. For users, this is e.g., observable in Facebook's "Timeline" feature (Panzarino 2011; Kupka 2012) which enables to browse through one's activities ranging from his "birth" in Facebook, i.e., registration and first login, up to the very last action taken in the network environment. Hence, every piece of information provided to the platform is stored and categorized for further processing. For instance, every time a user uploads an image to Facebook, an automated algorithm integrates tags into these images in order to enrich them with additional, searchable information (e.g., persons visible, description of the scene such as "eating", "sitting", meta-description of the landscape such as "beach" etc.)[64]. These keywords provide additional ways to explore content, relationships, activities etc. of users. Furthermore, the social graph provides many options to exploit social media content and

[61] As e.g., observable in Facebook where users are prompted to enter their relationship status, employers, interests, feelings, favourite music, films, books, visited events etc.
[62] Facebook's real name policy: https://www.facebook.com/help/292517374180078
[63] Google+ profile name information https://support.google.com/plus/answer/1228271?hl=en
[64] An open source browser add-on called "Show Facebook Computer Vision Tags" visualizes the tags of an image https://github.com/ageitgey/show-facebook-computer-vision-tags http://nymag.com/selectall/2017/01/see-what-facebook-thinks-is-in-your-photos.html

re-use it in other application contexts. Consequently, there are quasi-unlimited contexts in which users' identity information can be processed.

A number of developments contribute to a further expansion of the social media landscape. As outlined above, personalization of services and profiling play an important role for the success of social media. Their providers are not merely acting altruistically but their services are key parts of their business models and thus social media also represent commercial infrastructures. In general, social media mirror the peculiarities and dynamics of software development practices that became widespread with Web 2.0: the quick release of new technologies, applications, features, modalities etc. accompanied by continuous adjustments and reconfigurations depending on user behaviour and feedback resulting in a sort of "perpetual beta" status (cf. O'Reilly 2005). In this regard, social media can be seen as a test bed or playground for developers and commercial stakeholders. The Internet economy profits in many respects of social media e.g., for CRM, customized advertising etc. which is a main reason for the significant increase in personalization of online services in the last decades. While several years ago, the outreach of SNS used to be limited to entities within the social media environment, it grew with the increasing societal relevance and further integration of other applications. As shown, there are plenty of possibilities and tools to use social media and exploit the vast arrays of information. APIs, the social graph and social plugins provide manifold ways to integrate and link applications and services with social media environments and vice versa. This amplified the further expansion of social media towards "outer" spaces in the Web. As a consequence, social media win additional means to gather identity information from external sources and thus gain an even more detailed picture of individual users. This is boosted by the growth in mobile computing (smart phones, tablets etc.) that enabled portable usage of social media and thus serves as unfolding platform technology for a variety of applications (e.g., the integration of apps). Consequently, the social media universe further expands with mobile social media usage, which is quickly increasing and a promising market (Perrin 2015; Buhl 2015). Marketers assume that about 80% of social media usage is mobile (see e.g., Sterling 2016). This is accompanied by novel services appearing that are often absorbed in no time by the major players, such as Facebook's takeover of the photo sharing application Instagram in 2012 or of the popular mobile messenger "WhatsApp" in 2014 (O'Connell 2014). A result of these takeovers is that Facebook gains additional user data from these services (such as users phone numbers), and user profiles grow further. Amongst other things, user data from WhatsApp such as mobile phone numbers are forwarded to Facebook (Gibbs 2016a). Other activities of Facebook concern the integration of payment services into its platform (Constine 2016; Russell 2016). A further trend is observable in the field of identity management where providers attempt to enter the IDM market. Social media platforms used to apply softer forms of IDM limited to their own domains (e.g., to provide and manage user accounts). However, recent developments such as the use of "social logins" (as outlined in Section 4.3.1) led to a significant extension: larger providers such as Facebook, Google, LinkedIn or Twitter) provide tools to integrate their login procedures and thus users' profiles into other (external) web applications. Figure 13 shows typical examples of social logins as embedded in many websites:

Figure 13: Google+ and Facebook social login buttons

Other websites, services, platforms etc. can use these features and outsource the management of their users' identities to social media platforms, who position themselves as central identity providers. As a result, user identities can be integrated in other contexts whereas in return, the platforms gain additional identity information from these external sources. With these approaches, social media platforms have the potential to become pervasive online identity systems. Besides these approaches to integrate social media identities in other web domains, there are also trends to apply them in real world IDM contexts. The growing relevance and widespread diffusion of social media partially lead to "growing pressures to use this (social network) data about connections and endorsements as part of the identity proofing processes for more formal identification processes" (Whitley et al. 2014: 26). Although the creation of online profiles currently does not require formal identity checking, there are approaches to change this. One attempt came from Facebook which attempted to enforce its real name policy by de-activating user accounts and prompting users to prove their real identities by providing official ID (e.g., passport or driving license) to the social network. After heavy user protest, Facebook conceded to have made a mistake and apologized (Raeburn 2013). However, the real name policy is still valid but not strictly enforced anymore (cf. Edwards/McAuley 2013). Attempts like this do not merely come from social media providers. In some countries, similar ideas come from political decision-makers. Considerations include, for example, to link social media accounts (such as Facebook) with national identity documents or governmental eIDMS approaches (cf. De Andrade et al. 2013; Martin/De Andrade 2013; Whitley et al. 2014). Ideas like this raise a number of critical issues, such as: economic dependency and risks of monopolization of identification, threatening its role as a genuine societal and governmental function, lack of control over IDM, problems of data ownership, international legal and regulatory issues, lack of network and technology neutrality, lack of trust and liability, and extensive security and privacy issues; to name just a few (De Andrade et al. 2013; Whitley et al. 2014). Regardless of these risks and problems, there are ongoing trends to merge social media and real world identities. This is inter alia observable in countries like China with problematic situations concerning human rights: Chinese authorities aim at legally forcing citizens to use their real names in social media (Chin 2015). But similar trends exist in other countries as well: such as the registration of voters via Facebook in the US state Washington (Farivar 2012); the use of Facebook accounts for official identification to access public online services in the UK (Lee 2012); discussions about integrating the governmental eID system into Facebook in the Netherlands (Martin/De Andrade 2013); recent plans of the US department of homeland security[65] to use social media accounts for border control, i.e., travellers to the US should expose their

[65] The US department of Homeland security already conducts surveillance of social media such as Facebook, Twitter and online blogs as well as news organizations. The US privacy center EPIC (electronic privacy information center) pursues a lawsuit against the department: http://epic.org/foia/epic-v-dhs-media-monitoring/ The monitoring activities are outsourced to a private contractor.

social media IDs (e.g., on Facebook or Twitter) to law enforcement (Gibbs 2016b; McCarthy 2016); or considerations of the European Commission to link social media and governmental IDs to stimulate the EU digital market (Best 2016; EU-C 2016b). These developments indicate further trends of extending digital identification practices and their application contexts.

4.4 Transition paths of digital identification

As shown, the representation and processing of identity information, i.e., identification practices are part of wider socio-technical transformation pattern in the realm of ICTs. Explored through the lens of a metasystem transition (as outlined in Section 4.1), different development stages of these phenomena can be grasped from a wider, system-theoretical perspective. A transition is here not seen as a shift from an old to a new system but rather as a continuous, dynamic transformation pattern where different technological developments influence each other. Indeed, the mapping of socio-technical developments against the main MST stages does not allow to unambiguously identifying the advent of a particular phase as well as the transition from one phase to the next. But it is a useful heuristic tool to structure and grasp the dynamics of socio-technical change. As outlined in Section 2.2.1, a transition is an enduring evolutionary process which involves various transformations resulting from dynamic interactions between socio-technical niches and regimes (cf. Kemp et al. 2001; Geels 2004; Geels/Schot 2007). Technological change starts in niches, i.e., "limited domains in which the technology can be applied" (Kemp et al. 2001: 274), where new socio-technical functionalities are explored and employed in specific contexts, which can lead to specialization and further expansion of the technology. With wider usage and diffusion (e.g., by entering new economic markets), a technology can become established in one or more regimes, which then entails wider societal impacts such as changing user practices. According to (Kemp et al. 2001: 273), technological regimes "are configurations of science, technics, organizational routines, practices, norms and values, labelled for their core technology or mode of organization". A (socio-technical) regime thus comprises the rules and dynamics that result from the complex of processes, organisational and institutional settings, infrastructures etc. and their dynamics that shape the socio-technical configuration of a technology or system (cf. Kemp et al. 2001; Geels/Schot 2007). Each of the domains involved in a transition has their own dynamics of self-organization processes which enable as well as constrain different development paths (Geels/Schot 2007; Hofkirchner 2013). This means that a technological development may enable a new practice but at the same time, complicate others. The dynamic linkages between different system elements create a semi coherent structure as several tensions may occur such as competing designs, practices, social and economic interests, policies, neglected consumer rights, legal and ethical issues etc. This may lead to further (intended or unintended) changes regarding structure and organization, re-configurations or adjustments which affect the design and usage patterns, and thus the societal impacts of a technology.

As shown in the previous Sections, similar dynamics of change can be found in the socio-technical transformations related to ICTs. The progress/development of ICTs can be seen as an evolutionary process where socio-technical systems become increasingly integrative parts of society. Furthermore, these different systems are interconnected by digital information. Conceptualised as MST, ICTs (and primarily the Internet) are perceivable as a set of intertwined socio-technical systems or metasystem which allows its multiple subsystems to interact, combine and integrate information from different domains. Formerly rather isolated, siloed systems (e.g., specific websites, single online services, applications etc.) partially transformed towards highly networked, interactive platforms (such as Web 2.0, social media, e-commerce platforms etc.) which deeply and constantly influence many domains of society as being an integral part thereof. As shown, these transformations also affect the way identities are represented and managed online as well as the related identification practices. Hence, there are certain transformation patterns of digital identification observable: in the early days of the WWW, before Web 2.0, online applications had no or limited relations and were mostly separated from other application contexts. User profiles and personal information were widely isolated, not extensively networked, and rather used on occasion but not as part of a strategic IDM concept. With increasing ICT diffusion and particularly Web 2.0 and social media entering mainstream, interactions between different application systems increased. In parallel, personalized services gained momentum, requiring or suggesting individual users to provide personal details. At the same time, stimulated by different economic and security interests, IDM emerged with attempts to unify and standardize identification and authentication procedures, enable interchangeable user profiles etc. Hence, personalization, digital identity representation and identification expanded, and became more networked.

Today, there is a broad range of different services, platforms and technologies covering various domains of every-day-life. To some extent, ICTs thus represent extensions of personal identities as individuals present themselves in online profiles, and use personalized services as well as technical devices. Entailed is the processing of extensive amounts of identity-related information; as exemplified by social media platforms and other global internet companies. Boosting smartphone usage underlines the convergence of physical and digital environments with integrated digital identity representation. This is accompanied by accelerated trends of further networking and integration of different online applications and services (social media, IDM systems, apps, social plugins and logins etc.). There are many tendencies to extend to scope of social media: with social plugins and logins becoming integrated into other systems, the platforms themselves gather information also from applications outside their environments. As a consequence, user profiles are enriched with information from external usage contexts. In this regard, there are meta-profiles emerging, which is also observable in the realm of Web search provider Google which kept user profiles from all its different services (ranging from search, e-mail, geo-location services, social media, chats, image processing, bulletin boards etc.) separated from each other. Since 2012, Google aggregates all user data (cf. Suhl 2012; Reitman 2012) from its broad scope of services and applications (e.g., web search, Gmail, G+, Maps, Google Docs etc.) into one centralized meta-profile. This meta-profile thus contains extensive collections of individual online

activities. But also besides Google and Facebook, several other big internet companies (e.g., Amazon, Microsoft, Apple) have extensive amounts of information from their users. Technologies such as the social graph enable to analyse and visualize aggregated user information from multiple sources. Moreover, they allow searching for particular individuals, based on their relationships, activities, interests, behaviour etc. and to embed this functionality into other applications as well.

These developments do not imply the emergence of a sort of global meta-ID system, or universal identification mechanisms. Nevertheless, identity information is broadly and increasingly available, can be aggregated from multiple sources and thus allows for sophisticated meta-profiling which may have similar effects. Moreover, IDM systems and identification mechanisms generally increase with trends towards (quasi-)centralization. There are also tendencies to integrate identification mechanisms into applications which actually do not need formal user identification. For instance, real name policies in social media and other online platforms (cf. Whitley et al. 2014) or the use of governmental eIDs for libraries, public transport, online games, chat rooms etc.; as given in some countries (cf. Bennett/Lyon 2008; Lyon 2009), as well as the previously outlined trends to expand the scope of social media identities. With social logins, global Internet companies (e.g., Facebook, Google+, LinkedIn) began to position themselves as central identity providers for online services: they offer particular login services to integrate their user profiles into websites and applications (e.g., "Login with your Facebook account"). In return, additional web content and user's identity information can flow into the social media platforms. Furthermore, the broad availability of identity information and user profiles triggered desires to combine social media accounts with formal identification. The increasing trends to use social media profiles for identity verification and, e.g., link it with national identity procedures including governmental IDs highlight that digital identities are more and more entering "real world" contexts and vice versa. Also smartphone technology is increasingly linked to IDM as they become popular, multifunctional carrier devices for identification and authentication. This is another example for technological convergence and integration. For instance, e-banking systems employing smartphones as authentication tokens, e-commerce platforms (e.g., Amazon and others) suggesting to provide the mobile phone number as user credential. Further examples can be found in the field of biometrics (see Section 5.3.3).

Hence, there are many empirical examples for the increasing convergence between analogue and digital environments where identification is involved. This is not just a consequence of technological progress but results from several socio-technical transformation processes. The outlined developments involve a number of societal transformations and patterns of change. Digital identification emerged in several technological niches and incrementally gained momentum which contributed to increasing interactions between different regimes. This is accompanied by various transformations and re-configurations in social, economic and political domains. The introduction of IDM systems entails socio-technical change processes in the public and private sector, alter the relationship between citizens and their governments as well as between customers and businesses; the dynamics of social interactions in Web 2.0 and social media altered socio-technical interactions and identity representations of individuals in online environments.

As shown, the increase in digital identification is driven by a number of interrelated socio-economic and political factors with two dominating domains (or regimes): digital economy and security. Economic considerations include e.g., to foster service efficiency, stimulate market development, CRM, behavioural advertising, service-for-profile business models etc. Security objectives range from providing secure online transactions, reliable identification and authentication up to combating fraud, crime and terrorism, i.e., national security. Considering the self-dynamics of both regimes (economy and security) which may stimulate further identification mechanisms, we can speak of economization and securitization. In our case, economic and security rationales have an influence on the implementation and use of digital identification mechanisms (as discussed more in-depth in the next chapter). In general, transitions include enabling as well as constraining mechanisms (cf. Giddens 1984/1997; Hofkirchner 2013). Socio-technical practices change, new ones emerge; and these changes also entail increasing pressure on existing societal functions which may reinforce existing societal conflicts. Amongst other things, ICTs enabled new modes of information processing, communication, interaction and networking. They can be seen as socio-technical artefacts which extend the representation of their users' identities and thus enable novel forms of identification. But to some extent, they also constrain individual privacy. The emergence of digital identification puts high pressure on the effective protection of privacy, or more precisely the boundary control function inherent to privacy. The increasingly seamless, frictionless processing of identity information aggravates privacy risks. Digital identification is thus confronted with a certain control dilemma which is explored in the following Sections.

CHAPTER 5

The privacy control dilemma of digital identification

The previous Sections analysed the emergence of digital identification. This Section sheds light on the privacy implications and challenges resulting from this development. As elaborated in Section 3, there is a naturally close relationship between identity and privacy. Privacy is vital for identity-building and autonomy. With its inherent boundary control function (see Section 3.2.2), privacy enables individuals in regulating and self-determining the extent to which they provide informational details of their identities to other individuals or institutions. Identification includes control functions as well but for other purposes: identification is a control mechanism that serves a connecting function between two or more entities. It contributes to gain certainty about particular characteristics of a person making her distinct from others. A central aim is the controlled processing of identity information. This aim is inherent to digital identification and IDM approaches. However, less in the sense of fostering privacy controls for the individual concerned but rather in the sense of ascertaining her identity. Identification processes are an essential part of socio-technical systems in many respects, vital for the functioning of society, the state and the economy. Means of identification are crucial to build trust among different entities interacting with each other in social, political and economic contexts, support administrative procedures, connect different technical systems etc. Hence, identification is an important instrument of governance but it also represents a control mechanism (cf. White 2008). In general, the striving for control involves a quest for security and stability of a matter. This is valid for political and administrative power, national security as well as for economic growth. Identification is a means towards this quest with the basic aim to reduce uncertainty and improve security in particular settings. It can contribute to build ties in social, political, economic as well as in technical contexts (see Section 3.1). As outlined in the previous sections, the field of IDM is a consequence of informatisation aiming at regaining control over digital identity information processing and tackling its increasing complexity in order to enable secure identification for a variety of purposes also in digital environments. The state also governs by identity (Amoore 2008; Whitley et al. 2014) and governments set up IDM systems to integrate standardized citizen identification in services of public administration, stimulate e-commerce and the development of digital markets. Security authorities apply identification mechanisms to control individuals and protect the interests of their states. Economic actors, online platforms, social media etc. use digital identification to protect and control login procedures, to provide personalized e-business and e-commerce services for CRM as well as to create new business models[66]. A crucial

[66] Identity information is also seen as lucrative business factor, as, e.g., indicated by these promotion articles: "Monetizing identity – pay by face", TMForum, March 25 2014 https://inform.tmforum.org/nfv-it-transformation/2014/03/monetizing-identity-pay-by-face/ or "How to monetize your customer data", Gartner, December 10 2015, http://www.gartner.com/smarterwithgartner/how-to-monetize-your-customer-data/

aim in each case is to govern the information flows related to the digital identities of their holders. These information flows serve many commercial interests, such as CRM, monitoring of potential customers for market research, targeted advertising and behavioural profiling. Hence, there is variety of different actors who benefit from identification practices for various purposes ranging from fostering administrative procedures, stimulating economic development as well as for political objectives. In all these approaches, the processing of identity information is primarily controlled by institutional/organizational actors (e.g., public sector institutions, security authorities, social media providers, businesses etc.).

Although there are several plausible reasons for identification, an imbalanced control over identity information and lacking ISD (see Section 3.2.3) from an individual's perspective hampers privacy protection. Furthermore, even when IDM is designed to enhance user control (e.g., where a person can proactively decide when to be identified) this may not prevent from misuse of information for other purposes (De Hert 2008; Strauß 2011). The information stored in centralized databases and registers applied in public and private sector may provide a detailed picture about the trails and contexts of individuals' identities. Similar is given in case of the broad array of identity information collected and processed in digital environments such as social media and the like. They serve a wide range of purposes that often remain unknown to the individual concerned. The de- and re-contextualisation of this information for other purposes than the originally intended is often hardly controllable; identification can occur in manifold ways with or without the consent of the individual. Personal information can generally flow limitless between information systems in public as well as private sector. Individuals are thus significantly hampered in ISD as regards the processing of their information. Attempts to improve control over digital information, such as IDM aiming at standardising the handling of identity information, bear some potential in this regard. However, when they mainly serve institutional entities, they contribute little to improve ISD. Especially as IDM implementations tend to neglect privacy: compared to security and economic interests, the protection of privacy is often not a primary issue of IDM (cf. De Hert 2008; Nauman/Hobgen 2009; Strauß 2011). Although some efforts in the field of governmental eIDMS exist in this regard, they have serious limits and mostly lack in effective privacy protection. While institutional entities benefit from IDM, the advantages for the individuals concerned are rather marginal (cf. Kubicek/Noack 2010b; Strauß/Aichholzer 2010); at least as regards ISD and enhanced control over her information. Hence, there is an imbalance of power between the individual and the institutional entities as regards the control over identity information. This problem can be described as a privacy control dilemma of digital identification: although IDM creates digital identification mechanisms in order to regain control over personal information flows in digital environments, it may ironically lead to a further loss of control over this information, at least from an individual's perspective (Strauß 2011). Effective privacy protection requires some friction in the processing of personal information with unlinkability as important requirement. In this regard, there is a smouldering tension between the connecting function of identification and the boundary control function of privacy requiring informational frictions.

Many socio-political, economic and technical issues shape this dilemma: in general, digital identification is an important tool of governance in the public and in the private sector. However, there are also several tensions as regards the imbalanced control over information between individuals (persons, citizens, customers etc.) and institutional entities (organizations, governments, businesses etc.) who apply identification processes. This control gap between individual and institutional entities is a general problem of imbalanced power structures inherent to the surveillance discourse, as presented and discussed in the following Sections. As will be shown in Section 5.1, there are certain overlaps between identification practices and the functions and mechanisms of surveillance, i.e., panopticism. A core issue in this regard are information asymmetries which reinforce with ICTs and digital identification. The implementation of IDM and means of digital identification is conveyed by securitization and economization of personal information which contribute to a further extension of identification practices. Conflicting interests as regards individual privacy, economic objectives and national security may complicate the challenge to balance different interests and conciliate those of the individual with those of the institutions processing her identity information. Embedded in the privacy control dilemma is an inherently conflictual relationship between privacy protection, security and surveillance which mirrors in the discourse about digital identification. Empirical results from the SurPRISE project provide insights into the perceptions of European citizens on the interplay between surveillance, privacy and security in relation to surveillance-oriented security technology (see Section 5.2). The growing concerns and fears about privacy intrusion among European citizens underline that there are several tensions between individual and institutional transparency as regards the processing of personal information. Furthermore, the design and use of ICTs reinforce these tensions as they boost (personal) information processing and stimulate a further expansion of explicit and implicit forms of identification. A central problem of contemporary privacy protection is thus socio-technically enforced identifiability (see Section 5.3). To tackle this problem requires enhanced control and informational self-determination. Section 5.4 presents and discusses the prospects and perils of privacy by design and related technical privacy control concepts in this regard.

5.1 Surveillance, identification and control

5.1.1 Overview on basic functions and practices of surveillance

In general, surveillance is a common cluster term to describe practices of observation, monitoring, and controlling individuals which involve the gathering and processing of information. Surveillance as a hierarchical modality of power and disciplinary practice is a core issue of security and surveillance studies. Research in this field mainly involves a sociological perspective on the general functioning and societal impacts of surveillance in different (public and private) domains (e.g. Foucault 1977; Clarke 1988; Lyon 1994; Haggerty/Ericson 2000; Lyon 2001; Lyon 2003; Lyon 2006; Bennett/Haggerty 2011; Ball

et al. 2012; Marx 2015; Wright/Kreissl 2015). This section includes a review of relevant literature of surveillance studies with an explicit focus in issues of identification.

Practices of identification are closely related to power structures and modes of surveillance. The gathering of "some form of data connectable to individuals (whether as uniquely identified or as a member of a category)" is a central feature of human surveillance (Marx 2015: 734). For Lyon (2009: 4) all forms of surveillance even begin with identification. But this does not imply that all forms of identification are equivalent to surveillance. Nevertheless, identification can be used for and can result in surveillance practices. Identification and surveillance can be implicitly and explicitly linked. Haggerty and Ericson (2000: 610) argue that "(...) surveillance is driven by the desire to bring systems together, to combine practices and technologies and integrate them into a larger whole". With its connecting function, identification can contribute to implement this combination of practices, technologies and systems in many respects. This contribution often occurs implicitly. Identification is thus not to be misunderstood as form of surveillance per se or as mechanism to serve surveillance purposes. However, the boundary between appropriate identification and excessive surveillance is not always clear but often ambiguous. The more intrusive security and surveillance practices become, the more individuals' identities can become affected. This section ties in here and elaborates the extent to which the control mechanisms inherent to identification processes may overlap with functions of surveillance. The basic premise is that identification, i.e., the processing of information referring to an individual, is a core process of privacy affecting surveillance. Hence, if a surveillance practice includes the identification of individuals, then privacy of these individuals is affected. Conversely, central mechanisms of surveillance (such as panopticism and asymmetry of power) can also mirror in identification practices. Before these issues are explored, the basic functions and modalities of surveillance are outlined.

Giddens (1984/1997) argued that surveillance is a basic practice of modern national states to exercise administrative power. It includes a machinery of processing large amounts of information which enables and supports control mechanisms across space and time. Correspondingly, for Giddens, surveillance is "the coding of information relevant to the administration of subject populations, plus their direct supervision by officials and administrators of all sorts" (1984: 183f. cited from Fuchs 2010). Today, this machinery is observable in many respects. A variety of institutions (e.g., governments, security authorities, intelligence agencies, businesses, public or private organisations etc.) collects and uses information about individuals for purposes of administration as well as of control. Surveillance is an intrinsic part of modern bureaucracy, but it is often more than a sheer practice of organizing information to serve administrative procedures in the realm of governance. It is inter alia employed to protect national security and can also entail various mechanisms of population control. As form of disciplinary power (Foucault 1977), surveillance provides several options to influence and manipulate individual as well as collective behaviour. Surveillance thus entails serious risks of social sorting and discrimination as, e.g., Lyon (2003) pointed out. Inherent to surveillance is the existence, creation or reinforcement of asymmetrical power between the observers and those that are being observed. This power is not necessarily exercised instantly or visibly. Though, it can

include a sort of preparatory action in terms of pre-emptive and preventive security measures (such as surveillance measures increasingly aiming at preventing crime and terrorism). This may include preventive information gathering and monitoring of individual behaviour.

Surveillance can both be enabling and constraining power (cf. Lyon 1994; Fuchs 2010). It enables as it gives power to those that conduct surveillance and allows exercising this power over those that are monitored. This power can include physical power (direct violence at worst, punishment etc.), control over information, violations of privacy, or discrimination and power of repression that leads to overwhelming self-control or self-discipline of the individuals and their behaviour being subject to surveillance (cf. Foucault 1977; Lyon 2003; Fuchs 2010; Ball et al. 2012). In many cases, surveillance is "implemented as a security mechanism so that surveillance and self-control are used as two mechanisms (one of direct violence and one of ideological violence) for reproducing and securing domination" (Fuchs 2010: 11). In totalitarian regimes, surveillance is employed to identify, classify, control, discriminate and repress civil society. There are thus risks inherent to surveillance that population control and social sorting reinforces discrimination. As highlighted by many scholars, history provides several examples for violent social discrimination and population control which drastically demonstrate how surveillance can serve destructive forces (such as in the Nazi regime, the Stasi dictatorship in East Germany, or in South Africa during apartheid). In its worst cases, this can even lead to deportation, murder and genocide with millions of people killed based on their ethnical identity. The abuse of population registration and census data played a significant role in Nazi Germany (cf. Lyon 2003; Bennett/Lyon 2008; Lyon 2009; Ball et al. 2012). It is thus essential to have effective forms of checks and balances to continuously question surveillance practices (cf. Wright et al. 2015) in order to control and stem its destructive capacity.

Irrespective of its manifold forms, purposes, and risks, surveillance generally represents a security practice aiming at reducing uncertainty and fostering control. To achieve these aims, the gathering of information is a core feature of surveillance. Although often understood in that sense, surveillance is not the nemesis of privacy per se; and vice versa, privacy was never meant to be "the 'antidote to surveillance'" Bennett (2011: 493). In the broad sense, the term surveillance (deriving from the French word surveiller which means to watch over) addresses the "close watch kept over someone or something".[67] Not every form of surveillance directly targets at individuals; such as surveillance in the field of public health, environmental protection or aviation control, where personal information is usually not gathered directly. For example, aviation control conducts surveillance to monitor planes and other objects in airspace. Privacy is not an issue of this monitoring as long as there is no processing of identity information involved. When aviation control includes the collection of information about flight passengers it represents a surveillance practice with a privacy impact. This simple example points out that the privacy impact of surveillance depends on the processing of personal information and thus the capacity to identification. Some forms of identification processes are often inherent to surveillance.

[67] As e.g., the Merriam Webster Dictionary defines https://www.merriam-webster.com/dictionary/surveillance

For instance, traffic control includes monitoring of streets in order to ensure that speed limits are kept. There is no identification process involved in this monitoring as long as number plates of vehicles are not automatically gathered. But identification is triggered when a vehicle breaks the legal speed limit. Then, the police or a speed camera automatically shots a picture of the vehicle's number plate. This picture then is used to identify the holder of the vehicle to deliver the speeding ticket. This practice thus involves conditional identification. The use of surveillance technology can complicate conditional identification when it permanently gathers information about individuals. Traffic control with automated number plate recognition without any limitation (i.e., permanent recording) is more privacy-intrusive than a practice with conditional identification. A basic example of a surveillance technology is CCTV[68] which usually captures images of persons per default. As a consequence, information referring to the identity of that person is gathered which is a form of unconditional identification. These examples point out that the modalities of identification as well as the design and use of a technology determine the extent to which privacy is affected by surveillance.

Contemporary surveillance benefits in many respects from the employment of technology as highlighted by many scholars of surveillance studies. For instance, Clarke (1988) used the term "dataveillance" which he defined as "the systematic monitoring of people's actions or communications" to describe electronic surveillance practices. For Gary Marx (2002: 12) the "new surveillance" is more extensive than traditional surveillance, characterized by the "use of multiple senses and sources of data" and includes "the use of technical means to extract or create personal data". Some years later, he defined new surveillance "as scrutiny of individuals, groups, and contexts through the use of technical means to extract or create information" (Marx 2015: 735). Examples of "new" technological surveillance practices are e.g., "computer matching and profiling, big data sets, video cameras, DNA analysis, GPS, electronic work monitoring, drug testing, and the monitoring made possible by social media and cell phones" (ibid: 735). Marx assumes that contemporary (or "new") surveillance is more extensive because of technology but he does not explain for what exact reasons. In my view, an important reason concerns the modality of information processing. The gathering of identity information basically creates the possibility of intruding into privacy. Depending from the storage capacity of a technology, the factual intrusion can happen instantly or at a later point in time. ICTs provide various ways to process information widely decoupled from space and time. This enables and amplifies surveillance practices in many respects. In its most general sense, surveillance involves the gathering of information to reduce uncertainty about an issue and thus gain in power and control. This rationale remains the same with or without the employment of technology. However, ICTs inter alia enable to gather and reproduce information remotely, decoupled from the physical location of a surveillance subject. A further aspect concerns the networking structure of ICTs that allow to aggregate different kinds of information from different domains and collecting them unbound from physical presence. Hence, through ICTs, surveillance gains an additional, non-physical dimension which contributes to extend its outreach.

[68] Closed Circuit Television

Besides these technological aspects, networking structures are also observable among the actors of surveillance. In a traditional sense, surveillance was mainly conducted by the state and public sector institutions for all kinds of administration and governance modalities including internal and external security. Contemporary surveillance is somewhat different representing a complex nexus of many different actors in the public as well as in the private domain. With the so-called "surveillant assemblage" and its rhizomatic surveillance Haggerty and Ericson (2000) use an interesting model to highlight this nexus. In this regard, surveillance spreads over various branches whereas there are two essential characteristics of the surveillant assemblage: "its phenomenal growth through expanding uses, and its leveling effect on hierarchies" (Haggerty/Ericson 2000: 614). Hence, in many cases, surveillance does not merely include a single, operating entity (e.g., the state) but a variety of different, partially interrelated actors. Surveillance practices can thus result from a functional conglomerate of multiple actors. The surveillant assemblage "operates by abstracting human bodies from their territorial settings and separating them into a series of discrete flows. These flows are then reassembled into distinct 'data doubles' which can be scrutinized and targeted for intervention. In the process, we are witnessing a rhizomatic leveling of the hierarchy of surveillance, such that groups which were previously exempt from routine surveillance are now increasingly being monitored" (ibid: 606). In the view of Haggerty and Ericson (ibid), the different actors "work" together as "a functional entity" meaning that the gathered information feeds into this system of surveillance. Cohen describes the surveillant assemblage as "a heterogeneous, loosely coupled set of institutions that seek to harness the raw power of information by fixing flows of information cognitively and spatially" (Cohen 2012: 9). Regardless of the multitude of constituting actors, the actions of the surveillant assemblage basically include control mechanisms, mainly driven by political and economic interests in the public as well as in the private sector. Or in other words: the surveillant assemblage is often accompanied by public-private-partnerships. Surveillance and monitoring practices play an important role for national security as well as for the digital economy. Put simply, governmental actors aim at improving (national) security and economic actors to serve their business models and secure their commercial interests. Indeed, these interests are often intertwined and therefore hardly distinguishable.

There is a prominent showcase of the surveillant assemblage with its complex interrelations between security authorities and private companies: namely the so far biggest case of electronic surveillance revealed in 2013 by whistle blower Edward Snowden. Since the Snowden revelations, there is hard evidence for global mass surveillance programs that exploit ICTs in many respects. These programs are conducted by the US intelligence service NSA, the British intelligence organization GCHQ (Government Communications Headquarters) as well as other security agencies worldwide being their strategic partners (Greenwald 2014). The NSA became the major synonym to surveillance. However, in fact, irrespective of its powerful role, the agency is by no means the only entity conducting surveillance. Besides the NSA, also intelligence agencies in Europe (such as the GCHQ) and in other regions are deeply involved in global surveillance activities. Furthermore, the variety of surveillance programs serves as a hub for other security agencies in the US as well as for countries with special spying agreements. This primarily includes the so-called

"*Five Eyes*" partners Australia, Canada, New Zealand, UK and the US. Besides this so called Tier A allies, there are also Tier B allies that involve focused cooperation with a variety of countries in Europe as well as in Asia.[69] The spying partnerships also include a number of international organizations and enterprises that are involved in the surveillance programs (Greenwald 2014). Among the multitude of surveillance programs is the PRISM[70] program which gathers data from the servers of the most prominent online services (such as Facebook, Google, Microsoft, Apple, Youtube, Skype etc.). The massive data collection enables intelligence agencies to spy in real-time on their targets. Put shortly, PRISM monitors all kinds of online communications including e-mail, voice and video chat, photos, videos, stored data, social networking and so on (ZDNet 2013; Greenwald 2014: 108ff.). In the same vein, the GCHQ operates programs striving for all-encompassing surveillance such as MTI -"mastering the internet" (MacAskill et al. 2013) and Tempora[71], which gathers raw data directly from fibre-optic cables. In general, the Snowden files drastically highlight how sophisticated and deeply privacy intrusive global mass surveillance has become. These surveillance practices demonstrate how ICTs are exploited to monitor personal communications of every-day-life. However, irrespective of its enormous impact on society, the Snowden case should not blur the view on the very mechanisms and drivers of contemporary surveillance which exist much longer than the revealed programs and function in other contexts as well.

5.1.2 Securitisation and economisation[72] of digital identification

Contemporary surveillance practices including the implementation and use of according technology is driven by several socio-political and economic developments. Security and economic interests are core determinants of surveillance as well as of identification practices. To some extent, both are interrelated. Security and economic objectives play a dominant role for the emergence of digital identification and IDM (as e.g., outlined in Section 4.2 and 4.4). There is a securitisation and economisation of digital identification observable as the processing of information (directly or indirectly) referring to individual identities serves a complex mixture of security and economic interests. Similar is given in surveillance contexts which benefit from identification processes. Hence, metaphorically speaking, securitisation and economisation of identity information meet in the shadows of surveillance.

As outlined in Section 3.3.1, a wider paradigm shift in security policy and practices took place since the 1990s, which is carried forward by securitisation. Securitisation describes the phenomenon of an expanding security discourse spanning across multiple domains as security is framed as a perpetual process which then justifies extensive security

[69] (e.g., Austria, Belgium, Croatia, Czech Republic, Germany, Italy, Denmark, Greece, Hungary, Iceland, Norway, Spain, Sweden, Switzerland, Japan, Israel, Saudi Arabia, South Korea, Turkey)
[70] Trivia: a prism is an optical item to decompound a ray of light into its constituent spectral colors. https://en.wikipedia.org/wiki/Prism
[71] For instance, the GCHQ ran the Tempora project to gather raw data directly from fibre-optic cables see e.g.: GCHQ taps fibre-optic cables for secret access to world's communications", *The Guardian*. 21 June 2013, http://www.guardian.co.uk/uk/2013/jun/21/gchq-cables-secret-world-communications-nsa
[72] The terms securitisation and economisation are here used to point out the predominant role of security and economic rationales.

and surveillance activities. At the political level, this paradigm shift is accompanied by intensified use of surveillance technology to cope with novel security threats and challenges. In the aftermath of the 9/11 terrorist attacks it came to a significant extension of surveillance on a global scale. However, similar developments with a stronger focus on holistic security approaches including pre-emptive and preventive surveillance activities already emerged beforehand, as many scholars observed (e.g. Lyon 2003; Ball/Webster 2003; Bigo 2008; Haggerty/Samatas 2010; Bennett/Haggerty 2011). The arbitrary self-dynamic of securitisation, i.e., the continuous pursuit of security, reinforces surveillance. Consequently, surveillance practices increasingly strive for the preventive detection of risks before they become factual threats. The use of surveillance technology with its capability to easily gather large amounts of information fits perfectly into this framing and is thus presented as the preferable "weapon of choice" to tackle all kinds of security threats (Strauß 2017a). Furthermore, there are overlaps between economic drivers, the industries implementing security technology and security policy asking for this technology. In this regard, the logic of securitization with its perpetual striving for security widely corresponds to the quest for economic growth. Thus, securitization and economization are intertwined. Several scholars (e.g., Haggerty/Ericson 2000; Ball/Webster 2003; Lyon 2009; Bennett/Haggerty 2011; Ball et al. 2012) pointed out the global growth in security and surveillance modalities and the strong influence of economic mechanisms to stimulate new markets. The framing of security as a holistic concept and "moving target" stimulates demand for security and surveillance technology which benefits the according markets and vendors. The OECD uses the term security economy (OECD 2004) for the intertwining between the economic and the security sector. Examples for the influential security economy can be found in the enormous efforts made for security and surveillance at large-scale events such as the Olympic Games or World Cups. As Bennett and Haggerty (2011) highlighted, such mega-events obviously provide a good occasion for vendors to demonstrate their latest technology and experiment its usage in monitoring people and places which then may be adopted by security authorities. Hence, the economic rationale behind the implementation of security technology is often to stimulate new security markets and innovation.

A further issue is the widespread belief in technology as the best means to foster security in the political discourse (cf. Guild et al. 2008). Hence, technological push reinforces securitization and vice versa. The nature of ICTs and the broad scope of technologies seem to seduce policy makers to believe that complex security measures could be simply improved or even automated by technology. The basic rationale here often is an economic one in the sense of lowering costs, improving efficiency and effectiveness of security measures (Wright/Kreissl 2015). The employment of security technologies often happens without evaluating its effectiveness as well as its risks for privacy, other human rights and liberty which could raise more insecurity (Guild et al. 2008). Consequently, there are certain tensions between surveillance practices and human rights (ibid; De Hert 2012; Ball et al. 2012; Wright/Kreissl 2015). An example for these

controversies of securitization related to ICTs is the European data retention directive[73] that obliged EU member states to pre-store communication data of all citizens for at least six months. On the one hand, law enforcement repeatedly proclaimed necessity of this technological measure to combat crime and terrorism; on the other, the high and growing number of critics from experts and civil society alike argued that this practice of mass surveillance violates human rights. Among the risks is the inversion of the presumption of innocence as every citizen becomes a potential surveillance target without concrete suspicion (cf. FRA 2016). In 2014, the EU Court of Justice declared the directive as illegal for its violation of fundamental human rights and particularly the right to privacy (CJEU 2014). Irrespective of this landmark verdict, some countries (e.g., Germany) did not fully abandon data retention but made some legal readjustments to continue this form of surveillance (FRA 2016). In 2016, though, the EU Court again declared that data retention is incompatible with fundamental rights (CJEU 2016). The data retention case is one of many other examples of increasing pre-emptive surveillance practices which naturally affects all citizens regardless of their factual relation to illegal behaviour. These practices are particularly critical as they deeply intrude into individuals' fundamental human rights and bear certain risks of discrimination (see Section 5.1.3). The high intrusive capacity of data retention results from its possibilities to create extensive identity profiles of all citizens based on their communications and online activities.

As outlined in the previous sections, privacy intrusive surveillance implies the processing of information that refers to an individuals' identity. Consequently, identity and identification play a certain linking function in the complex relationship between privacy, security and surveillance. The nexus between securitization, extended surveillance and identity is observable in many respects. Identity represents a referent object of securitization related to the collective identity of a community or a national state (cf. Buzan et al. 1998; Bigo 2000; CASE 2006; Ajana 2013). National identity schemes such as passports, for instance, are instruments of security governance, showing that a person has a particular nationality and thus is a citizen of the according country. This identity representation allows categorizing individuals in citizens and non-citizens which is a typical security practice of border control. Similar deployment of (digital) identification as security practice is observable in many other contexts as well, where identity information is gathered to e.g., control access to a border, building or service or categorise individuals based on their characteristics. This includes a broad range of applications and objectives: e.g., measures for national security such as border control with ID documents or plans towards "smart" automated border control systems ("entry-exit") with biometric scanning[74], data retention approaches to preventively collect individual communications data or flight passenger records, secret surveillance programs, as well as identification procedures and IDM systems to improve the security of e-commerce, e-government or other online services (as outlined in Section 4.2).

[73] Directive 2006/24/EC of the European Parliament and of the Council on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:105:0054:0063:EN:PDF

[74] Plans are to use four fingerprints as well as facial images, see, e.g.: EU Commission (2016): Stronger and Smarter Borders in the EU: Commission proposes to establish an Entry-Exit System. Press Release, April 6 2016, http://europa.eu/rapid/press-release_IP-16-1247_en.htm

The push of digital identification and identity management systems is related to the security economy and the overlap between political and economic domains (cf. Lyon 2009; Bennett/Lyon 2008) Lyon (2009) uses the term "card cartel" to highlight the overlaps between these sectors and the "oligopolization of the means of identification", i.e., several players pushing and implementing eID cards (Lyon 2009: 16). Electronic identity cards thus "have become the tool of choice for new forms of risk calculation" (Amoore 2008: 24) which facilitate a "mode of pre-emptive identification" (ibid). In many countries, the planning of governmental eIDMS was linked to debates about national security and cybersecurity, reinforcing counter-terrorism, fighting organized crime, identity fraud as well as illegal immigration (cf. Bennett/Lyon 2008; Lyon 2009; Whitley/Hosein 2010). For instance, the US in their role as a member of the International Civil Aviation Organization (ICAO), a sub-organization of the United Nations, induced the creation of a standard for a travel ID (passport) with RFID-chip and biometrical data (Stanton 2008). Some EU member states considered this standard while planning and implementing their national eID systems. For instance, in Germany, Portugal, Spain the national eID is equipped with biometrical data and can both be used as travel document as well as for online transactions (ibid; Bennett/Lyon 2008; Kubicek/Noack 2010b). Hence, national security considerations also had some impact on the creation of governmental eIDMS, though they mainly aim at stimulating e-government and the digital economy (as outlined in Section 4.2). The political strategies inducing the emergence of these systems include an often indistinct mix of economic and security objectives. On the one hand, digital identification is seen as key enabler for different kinds of online services in public and private sector. On the other hand, it is framed as tool to achieve national security objectives. Approaches in the field of identity management are also seen as means to improve the ability of governments and intelligence agencies to identify potential security threats such as fake identities in the realm of crime and terrorism (cf. Glässer/Vajihollahi 2010). From a wider view, the increase in (digital) identification corresponds with fears and uncertainties of society as regards global risks and security threats. The employment of digital identification is thus also linked to security issues as well as preventive and pre-emptive surveillance mechanisms.

Besides the complex interrelations between national security and economic rationales in the government sector, there are many other forms of surveillance related to the securitisation and economisation[75] of digital identity information. Basically, big data and boundless information flows are framed as an economic core value. Big data serves as a cluster term for a new technological potential to exploit digital information for a broad scope of applications. It boosts the digital economy in the spirit of a "digital gold rush" to gather and exploit maximum data for socio-political and economic purposes of various kinds including analysis of consumer preferences, predicting behaviour, to automated risk calculation and predictive policing (cf. Mayer-Schönberger/Cukier 2013; Strauß 2015a). A working group of the World Economic Forum promoted the cross-border data flows to foster the global trade and investment system in the digital economy (WEF 2015). Personal

[75] Economisation here means that identity information is primarily treated as an economic factor, including its processing for commercial purposes but also beyond that framed as a quantitative figure.

data in general represents a crucial asset for a growing number of digital markets of any kind (cf. Spiekermann et al. 2015; Christl/Spiekermann 2016; Acquisti et al. 2016). Consequently, identity information is exploited for an extensive range of economic purposes. In contrast to its originally decentralized nature with some possibilities of anonymous interactions, the Internet became a central hub for a variety of options to capture vast arrays of information about personal identities, their behaviour and actions, interests etc. (cf. Acquisti et al. 2016). "As a result, chronicles of peoples' actions, desires, interests, and mere intentions are collected by third parties, often without individuals' knowledge or explicit consent, with a scope, breadth, and detail that are arguably without precedent in human history." (ibid: 3). The Web 2.0, social media and the like are prominent showcases for extensive clusters of information that enable deep insights into their users' identities, which serve various commercial purposes (see Section 4.3). The very design of social media platforms entails permanent availability of personal information. This availability makes social media but also other ICTs a very attractive source for all kinds of economic purposes.

The gathering and trading of personal information is literally big business: according to figures of 2015, the company BlueKai (a sub firm of Oracle) has about 750 million profiles of internet users with approx. 30,000 attributes about these individuals (Spiekermann et al. 2015). This example is only one among many others. There are several enterprises specialized as data brokers which conduct large scale profiling as their business model. Enterprises like e.g., Acxiom, Datalogix, Experian, or LexisNexis offer services that focus on exploiting massive amounts information about individual consumers (Christl/Spiekermann 2016). The data comes from various sources whereas the Internet and ICTs are generally highly valuable. For instance, many online services have several third parties involved (e.g., online marketers and advertising networks like DoubleClick or ScorecardResearch, content delivery network providers such as Akamai etc.) which gather information, e.g., for targeted advertising to serve their customers. Specialized data brokers aggregate information from many sources to create in-depth profiles. Similar to social network sites like Facebook building social graphs (see Section 4.3.2), data brokers employ identity graphs (e.g., Oracle's ID graph[76]) to map, structure and aggregate information about individuals. Oracle, for instance, promotes its graph as tool that allows to "unify addressable identities across all devices, screens and channels" as well as to "create a comprehensive consumer profile including what people say, what they do and what they buy" (Oracle 2015). Moreover, it "connects offline and online ID spaces with the goal of maintaining an industry-leading, comprehensive, and accurate ID graph in which all links are 'probabilistically validated'" (Oracle 2016: 20).

Oracle is one among several other companies with access to large sources of personal information being a strategic partner of the NSA (Greenwald 2014). Features like the identity graph thus serve economic as well as security and surveillance purposes. Besides data brokers that mainly conduct consumer profiling, there are other companies like

[76] See, e.g., Data management platform ID graph, https://www.oracle.com/marketingcloud/products/data-management-platform/id-graph.html
Oracle (2015): Oracle Data Cloud: The new driving force of data-driven marketing, online broschure, http://www.oracle.com/us/products/applications/brochure-data-driven-marketing-odc-2894231.pdf

Palantir[77] technologies with a special focus on data analysis and profiling for security authorities. Among its clients are several US security authorities such as the NSA, CIA, FBI, Department of Homeland Security - DHS, Air Force etc. (Burns 2015). Palantir has thus close connections to the intelligence community, has a strategic partnership with the NSA and was involved in developing its surveillance software XKeyscore (Biddle 2017). XKeyscore is a sophisticated search tool allowing intelligence agencies to explore all kinds of information about particular persons (including social media activity or online communications) based on their internet usage (Greenwald 2014).

Hence, social media and ICT usage in general provide vast sources of all kinds of information about personal identities which feed into the surveillant assemblage. As shown, the Snowden case exemplifies this in many respects. As highlighted in some of the Snowden slides, online social networks are attractive to intelligence agencies because the provide "insights into the personal lives of targets" including "communications, day to day activities, contacts and social networks, photographs, videos, personnel information (e.g., addresses, phone, email addresses), location and travel information" (Greenwald 2014: 158). They represent "a very rich source of information on targets" such as "personal details, 'pattern of life', connections to associates, media" (ibid: 161). But social media is only one of many other sources to follow the ambitions of the NSA and its partners to "collect it all" (ibid: 90). As outlined, PRISM exploits information from very widespread online services. Other surveillance practices even include the capturing of raw data streams (so-called upstream collections, as in the Tempora or the Stormbrew project which includes tapping of network devices) (ibid). With tools like XKeyscore analysts can enter search queries to seek for information about their targets similar to web search (Greenwald 2013; Greenwald 2014). It enables to find "nearly everything a typical user does on the internet" (ibid: 153). To find a particular person, keywords (so called selectors) can be used that refer to individual identities (e.g., IP address, phone number, e-mail address, usernames etc.) which then allow learning more about e.g., personal communications of a target.

As shown, there are manifold options for a variety of actors of the surveillant assemblage to monitor and track the behaviour of billions of individuals by exploiting digital information about them. The increasing availability of digital identity information stimulates surveillance desires in many respects. Hence, our digital identities are subject of manifold forms of surveillance. These surveillance practices are particularly driven by a complex nexus of economic rationales and objectives in the realm of national security. The Snowden disclosures underline that big data generally has a "supportive relationship with surveillance" (Lyon 2014: 1). Platforms, services and technologies with inherent identification mechanisms of various kinds can be very supportive to the practices of the surveillant assemblage. Entailed is an increasingly difficult distinction between personal and non-personal information, which facilitates de-anonymization and re-identification techniques (Section 5.3 takes a closer look at these issues). Regardless of whether identity information is exploited for economic or security purposes, these practices are mostly

[77] Palantir is co-founded by Peter Thiel, who is one of the technological advisers of US president Donald Trump. According to media reports, the company is involved in plans of the US Trump-administration to deport millions of immigrants (Woodman 2017).

hidden to the individuals concerned. In this regard, there are certain similarities between forms of identification and panopticism as presented and discussed in the next Section.

5.1.3 Panopticism and information asymmetries

As shown, there is a strong interplay between surveillance and digital identification. This interplay also concerns the functioning of surveillance and its panoptic features. Besides the all-seeing character "big brother" from George Orwell's novel "1984", Jeremy Bentham's panopticon (dating back to the eighteenth century 1791) is the most prominent metaphor of surveillance. This prominence is mainly due to the work of Michel Foucault, particularly his book "discipline and punish" (1977) where he analyses the emergence of the prison and mechanisms of control. Foucault interpreted the panopticon as "a generalizable model of functioning; a way of defining power relations in terms of the every day life of men" (Foucault 1977:205).

Figure 14: Presidio Modelo, prison built in the design of the panopticon, located on the Island de la Juventud, Cuba. (Source: Wikipedia:user Friman 2005, licensed as CC-BY-SA 3.0).

The aim of the panopticon is to create a state of conscious and permanent visibility for the automatic functioning of power. At the same time, as the insides of the watchtower remain opaque, "this model of undetected surveillance keeps those watched subordinate by means of uncertainty" (Lyon 1994: 60). Panoptic power is automated and de-individualized as the principle of power is not in the hands of a single entity but rather results from organisational settings, i.e., a hierarchical order with a permanently visible watch tower as a centralized control unit, which entails a lack of private sphere for the observed. With its strict hegemonic structure, the panopticon divides community and instead constitutes a sorted collection of separated individuals which are subjected to disciplinary power. Its aim is to control and normalize social behaviour through permanent (at least perceived as such) surveillance. Panoptic features of surveillance thus undermine the boundary control

function of privacy. Individuals are significantly hampered in their efficacy to self-determine their interactions with others without being controlled.

With the emergence of ICTs and their quasi-ubiquitous features, surveillance scholars began to rethink the panoptic metaphor in this altered technological context and introduced neologisms such as "electronic panopticon" (Gordon 1987), "superpanopticon" (Poster 1990), "panoptic sort" (Gandy 1993) or similar. There are several studies about the manifold aspects of electronic surveillance, technologies (e.g., CCTV, wiretapping, internet and telecommunications, RFID, location tracking etc. just to name a few) creating "panoptic" situations in everyday life contexts, often entailing evident threats of surveillance such as social sorting, discrimination, exclusion etc. (cf. Lyon 1994; Haggerty/Ericson 2000; Marx 2002; Lyon 2006; Bennett/Haggerty 2011; Marx 2015). However, irrespective of their relevance in particular, there is a "general tendency in the literature to offer more and more examples of total or creeping surveillance, while providing little that is theoretically novel" (Haggerty/Ericson 2000: 607). Lyon (1994) argued that there is no common understanding about the extent to which electronic surveillance has panoptic features. "Different analysts focus on different aspects of panopticism" (Lyon 1994: 67), and thus, the relationship between electronic surveillance and panoptic power remains relatively diverse. In the same vein, Haggerty (2006: 26) notes that "the panopticon now stands for surveillance itself", which makes it difficult to understand "the complexity and totality of contemporary surveillance dynamics" (Haggerty 2006: 38). As a consequence, it often remains unclear, to what extent a technology actually enables or amplifies surveillance and which modalities are relevant in this regard. Some scholars argue that the panoptic metaphor is invalid because ICTs relativize panoptic power and enable everyone to conduct surveillance. Consequently, there are novel options for bottom-up surveillance, inter alia called "sousveillance" (cf. Dupont 2008) or "participatory surveillance" (Albrechtslund 2008). Dupont (2008: 265f.) even claims that there would be a "democratization of surveillance" as civil society (individuals as well as NGOs etc.) can e.g., use the Internet to monitor the activities of governments or large corporations. In his view, the Internet is rather an anti-panopticon due to its decentralized architecture and democratic capacity. Indeed, phenomena such as the Occupy movement, Wikileaks, or Anonymous (despite of the critical issues these phenomena entail) exemplify the potential of ICTs to empower civil society and enforce more transparency by scrutinizing political and economic activities. These and similar approaches have some impact on public opinion. However, at the same time, global mass surveillance steadily proceeds and power structures widely remain stable. The assumption that ICTs, online services etc. would be widely independent and decoupled from surveillance control attempts of public and private authorities (e.g., Dupont 2008) is a fallacy. The actions of the surveillant assemblage rather demonstrate the opposite. Global mass surveillance programs drastically highlight that there is no need to control an entire online service as long as digital information flows are gathered. This is among the essential lessons of the Snowden revelations. Hence, even though there is no doubt about the enormous democratic and innovative potential of ICTs, this does not imply that panoptic power ceases to exist. The framing of ICTs as sort of anti-panopticon is as misleading as

the view on ICTs as primary instruments of surveillance. Social media, for instance, highlight that ICTs bear potential to serve democracy and hegemony alike. An example is the Arab spring which used to be falsely presented as a result of social media empowerment. While social media served as a tool to convey pre-existing democratic movements, the technology as such did not induce a democratic shift. Moreover, also the regimes (e.g., in Syria, Egypt or Tunisia) used social media to monitor and control the activities of the counter-movements (Benkirane 2012; Skinner 2012; Dewey et al. 2012). This example highlights the simple fact that power structures are complex and not simply changeable by technological means only but require a deeper change of socio-technical practices. Thus, hegemonic power structures may be reproduced, reinforced as well as relativized by the use of technology but they do exist and function without technology. In this regard, Foucault's interpretation of the panopticon is still very useful to explore modalities of power without a need to focus on technology. There seems to be a certain fallacy in surveillance studies to reduce the panoptic metaphor to an issue of architecture and design of a technology in order to explain its panoptic potential. Such attempts seeking for architectural analogies between the panopticon and surveillance technology are rather doomed to fail, especially in case of ICTs. Of course, the Internet and other ICTs are mostly multimodal, widely decentralized networks while the panopticon is centralized by its very design. In this regard, the panoptic metaphor is of limited use to explain the modalities of contemporary surveillance. However, there seems to be a misunderstanding about the universal character and functioning of the panopticon. As e.g., Fuchs (2010) pointed out, Foucault's theory does not exclude decentralized forms of surveillance. In his book discipline and punish, Foucault (1977) analysed the prison system which is not and never was completely centralized. There is no permanent connection between different prisons, but a prison as such is still a panoptic unit that exercises panoptic power. The same is given for surveillance technology and ICTs which are neither fully centralized control units. The Internet, for instance, is decentralized with myriads of websites, services, platforms etc. But this decentralized architecture does not change the fact that some providers may centralize their services or that information is gathered and stored in a centralized form. Much more important than the architectural design of the panopticon is the way it enables and reinforces disciplinary power. In this regard, Foucault's interpretation provides a valuable analytical lens to reveal the general functioning of surveillance and of control mechanisms. While a broad range of surveillance practices exist, their basic functionality is widely similar: namely to create asymmetry of power in order to exercise control.

Information processing obviously plays a crucial role for panoptic power and the way it exercises control over individuals. As outlined in the previous sections, there is a close relationship between identification and surveillance. Thus, identification practices can have similar effects than panoptic power inherent to surveillance. Lyon (2001: 2) defined surveillance as "any collection and processing of personal data, whether identifiable or not, for the purposes of influencing or managing those whose data have been garnered". Although surveillance can include information that does not directly identify a person, some form of (perceived or factual) identification occurs as soon as a particular individual is exposed to privacy-intrusive surveillance. At some stage in the surveillance process,

information about the very individual being monitored is required to exercise power over her. This power can have different shades and can function when surveillance has the effective capability of identification or give the impression to have this capability. This is given for traditional as well as for electronic surveillance. This does not necessarily imply that the identified person is aware of being identified. Particular not when surveillance aims at secretly controlling or manipulating persons. At the same time, surveillance does not have to effectively identify the individual to exercise control. It can be sufficient to create the notion of identifiability among individuals. This can trigger the so-called "chilling effect" meaning that individuals tend to celf-censorship and avoid exercising their rights such as free speech when perceiving to be under surveillance (cf. Lyon 2003; Raab/Wright 2012). Traditional surveillance usually includes observation, which, if targeted at a particular person, also implies some (cognitive) processing of information about this person. For instance, the guard in the panoptic watchtower exercises control over an individual by an informational advantage: he has more information about the individual than vice versa. This gives him the power to factually decide to observe a particular individual as well as to let the individual believe to be observable which creates uncertainty for this individual. The same principle is given for modern surveillance. But in contrast to traditional surveillance, observation is only one of many other surveillance practices while the processing of information about individuals is involved in any case. Identification mechanisms thus can have similar effects than observation in the panopticon. The table below shows some similarities between panopticism and (digital) identification. It compares basic characteristics of the panopticon (as identified by Foucault 1977) on the left column with common characteristics of digital identification.

Table 1: Similarities between digital identification and panopticism

Panopticism	***(Digital) identification***
Political technology	Tool of political and economic governance
Centralized control unit (watchtower)	Centralized data processing and meta-profiling
Trap of permanent visibility	(Quasi-)obligation to provide identity information; identifiability
Automatisation and de-individualisation of power	Pre-defined identity criteria, automated categorization and risk calculation
Asymmetry of power as core principle	Opacity of information processing and imbalanced control over information

The panopticon represents a political technology as its functioning can be used in manifold ways to support and reinforce control. Identification in general is an organizational mechanism aiming at reducing uncertainties in different interactions between individual and institutional entities. As shown, digital identification is also used as a governance tool and its mechanisms serve both political as well as economic objectives to implement various forms of control. In this regard, it is an effecting tool (cf. Hood/Margetts 2007; Bennett/Lyon 2008). Centralization is a crucial aspect of the panopticon which consists of

a centralized control unit (i.e., the watch tower). Identity information is often stored and processed in centralized databases and registers. Moreover, (as shown in Section 4), ICTs in general and the extension of digital identification in specific reinforce meta-profiling and the aggregation of information from various sources. The panopticon features a twofold state of permanent visibility and transparency: given the constant presence of the watchtower and the invisibility of the guard, individuals have no to place to hide and thus perceive to be under permanent observation but can never be certain whether to be observed or not. Foucault (1977: 203) called this the "trap of visibility". Similar situations may occur when individuals are requested to reveal their identities or provide identity information without knowledge about the actual purpose and use the information they provide. Identification puts the spot light at the individual as it (directly or indirectly) aims at revealing her identity. At the same time, it is often highly opaque and features transparency in the sense of hiding information (see Section 3.3.2). Incremental increases in contexts where the provision of identity information is factually or quasi obliged to e.g., use a service, access a platform etc. Furthermore, as shown, identity information is exploited for various contexts often beyond the knowledge and control of the individual concerned. The information processing inherent to digital identification is often standardized and automated based on pre-defined criteria to, e.g., categorize users, customers etc. serving various purposes such as CRM, profit calculation, scoring or risk management (cf. Bennett/Lyon 2008; Lyon 2009). Hence to some extent, individuals become de-individualized by organizational and structural settings as in a panoptic. Against the background of big data and similar trends, individuals may increasingly become subjected to automated control and risk calculation based on their digital identity information.

Finally, the fundament of the panopticon lies in its asymmetry of power between the watchers and the watched. More precisely, this asymmetry is an informational one as the watchers have more information on the observed than vice versa. As the chilling effect demonstrates (Lyon 2003; Raab/Wright 2012), the sheer perception or fear among individuals of being under surveillance, i.e., observable or identifiable may have a controlling effect. This asymmetry of information is the basic motor of panoptic power. Identification can create similar effects. This is particularly the case, when an asymmetry of power over identity information occurs which implies some form of information asymmetry. This includes insufficient ISD as well, i.e., when an individual is identified or perceives to be identifiable but lacks in knowledge or control about her as well as its usage purposes. In this regard, identification and identifiability of individuals can be related to forms of panoptic power. Therefore, in my view, identifiability, i.e., the possibility to become identified can lead to a form of information asymmetry from an individuals' perspective. The increasing tendency to digital identification in various contexts may entail and reinforce its panoptic features in many respects.

5.1.3.1 Information asymmetries, agency problems and social control

As argued, the essence of panoptic power is the creation and maintenance of information asymmetry. Information asymmetries occur in various contexts which are not per se panoptic. In general, this imbalance is a classical problem likely to occur between

individual and institutional entities but also between interaction partners in economic markets. Information economics deals with information asymmetries in relation to macro- and micro-economic issues such as market development and contracting. Asymmetric information is given if entities with different amounts of knowledge about an issue interact with each other (cf. Akerlof 1970; Stiglitz 2002). A simple example is an economic transaction such as a consumer buying a good from a vendor. The consumer usually has less information about the product (e.g., origin, ingredients, production process, the supply chain etc.) than the vendor. Agency theory[78] differs between two roles: the principal and the agent. The agent usually has more information than the principal and thus more power to take advantage of the principal. This setting is also known as the principal-agent or agency problem (Shapiro 2005). The fact that the agent has more information does not necessarily imply that he exploits the principal though, there is a certain risk in this regard (moral hazard). To reduce this risk, transparency and accountability of the agent, his motives and actions are essential.

Similar problems mirror in ICT usage in general and more specifically in digital identification practices and the privacy discourse. Digital environments are often accompanied by various information asymmetries. For instance, between users and providers of commercial social media platforms, which also bear tensions between collective and individual agency (cf. Milan 2015). In contrast to optimistic views on social media (seen as means of prosumption, user empowerment etc.), several scholars (e.g., Comor 2010; Fuchs 2014; Milan 2015) argue that the commercialization of social media and ICTs affects alienation and increase disparities. Hence, the empowering potential of social media and other ICTs became quickly assimilated by existing power structures and market mechanisms. In general, the way technology is developed and applied often reproduces and reinforces given power and information asymmetries.

Institutional entities usually have an informational advantage while the individual concerned lacks in knowledge about whether and how her information is used. Entailed is a state of insecurity and uncertainty for the individual. Consequently, there is an imbalanced control over identity information between the individual and the various institutional entities gathering and processing her information. Digitally networked environments, technologies and applications complicate this problem. Today, identity information is used in various contexts by a conglomerate of numerous institutional actors. While there is a general growth in identification systems and institutional control over digital identities observable, its purposes and functions increasingly blur for individuals who have to reveal their identity often without knowing whether and how it is used. This form of information asymmetry thus entails a lack in informational self-determination and individual control over digital identities. The increase in centralized databases, national registers, online platforms etc. and identification practices enhances the informational power of those entities holding information about individuals. Particularly, when different databases and systems use the same identification method, as e.g., intended by some governmental IDM approaches (cf. De Andrade etal. 2013) as well as IDM concepts of social media platforms (see Section 4.2 and 4.3). With an increase in standardised IDM

[78] Also called principal-agency theory

implementations, there is thus a risk of an incremental quasi-centralisation of identification systems or development towards a "pervasive IDM layer" (Rundle et al. 2008: 19). Tendencies to centralise data storage or use a single identification system for multiple purposes facilitate cross-linking and aggregating identity information from different sources, and thus profiling. But also besides explicit IDM implementations, information can be aggregated from multiple contexts without the knowledge of the concerned individuals. The manifold forms to explicitly and implicitly identify individuals thus reinforce information asymmetries and aggravate effective privacy protection (Section 5.3 deals with this issue more in-depth).

Information asymmetries can entail multiple forms of social control up to discrimination and manipulation. Individuals mostly lack options to reduce these asymmetries. Regardless of their relevance for various governance modalities, identification mechanisms, IDM etc. bear certain risks of social sorting, i.e., a means of verifying and classifying individuals to determine special treatment, e.g., for purposes of administration, CRM or risk management (Lyon 2003; Bennett/Lyon 2008; Lyon 2009; Ball et al. 2012). This "special treatment" can involve several forms of discrimination such as social, racial, or ethnic exclusion and thus reinforce social disparities, mistrust and racism (Lyon 2003). False positives of surveillance activities are evident, where innocent people were secretly monitored and suspected to be involved in criminal activity. Various examples exist where people falsely became classified as suspicious and subject to surveillance without concrete suspicion, such as false positives on the "no fly list" of the United States (Schneier 2006b; Krieg 2015) or cases in the UK, where innocent people become suspects due to profiling activities deploying DNA databases (Travis 2009). Another example is the case of a public servant and pro-democracy activist in New Zealand being accused of planning a terrorist attack in 2012. Based on surveillance of his online activities (as monitored in the context of the PRISM program) the NSA treated him as suspicious person and passed the data to the New Zealand security agencies. As a consequence, he became a suspect at a top-secret surveillance list. The authorities revoked his passport and raided his home. According to reports in the media, all this happened without legal grounds (Gallagher/Hager 2016). These examples demonstrate threats inherent to mass surveillance. Hence, not without irony, extensive modes of security and surveillance can undermine and disable the essence of human security, i.e., the protection of the integrity of the individual and her rights from different kinds of threats (see section 3.3.1). Thus, to some extent the human security concept becomes inverted. Instead of being protected, individuals may be exposed to the risk of being classified as security threats and discriminated by security and surveillance practices. There is thus often a thin line between applying identification mechanisms for efficient and secure provision of services and overwhelming social control. The securitization and economization of identity information facilitate the maintenance and reinforcement of information asymmetries: security authorities and companies alike benefit from informational advantages. For instance, what businesses understand as customer relationship management (CRM) to personalize services and advertisements may include profiling and tracking of individual behaviour, activities etc. and lead to scoring as well as prize discrimination from the individual consumers' perspective. The measures and programs that security agencies conduct in

order to protect national security may involve surveillance, reduce privacy, informational self-determination and entail censorship. Practical examples of discrimination are profiling and scoring activities, i.e., the statistical classification and sorting of individuals based on information about them. For instance, in the banking and insurance sector, credit scoring is used to classify customers based on their financial situation and consumer patterns (Dixon/Gellman 2014). Personal information in social media is inter alia used for so-called "personality mining" to predict user behaviour and recommend products. Such an approach is e.g., proposed by (Buettner 2016) who (similar to other marketers) shares the belief that personality would be computable and thus also predictable. The approach uses a five factor model from behavioural psychology (openness to experience, conscientiousness, extraversion, agreeableness and neuroticism) to determine one's personality. The factors are explored by conducting a detailed statistical analysis of user information (e.g., profile information, time spent online, no. of logins, contacts, interests, posts, pictures etc.) that is inter alia provided by the social graph (see section 4.3). As a result the factors are represented by probabilities which then can be fed into applications to e.g., recommend different products or customize tariffs. For example, a British insurance company planned to calculate its tariffs based on Facebook posts of its customers: the idea is to analyse the personality profile of first-time drivers. Persons assumed by the algorithm to be conscientious and well-organised can expect a higher score than e.g., those assumed to be overconfident (Ruddick 2016). It is evident that social media is also classifies its users based on their profile information which in some cases even lead to ethnical and racial profiling such as Facebook's system provides a feature for its marketing customers "to exclude black, Hispanic, and other 'ethnic affinities' from seeing ads" (Angwin/Parris 2016). As already mentioned, security and intelligence agencies exploit ICTs and data of online communications for all kinds of surveillance activities. Social media is a particularly easy target and thus increasingly monitored by law enforcement and security agencies (cf. Greenwald 2014; Belbey 2016; Bromwich et al. 2016). Besides the various mass surveillance practices, users of privacy tools can become special targets of surveillance programs. An example is the previously mentioned NSA software tool "XKeyscore": the tool allows searching for particular persons based on pre-defined criteria. Security researchers and journalists who analysed the source code of the software revealed that it categorizes users of privacy-friendly software such as the Tor browser or other anonymization tools[79] (more details see Section 5.4) as "extremists" and thus as potential targets of surveillance (Rötzer 2014; Doctrow 2014). Another example for social sorting on a large scale can be found in the republic of China which is currently testing a (yet voluntary) social credit scoring system which is planned to be mandatory for the whole population by 2020. The system aims at collecting as much information about citizens' interests, actions, behaviour etc. as possible from all kinds of electronic sources. The system is inter alia backed by the national ID card system as well as several e-commerce and credit companies like Alibaba and Tencent who also run all Chinese social networks and thus have extensive personal data collections. The concept is partially similar

[79] Privacy tools like Tor are inter alia used by journalists, lawyers, human right activists and other sensitive professions to protect their communications from surveillance.

to credit scoring systems which are used to verify one's credit rating, but much more intrusive. Based on information it has about the population, the system creates a score for every citizen (ranging between 350 and 950) which alters with good and bad behaviour. The system determines various forms of bad behaviour such as running a red light[80], posting a political comment online without prior permission, buying a video game etc. Even bad behaviour of one's relationships may lead to a reduction of one's score. People with higher scores gain benefits (such as lower costs for car rental at a score of 650 or a travel permission to Singapore when reaching the score 700) people with lower values receive restrictions have less chances to get jobs (Storm 2015; Denyer 2016). This scoring system can be seen as digital panopticon par excellence. Besides its deep intrusiveness, this example can be seen as part of a general trend to quantify personal information and use it for different scoring systems in line with the global big data paradigm. Ideas to apply scoring not just for financial issues but also for law enforcement and crime prevention can e.g., be found in Germany where the Minister of the Interior made similar proposals (Krempl 2016). Related concepts already exist in the field of predictive policing which aims at identifying "likely targets for police intervention and prevent crime or solve past crimes by making statistical predictions" (Perry et al. 2013). Predictive policing systems such as IBM's "blue C.R.U.S.H." (Criminal Reduction Using Statistical History) or the software "TrapWire" are already in use e.g., in the US and the UK. Threat scenarios in the sense of the movie "Minority Report"[81] may be overestimated. However, the use of automated predictive analytics entails several risks and complicates to find the right balance between appropriate computation and excessive social control (Strauß 2015a).

5.1.3.2 Algorithmic authority – toward semi-automated identity scoring?

As outlined, there are many cases where personal information is used for purposes that are unknown and uncontrollable for the individuals concerned. Such approaches raise serious ethical concerns and can include various forms of discrimination, stereotyping, manipulation and abuse of power. The activities of profiling, scoring, monitoring etc. are widely automated by data mining and pattern recognition algorithms that analyse vast arrays of digital information. In this regard, "code is law" (Lessig 2006), as software and algorithms increasingly affect and determine society and thus entail a certain regulative capacity. This regulative capacity is often vague, hidden and opaque and this form of regulation is far from being open to public scrutiny and oversight. Shirky (2009) used the term "algorithmic authority" to describe this phenomenon. Today, this algorithmic authority occurs in various forms already. Considering a further increase in this regard, a digital identity may become increasingly co-referential with its (algorithmic) environment (as e.g. observable in social media). Hence, in some contexts it may be seen as a sort of self-referential machine, because it is permanently determined by the dynamics of its constituting digital information flows. For instance, algorithms of online platforms process user information to calculate what users may be interested in and then present customised information to them. This is a form of strict-determinism as information about (assumed) interests of the past is used to determine the amount of information presented to a

[80] This costs, for instance, 50 points.
[81] In the plot, persons become arrested based on statistical predictions that they will commit a crime in the future.

particular user. To some extent, this can limit individuals' view on reality at least within the scope of the information system. Practical examples can be found in many digital systems of every-day-use (e.g., online consumer or social media platforms, internet search engines etc.); where personal information is used to customize and filter the amount of information an individual receives from the system. This is, e.g., useful for marketing purposes where the detection of specific patterns in consumer habits may smoothly include soft manipulation to create demands such as in Amazon's recommendation algorithm (Van Dijk 2014). In the view of platform providers, these practices are reasonable in order to improve service quality and user experience. But this can lead to a "filter bubble" (Pariser 2011) or "echo chamber" (cf. Hosanagar 2016). Individuals then repeatedly receive information about their assumed interests while other information is automatically filtered out. This can complicate to receive other information. Moreover, free and self-determined content production in social media platforms can be very restrictive. Content which platform providers such as Facebook perceive as inappropriate may be censored or erased (Heins 2014; York 2016). Consequently, users mostly receive information based on their assumed interests as well as information aiming at creating particular demands. Similar mechanisms may also serve purposes of voter manipulation: according to media reports, big data analysis conducted by the company Cambridge Analytics was used to predict and influence the behaviour of voters in the US presidential election in 2016 in support of president-elect Donald Trump (Confessore/Hakim 2017; Beuth 2017).

The increasing trend to employ automated profiling, scoring, risk calculation etc. includes approaches to quantify information about individuals which feeds into various forms of social control. Entailed to these developments are manifold societal risks and above all, there is a certain risk that individual identities are reduced to the sum of their (digitally represented) informational parts and thus to quantitative, computable factors. If a person's identity is merely recognized based on her computable digital representation, the non-computable parts of identity (ipse) may be statistically normalized. This partially refers to the de-individualization mechanism of panoptic power as the uniqueness of an identity may diminish with its expanding digital representation. The mentioned plans of the citizen score system carries this to its extremes as an individuals' score may determine significant parts or socio-technical contexts of her life. But also less drastic examples bear many societal risks. While automated information processing aims at fostering efficiency and effectiveness, automated decisions affecting individuals are highly critical from a privacy perspective, as they facilitate social sorting, stereotyping and manipulation. The outlined examples highlight these problems. Additional problems occur as these automated socio-technical practices of big data may entail a drastic growth in complexity induced by massive data collections which can amplify the probability of errors, false positives and spurious correlations (cf. Strauß 2015a; Calude/Longo 2016). For instance, Hamacher and Katzenbeisser (2011) demonstrated that the preventive gathering of data to, e.g., predict criminal activities does not lead to better results as expected from big data and predictive analytics. The increase in complexity of information processing and the costs to reveal and correct errors may even entail opposing effects. These costs for detecting and correcting failure do not merely involve economic but above all many social ones (Wright/Kreissl 2015; Strauß 2017b). But also in economic terms the financial burden of surveillance is

enormous: for instance, the economic impact of NSA surveillance is expected to exceed 35 billion USD, as estimated by the US Information Technology and Innovation Foundation (Castro/McQuinn 2015).

As identity is not static but dynamic, approaches to gather identity information naturally entail some uncertainty as regards the validity of this information. This is particularly given for dynamic parts of identity information (i.e., ipse) which are naturally hard to quantify. Besides other things, reductionist attempts aiming at quantifying ipse-information bear risks of misleading interpretations, which may even harm the individual represented by this information. For instance, incorrect behavioural patterns or false correlations may lead to discrimination or false accusations of innocent individuals (as shown in the previous sections). Furthermore, the dynamic character of identity in general is vital for individual experiences, personal and societal development alike. Any limitation or restriction thus entails manifold social and ethical problems for personal and societal well-being. Attempts to increase data quality contribute little to reduce these problems when identification is employed for unethical purposes.

As shown, there are several overlaps and similarities between panopticism and forms of digital identification. While there is neither a universal identification system nor is identity information per se exploited for surveillance and panopticism, there are many examples that highlight how identification serves panoptic forms of power. Lyon (1994: 67) argued that with the panoptic metaphor applied to ICTs "(…) we see here nothing less than the near-perfection of the principle of discipline by invisible inspection via information-gathering." The myriad of global mass surveillance activities seem to underline Lyon's view. Lyon (1994) also discussed the question whether the panoptic can be generalized over different social spheres, which cannot be easily answered as society as such is too complex. I argue that the employment of identification mechanisms can contribute in various ways to the emergence of panoptic power in different social spheres. ICTs foster the inherent function of identification to connect and link different entities and systems. From this view, digital identification may be seen as a conductor of panoptic power. Identification is here not meant in a strict sense of uniquely identifying a person, but rather of processing information about an individual which can be used to reveal her identity. In the original concept of the panopticon, individuals are literally observed as they cannot move their bodies without being watched. This is not the case with modern surveillance where physical presence is not a necessary condition anymore. But the core feature, i.e., the construction of information asymmetry is widely similar. Instead of being under potentially permanent observation, modern surveillance practices exploit the various identity traces individuals create when using technologies, applications etc. These identity traces are widely available decoupled from space and time, though often invisible and uncontrollable for the individuals themselves. Considering the growth in big data and predictive analytics, the individual not just becomes increasingly controllable in the presence, but to some extent even in the future as her behaviour can be predicted and thus manipulated based on semi-automated systems.

In total, there are many privacy risks resulting from information asymmetries entailed to means of identification in surveillance contexts. However, the problem does not result from identification per se, which is a crucial governance instrument is serving many

important socio-political and economic purposes as well. But the contexts and purposes it is applied to are often opaque, lack in transparency and accountability. Consequently, individuals have limited options to control their information. Hence, privacy not merely suffers from extensive information gathering and surveillance but mainly also from the entailed lack of transparency and accountability of information processing. This also mirrors in the perceptions of European citizens, as will be shown in the following Sections.

5.2 Citizens' perceptions on privacy, security and surveillance[82]

Although security and surveillance measures directly affect individuals, little is known about their perceptions and opinions on contemporary surveillance. The EU-funded SurPRISE project[83] contributed to narrow this gap and explored the perceptions of European citizens on the interplay between surveillance, privacy and security with a focus on surveillance technology. A core part of the empirical work included a large-scale participatory approach with citizen summits held (in 2014) in nine European countries with about 200 participants each.[84] In total, the summits had 1780 persons participating with N=1772 valid responses. The participants were recruited differently across the involved countries. In Austria, Italy, Hungary, Spain and the UK, external contractors conducted the recruitment. In Denmark, Germany, Norway, and Switzerland, a mix of channels (such as postal invitations, announcements in online and print media) was applied to get people involved. In each case, the recruitment was based on particular criteria (age, gender, educational level, occupation and geographical area) to avoid bias and achieve a heterogeneous panel structure. Persons with particular expertise in the project issues privacy, security, surveillance or similar were excluded in order to avoid expert-driven opinions. A relatively even panel distribution could be achieved based on these criteria. As regards gender, there was a slightly higher share of male (52%) than of female (48%) participants. The age distribution ranged from 18 to 70+ with a slight majority (44%) of people belonging to the categories (ranging from 40 to 59). The participation processes followed the same basic design with a combination of quantitative and qualitative methodology: a predefined interactive survey was the basic instrument to gather quantitative data, and in addition, three thematically structured group discussions were held during each summit to gain more insights into the rationales behind the perceptions of the participants. To explore eventual differences in the perceptions as regards technology, a particular focus was set on surveillance-oriented security technologies (SOSTs). This term was introduced in the project to describe security technologies with inherent surveillance capabilities. The participants were confronted with the three different SOSTs smart CCTV, deep packet inspection – DPI (a specific form of internet surveillance) and smartphone location tracking – SLT. These technologies served as examples representing different

[82] Parts of the results presented in this Section refer to (Strauß 2015b) and (Strauß 2017a).
[83] The project received funding from the EU's Seventh Framework Programme for research, technological development and demonstration under grant agreement No.: 285492.
[84] In Austria, Denmark, Germany, Hungary, Italy, Norway, Spain, Switzerland, and United Kingdom. Further information about these national participation processes is available at http://surprise-project.eu/events/citizen-summits/

issues such as visual privacy, privacy of information and communication as well as locational privacy.[85] In following, some of the main quantitative results of the citizen summits are presented and discussed.

5.2.1 Major attitudes and concerns about surveillance technologies

Figure 15 below illustrates various attitudes and concerns of participants concerning privacy and security issues of SOSTs. The results show that for the majority of 64%, there is some necessity of using surveillance technology in order to improve security. Only a minority of the respondents (26%) thinks that SOST usage rather has the function to demonstrate action against crime. However, the general position that using surveillance technology can contribute to security does not imply a general acceptance of privacy intrusions and surveillance. The highly-expressed fears and concerns indicate that the problem of information asymmetries is also perceived among the survey participants. 50% of the respondents do not share the opinion that worries about surveillance technology are unfounded if one has done nothing wrong (this refers to the classical "nothing to hide" argument, which is discussed in the next paragraph). For 70%, it is likely that SOSTs contribute to an abuse of power. The concerns about the misuse of personal information point to perceived uncertainty and insecurity as regards information processing: 70% are concerned that too much information is gathered about them; 63% worry about inaccurate and thus misleading information held about them; about 80% are concerned about their personal information being used against them and 91% expressed concerns about their information being shared without their permission[86]. Thus there are several fears of privacy violations among the respondents due to the misuse of personal information as processed by surveillance-oriented security technology. In total, the results indicate that the majority perceives the intrusiveness of security and surveillance measures as threat to privacy. As a consequence this can also have negative impact on the perceived effectiveness of a security measure, as visible in the further results.

[85] For further details about the methodology and synthesis of the summits see Strauß (2015b). Further information about all national participation processes, individual country reports, research data, information material and impressions of the summits is available at http://surprise-project.eu/dissemination/research-results/

[86] Fear of unauthorized information usage also refers to the problematic aspects of informed consent as one cornerstone of privacy. While consent is essential for legal data processing, it is often difficult to effectively prohibit or permit the use of information due to a lack of alternatives from the individuals' point of view. See also Section 5.4.3.

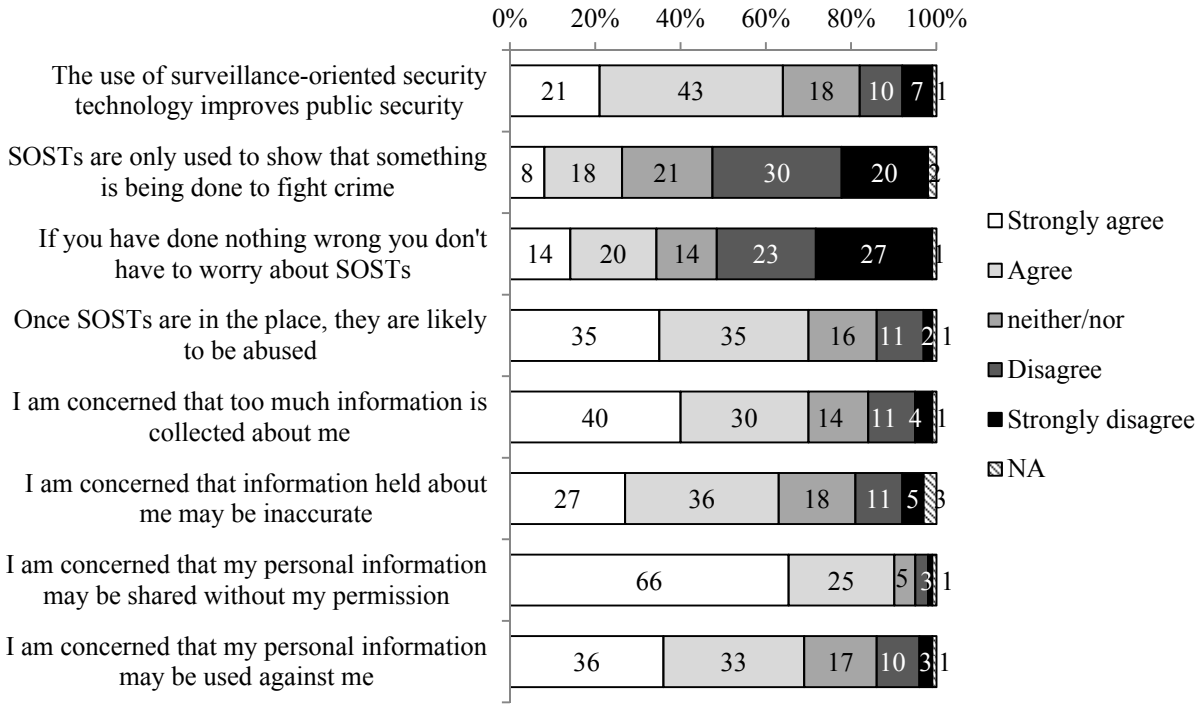


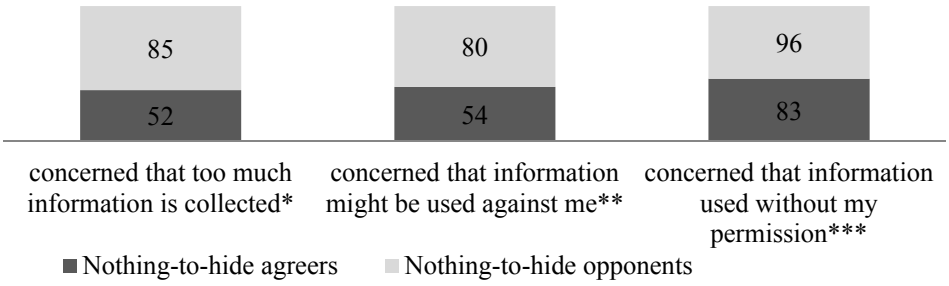
Figure 15: Major attitudes and concerns about SOST usage (N=1772)

5.2.2 "Nothing to hide" unscrambled

As outlined in the previous sections, privacy is often falsely framed as a form of secrecy which reduces it to a means of hiding things respectively information. This framing supports the misleading view that privacy would be in opposition to security and transparency (as discussed in Section 3.3). A very common argument in this regard is the statement "those who have nothing to hide have nothing to fear" (cf. Solove 2004; Schneier 2006a; Solove 2011). This statement is frequently used in the security discourse to justify surveillance measures and privacy intrusions. The rhetorical trick of this argument is to present privacy as equivalent to secrecy which conveys the impression of something shady and suspicious. As a consequence, the argument implies that one should not worry about surveillance and privacy infringement as long as one behaves correctly and does not has the desire to keep information secret. In this regard the argument corresponds with disciplinary power in a Foucauldian sense. It restricts options to express concerns against surveillance (such as limitations to privacy, freedom of thought and expression etc.) as they are automatically framed as something suspicious as those who worry may have done something wrong then. Consequently, the nothing to hide argument also implies that those who have concerns have something to fear then.

In order to learn what citizens think about this line of argumentation a similar statement (see statement two in Figure 15) was asked: "If you have done nothing wrong, you don't have to worry about SOSTs". The results show that 50% reject this statement while only 34% share this opinion. As shown previously, the majority expressed high concerns about malpractice of personal information, i.e., that too much information is

collected, information is used against them and information is shared without the permission of the concerned individuals[87]. These results were cross-linked with those about the perceptions on the "nothing to hide" (NTH) statement (see Figure 16 and Table 2 below):



* N=549 supporters, N=804 opponents. ** N=541 supporters, N=807 opponents.
*** N=549 supporters, N=803 opponents.

Figure 16: Concerns about information misuse among NTH agreers and opponents (percentages).

Table 2: Concerns of nothing-to-hide supporters and opponents (percentages)

		Concerned	Not concerned	Neither nor	NA
Too much information is collected	Supporters	52	29	18.5	0.5
	Opponents	85	6	8.5	0.5
Information might be used against me	Supporters	54	23	23	0
	Opponents	80	7	12	1
Information is shared without my permission	Supporters	83	9	8	0
	Opponents	96	1	2	1

It is expectable that most NTH opponents (85%) have concerns about the collection of their personal information. But it is rather surprising that the majority of NTH supporters (52%) share the same concerns about too much information being collected about them. Moreover, the same contradiction is observable as regards the other concerns: 54% of the NTH supporters expressed concerns about their information being used against them and 83% are concerned that their information is shared without their permission. Thus, even those people who perceive to have nothing to hide and thus claim not to worry about SOSTs in fact worry about extensive information collection and abuse. These results further confirm that the "nothing to hide" argument is, similar to the assumed privacy-security trade-off (Section 3.3.1), misleading. As shown, the perceptions of citizens on privacy and security are more differentiated than suggested by a narrow framing of privacy which neglects its public value. These results rather indicate that citizens do not accept the

[87] This broad disagreement to the nothing to hide argument also mirrors in the results of the different countries with the strongest opposition in Germany. Exceptions are given in Hungary and UK, where the respondents tended to agree with the statement. However, also in these two countries the participants expressed concerns about extensive information collection and fears of misuse.

gathering of their information for surveillance purposes without plausible reasons. Hence keeping information private seems to be an important issue among respondents, especially *because* of concerns about unjustified and uncontrolled information gathering for surveillance. Consequently, ISD and self-controlled handling of their information is an important issue as well as to what extent and for what purposes this information is being processed (Strauß 2017a).

5.2.3 Perceived intrusiveness and effectiveness of surveillance technology

The degree of intrusiveness, i.e., the extent to which surveillance technology intrudes into an individuals' privacy can be expected to have an impact on how the individual perceives the use of the technology. Different technologies usually have different intrusive qualities and surveillance modalities, which may also affect their perceived effectiveness concerning security. To explore people's perceptions on these different intrusive qualities as well as on the perceived effectiveness, the participants were asked to assess statements for each of the three SOSTs (as shown in Table 3). Each technology represents a different mode of intrusion and thus affects privacy differently: smart CCTV mainly involves visual surveillance, smart phone location tracking (SLT) monitors one's movement and thus intrudes into locational privacy, and deep packet inspection (DPI) represents surveillance of internet activity and thus includes intrusions into the privacy of information and communications. Hence, these technologies process different types of personal information which effect the perceptions on privacy intrusions. The results indicate that the more intrusive a technology is perceived the more concerns it raises which can also influence the perceived effectiveness.

Table 3: Major views on intrusiveness and effectiveness of SOSTs (percentages, N=1772)

		Agree	*Neither/ nor*	*Disagree*	*NA*
… is an effective national security tool	sCCTV	64	18	17	1
	DPI	43	24	32	1
	SLT	55	25	20	0
The idea of … makes me feel uncomfortable	sCCTV	39	20	40	1
	DPI	66	16	17	1
	SLT	45	24	31	0
I feel more secure when … is in operation	sCCTV	43	25	32	2
	DPI	12	25	61	2
	SLT	27	29	43	1
… is forced upon me without my permission	sCCTV	60	16	23	1
	DPI	87	7	5	1
	SLT	68	14	16	2
… can reveal sensitive information about me	sCCTV	52	15	32	1
	DPI	80	10	10	1
	SLT	60	20	19	1
… can lead to misinterpretations of my behaviour	sCCTV	67	13	19	1
	DPI	77	12	10	1
	SLT	66	16	17	1
… can reveal strangers where I am or was	sCCTV	57	16	26	1
	DPI	73	14	12	1
	SLT	68	16	16	1
I worry about how the use of … could develop in the future	sCCTV	67	13	19	1
	DPI	84	9	7	1
	SLT	65	18	17	1

As shown, respondents were relatively ambivalent in their perception on security provisions by the technologies. While at least smart CCTV (64%) and SLT (55%) were assessed as effective security tools for a majority, more than two-thirds also said to feel uncomfortable by these technologies. In the case of DPI, the position was clearly opposed as only 43% perceived as effective tool while 66% feel uncomfortable by the use of technology. None of the SOSTs was perceived as raising security when being in operation. DPI as being the most intrusive technology was also assessed by the participants as such. Furthermore, it is also received the lowest values as regards its perceived effectiveness to improve security. Only for 12% this technology contributes to a feeling of a security gain while the clear majority of 61% perceives the opposite. A similar pattern though less distinct is observable for the other SOSTs as well. Security gains are perceived by 43% in case of smart CCTV and in case of SLT by 27%. This reluctant view on security gains of the SOSTs indicates that privacy intrusions to serve security purposes are not (as often assumed) simply accepted but can even lead to an increase in subjective insecurity. This insecurity mirrors in the high concerns: for each SOST, the vast majority perceives that the technology is forced upon them (sCCTV – 60%, SLT – 68%, DPI – 87%). The same is given for concerns about the technologies revealing sensitive information (sCCTV – 52%, SLT – 60%, DPI – 80%), or information about one's locations (sCCTV – 57%, SLT –

68%, DPI –73%), as well concerns about misinterpretations of behaviour ((sCCTV – 67%, SLT – 66%, DPI – 77%). Considering the different values for each SOST, it is conspicuous that DPI raises the highest concerns and lowest acceptance mostly followed by SLT and then sCCTV. This indicates an interrelation between intrusiveness and effectiveness: the more intrusive the technology is perceived the less effective it is perceived. While all three SOSTs are assessed as privacy critical, smart CCTV raised slightly lower concerns in most cases compared to the other two technologies. One explanation is that CCTV is well-known, it is perceivable what the technology does (i.e., a camera that gathers images) and it is limited to particular areas in the public while the others are more abstract, hidden and intrusive as they can gain much deeper insights into one's private sphere. At the same time, the technologies behind SLT and DPI are more present in every-day-life as most people permanently carry smartphones and use the Internet every day. The related surveillance practices can gain very deep insights into one's identity by e.g., gathering one's movements, social relationships, communications, interests etc. These results thus indicate that the more details about individual identities are gathered from their very private sphere such as permanent monitoring of individual behaviour, communications and interests, the higher are the individual concerns about the collection and misuse of information. Also here, information asymmetries play a role. The information asymmetry entailed to a deeply intrusive technology (such as DPI) with multiple surveillance options may be perceived as higher compared to technologies with a more clear focus on gathering particular information (such as CCTV primarily gathers images). In total, the clearly expressed concerns represent a strong statement against privacy intrusive security and surveillance practices. Moreover, the vast majority is worried about how the technologies could develop in the future which indicates a certain fear about a further expansion of surveillance, function or mission creep (i.e., the uncontrolled, extensive surveillance).

These results allow drawing the conclusion that the intrusive quality of a technology, i.e., the way it interferes with privacy also have an influence on its perceived acceptability. To some extent privacy intrusion seems to be tolerable for the respondents if it is necessary for a security gain. This is highlighted in Figure 17. However, improving security is a necessary but not a sufficient condition for the acceptance of privacy-intrusive practices.

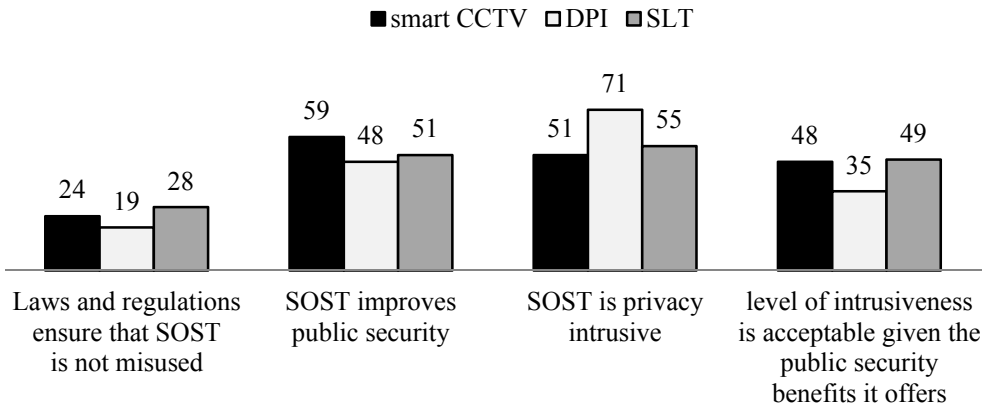


Figure 17: Intrusiveness and acceptability (percentages) (N=1772)

As shown, it makes a difference what types of information a technology gathers as well as how deeply it can intrude into an individuals' life. This is not just a matter of the technology and its design but also of the socio-technical practices in which it is applied to and the usage contexts. The crux is that the usage of surveillance technology is mostly opaque and obscure, which was intensively debated in the group discussions at the citizen summits. Discussants often mentioned the importance of legal control and oversight to ensure that surveillance is appropriate, lawful and not abused. However, there is only little trust that laws and regulation are sufficient to protect from malpractice and misuse of the technologies as Figure 17 shows: in each case, less than 30% think that this is the case while for the vast majority protection is insufficient. Hence, these results point out that transparency, accountability and effective oversight are among the essential issues for privacy protection.

5.2.4 Trust in security authorities

In accordance with the low amounts of trust in regulation, the respondents also expressed high uncertainty as regards trust in security authorities to use the surveillance technologies with respect for human rights. As Figure 18 illustrates, security authorities are tendentiously perceived as trustworthy though with some differences as regards the SOSTs (36% smart CCTV, 36% DPI, 46% SLT). However, only a minority of respondents expressed no doubts about the authorities abusing their power while there is a clear tendency towards the opposite: 46% in case of smart CCTV, 34% for SLT and 52% regarding DPI. Hence, the respondents have a certain fear that security authorities take advantage of surveillance technology and abuse their power.

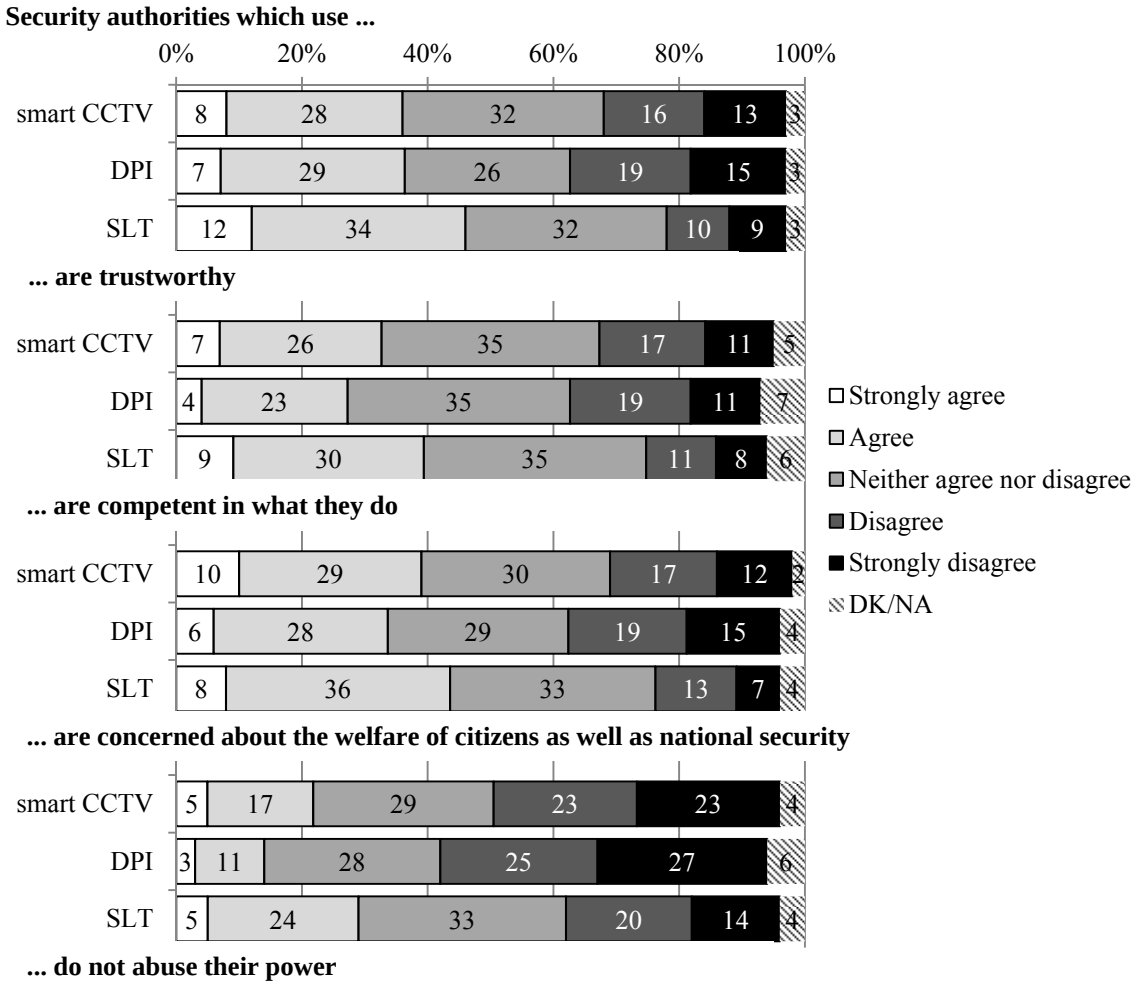


Figure 18: Trustworthiness of security authorities (N=1772)

Furthermore, these results have the highest values in the category "neither/nor" in the whole survey which mostly exceed 30%. This indicates a high amount of uncertainty and insecurity among the respondents as regards trusting the authorities. As the previous results reveal, for most of the participants, the use of SOSTs represents a very intrusive measure which raises enormous concerns about the misuse of personal information by extensive surveillance activities. Although security measures as such are not rejected but to some extent perceived as being relevant and useful, the clearly concerns dominate. Particular fears are observable about function creep and the authorities abusing their power by employing the technologies. These fears result from privacy intrusive technology usage, extensive surveillance as well as a lack of accountability and oversight of security authorities which altogether hamper trust. What complicates the building of trust is that security and surveillance practices mostly imply certain mistrust in the observed persons. This is particularly problematic in case of untargeted measures such as mass surveillance where everyone may represent a potential suspect. Consequently, such measures rather raise insecurity and uncertainty. Not least because mistrust can reinforce itself: security and surveillance practices that mistrust citizens likely lead to mistrust of these citizens in the authorities conducting these practices.

In total, these results significantly relativize the assumption that the majority of persons lacks awareness about privacy and has little concerns about security and surveillance measures.[88] The perceptions of the citizens also refute traditional conceptualizations of the privacy-security interplay in terms of a trade-off (as discussed in Section 3.3.1). Several other studies dealing with related issues came to similar results[89] (see e.g., Friedewald et al. 2017). The results indicate that security measures which neglect privacy protection do not contribute to raise security perceptions but can even entail opposite effects. This is particularly the case when surveillance is based on extensive collections of personal information and operates beyond the scrutiny of the public. Privacy protection suffers from privacy-intrusive security and surveillance measures which directly and indirectly affect individual identities as their private spheres are intruded in manifold ways. Critical issues concern insufficient control for individuals over their information (ISD) as well as lacking transparency and control of surveillance practices and their operating entities. These issues are of increasing public concern, as other recent studies about online information processing confirm. For instance, the special Eurobarometer 431 on data protection of June 2015 (EB 2015) shows that citizens have major concerns regarding the control and protection of their personal data: 67% of the EU citizens perceive a lack of control over the information they provide online and for 37% there is no control at all over their information. The majority of respondents (57%) disagree to the statement that "providing personal information is not a big issue" and 53% feel uncomfortable about the use of information about their online activity by Internet companies. For nine out of ten Europeans it is important that their personal information is thoroughly protected irrespective of the country a public authority or private company processing the information is located. Only a minority of 24% trusts providers of social network sites to set appropriate privacy settings. Seven out of ten respondents say that personal information should not be collected and processed without their explicit permission. In general, there are widespread concerns among Europeans about their information being misused (ibid). The flash EB 443 on e-privacy of July 2016 provides similar findings (EB 2106): for over 70%, guaranteed confidentiality of their internet communications is very important. Nine in ten persons (90%) are in favour of having their communications (messages and calls) encrypted so that only the recipients can read them. 78% of the respondents find it very important that the use of their personal information on their devices (e.g., computers, smartphones and tablets) requires their permission. Near to two-third (64%) find it unacceptable that they only get unrestricted access to a website or service by having their online activities monitored. 89% of the respondents think that the default-settings of internet browsers should be sufficient to protect from tracking (ibid).

The presented results in total further indicate that information and power asymmetries are among the core problems of contemporary privacy protection. Individuals generally

[88] It has to be noted that the presented perceptions of the participants may differ from their views in "real world" settings. Nevertheless, the expressed attitudes are important to improve the understanding of the interrelations between privacy, security and surveillance.

[89] Such as of the EU projects PRISMS (privacy and Security mirrors - http://prismsproject.eu/) and PACT (public perception of security and privacy- http://www.projectpact.eu/). Similar to the SurPriSE project, these projects also analysed the traditional trade-off between privacy and security though with different foci on technologies and practices. For a broader discussion, the three projects decided to organise a joint conference, held in November 2014 in Vienna. For selected papers of this event, see (Friedewald et al. 2017).

lack control over their information as well as effective agency to reinforce their right to privacy. As shown, this has less to do with low privacy awareness but is rather a result of deficient options for informational self-determination; as well as lacking transparency and accountability of the institutions that gather information about individuals beyond their scrutiny. Easing this problem requires more privacy control options as well as effective oversight and scrutiny of privacy-intrusive socio-technical systems and practices. However, the crux is that it is often unclear to what extent a socio-technical system or practice effectively intrudes into privacy and how privacy impacts emerge. It is thus important to come to a deeper understanding of the emergence of impacts on individual privacy to improve impact assessments and consequently also privacy protection. The next sections are dedicated to these issues with the main argument that (socio-technically induced and reinforced) identifiability is a core criterion in this regard.

5.3 Uncontrolled (socio-technical) identifiability as core challenge of privacy

As shown in the previous sections, identification is an intrinsic part of privacy-intrusive surveillance and there are many examples where identity information is exploited for a variety of purposes beyond the control of the individuals concerned. Problems of insufficient individual control also mirror in the perceptions of citizens on the interplay between privacy, security and surveillance. There are certain overlaps between panopticism and identification, in particular as regards the creation and maintenance of information asymmetries. The occurrence of information asymmetries is not limited to surveillance but a classical agency problem between individual and institutional entities. Individuals frequently experience informational disadvantages, lack in control over their information and have little options to protect and enforce their right to privacy. Information asymmetries and agency problems are likely to occur in digital identification processes, where identity information can be used for multiple contexts.

While the exploitation of identity information for different forms of surveillance is evident, the diversity of privacy-intrusive practices and technologies involved complicates to grasp the relevant factors enabling privacy intrusion. Irrespective of this diversity, a crucial determinant of the emergence of a privacy impact is identifiability. Identifiability is here understood as the possibility to process information that refers or relates to the identity of a person. This aspect is important for several reasons: firstly, because it implies the existence of one or more pieces of identifiable information, suitable to directly or indirectly identify an individual. Consequently, identifiability determines the occurrence of a privacy-affecting information asymmetry. This asymmetry occurs when an individual is identifiable and cannot control whether to be identified or not, i.e., lacks in ISD. Only in a state of anonymity, there is no such information. Secondly, identifiability is particularly relevant in socio-technical contexts as ICTs significantly reinforce options of identification. Many of the outlined surveillance practices benefit from the wide

availability of digital information in social media and other online platforms, and basically from the identifiability inherent to ICTs.

At first glance, the nexus between identifiability and privacy is obvious as the focus of privacy and data protection ever was on personal information. However, ICTs altered this nexus: as shown in the previous sections, different forms of identification mechanisms are embedded in a number of socio-technical contexts, entailing an expansion of identifiability and thus a reduction in anonymity. Today, our identities are widely exposed to the dynamics of ICTs. While basically, every interaction of an individual produces information that can refer or relate to her, in analogue settings without technologies, this information diminishes and is not persistently available (e.g., conversations, movements, activities etc. are usually neither recorded nor in any other way easily reproducible without technological means). Digital technology significantly changed this setting and reduced natural areas of anonymity as physical and digital interactions are observable and reproducible. ICTs generally contribute to a further growth in the amount of personal information not just because they reinforce its processing across multiple domains; but also because their usage can create additional information which may refer to individual identities. Hence, usually, every interaction in a digital environment creates explicit and implicit informational traces, often suitable for identification. This issue is described as identity shadow.

5.3.1 The identity shadow – explicit and implicit identification

Identifiability, i.e., the availability of suitable information to identify the individual is the precondition of identification. The extent to which a person is identifiable strongly depends on whether the information is sufficient for unique identification. If this information is not already an identifier, further information may be needed about the person. A combination of different sets of information then can enable identification. ICTs provide many options in this regard to combine and aggregate information, and thus facilitate different forms of identification. Identification may occur as an explicit as well as an implicit part of an information process. Explicit identification means that identity information of a person are being processed at her knowledge. In contrast to that is implicit identification, which can also occur as a "side effect" of an interaction in a socio-technical context; e.g., while a person is using a technology, service, visiting a website or searching the Web etc. Implicit identification such as targeted advertising, user tracking or profiling happens en passant, e.g., during an online session. Hence, identity information is often not gathered directly from an individual, but rather from the applications, technical devices etc. referring to her, which are then aggregated. The collection of different metadata, for instance, is a common practice of providers, who frequently argue in their privacy policies that this kind of information is merely collected to improve service quality and user experience.[90] In fact, this kind of information is often used for profiling activities of third parties as well. System-immanent or implicit forms of digital identification mostly proceed undetectable from the individual and are therefore difficult to grasp. An important issue is that

[90] This common argument refers to the issue of informed consent in privacy regulation, which is discussed in Section 5.4.3.

technological design can provide identifiable information as a by-product which significantly fosters implicit identification. A simple example is a Caller-ID of a phone call. This technology was introduced in 1988 (cf. Marx 2002) and before that, phone calls were usually anonymous per default. The caller-ID changed this practice of phone communication to the opposite as the phone number (which is an identifier) became visible by default. This made it easier to identify a caller and complicates anonymous phone calls as a caller then has to proactively hide her number. Moreover, the call-receiver may perceive this as suspicious. This is a simple example for a technology having an embedded identification mechanism. Certainly, in this case, there is usually no privacy problem given as individuals want to communicate with each other and mostly know themselves anyway. Nevertheless, this case demonstrates how technology changed the default setting to being identifiable. Similar examples can be found in ICTs and always-on devices, though more complex and with greater implications for privacy.

This setting inherent to ICTs can be called identifiability by default mechanism. From a privacy perspective, this mechanism is critical as it enables a myriad of ways to exploit information about individuals beyond their knowledge and control. In this regard, identifiability by default may be seen as antagonist to privacy by design. To conceptually highlight the problem of hidden or implicit identification I introduced the term "identity shadow"[91] (Strauß 2011: 210) in recognition of Alan Westin's (1967) data shadow who broadly framed data availability as general privacy problem. However, data per se does not affect individual privacy but the processing of data referring to an individuals' identity. Thus, the identity shadow "comprises all the data appearing in a digital environment which can be used to (re-)identify an individual beyond her control and/or infringe her privacy" (Strauß 2011: 210). Re-identification or de-anonymization can be achieved by, e.g., aggregating different data attributes from multiple contexts which enable the gathering of semi-identifying data or quasi-identifiers (cf. Sweeney 2002; Henriksen-Bulmer/Jeary 2016). Although these data are not necessarily unique they can at least reduce uncertainty about an identity as they refer to an individual. Thus even seemingly harmless (non-personal) information can affect one's privacy as they enable cross-linkage of identity information over different (normally separated) contexts. A typical example is the combination of date of birth, gender and zip code which can be used to create identifiers that are likely unique to e.g., determine major parts of the population in the US, as Sweeney (2002) demonstrated. Consequently, the absence of typical identity information such as a persons' name is not a barrier of identification. Similar approaches are feasible by combining other data sets. The use of ICTs in general entails manifold options for re-identification as many (semi-identifying) data attributes are required or created as a by-product during a user session. These data sets are here called common data (CD) and a distinction can be made between person-specific (typically name, date of birth, address, ZIP etc.) and technology-specific data (e.g., IP address, network address, device identifiers, metadata etc.) (Strauß 2011).

[91] The term identity shadow takes up the panoptic metaphor and the problem of information asymmetries. The interplay of light and darkness determines the cast of a shadow which is mostly beyond one's control. For one self, her shadow is mostly invisible and impossible to catch. The only option is to gain control over the amount of light by which one's identity is illuminated.

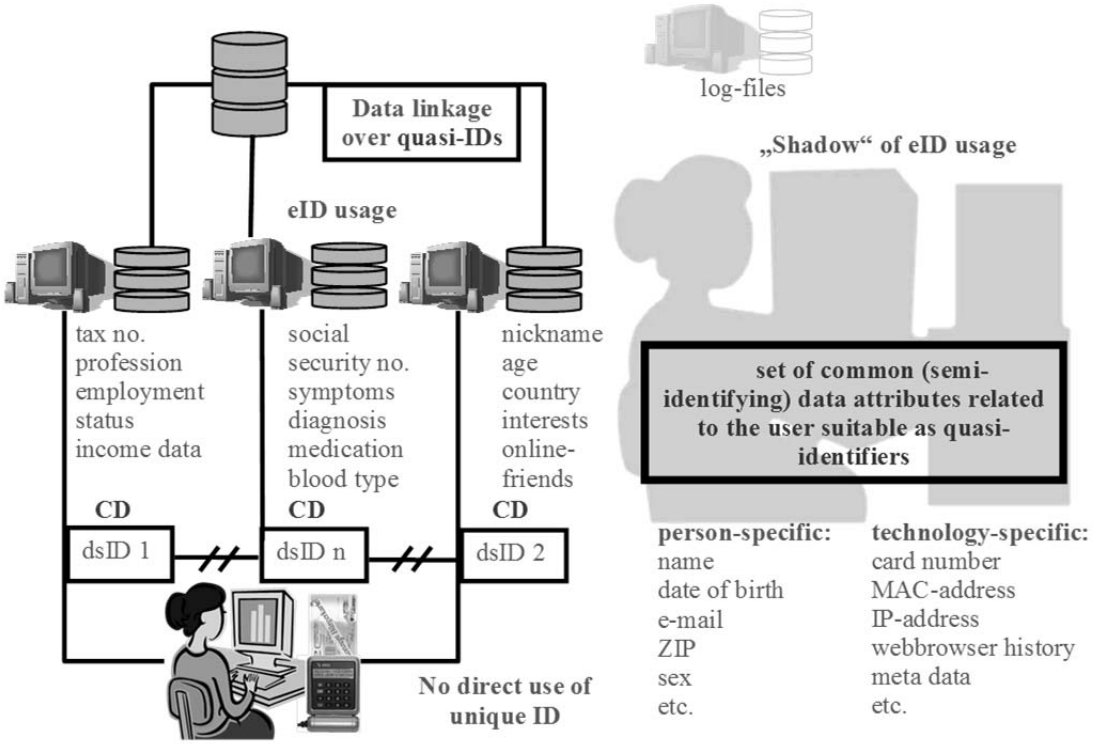


Figure 19: The Identity Shadow (adapted from Strauß 2011: 211)

Figure 19 illustrates the identity shadow on the example of Alice using a digital identity device (eID) for a wide array of different services such as e-government services, health services or accessing a social media platform. Each of these services may have either a specific IDM approach or the ID device applies an unlinkability concept (e.g., with domain- or sector-specific identifiers – dsID for each service as in some governmental eIDMS, see Sections 3.1.3 and 5.4.2). A separation of usage contexts can be achieved with both approaches, if the use of a unique ID is avoided. However, the main problem addressed here is that regardless of the use of domain-specific identifiers to avoid linkability of digital information over separated contexts, there is a risk of gathering quasi-identifiers to break the unlinkability concept and cross-link information, e.g., by exploiting common data. Even if information directly referring to one's identity is removed from a data set, there are other options for re-identification. A simple example concerns log files and protocols. Although their aim is mostly to detect unauthorized access and protect from abuse, they can be exploited for privacy intrusive profiling as they usually provide detailed information about user activities. The depth of the identity shadow in a particular usage context depends on the amount of common data available in that context and the technologies involved. For instance, online activities usually leave a number of data traces that can be exploited for de-anonymization: besides the very common form of user tracking based on web-cookies or click-tracking etc., web browser data (e.g., bookmarks, history of visited sites, browser configuration etc.) allows generating a digital "fingerprint" for unique identification of a person (cf. Eckersley 2010; Schrittwieser et al. 2011). A recent study demonstrated a fingerprinting-based identification rate of over 90 per cent

without using IP addresses or cookie data (Cao et al. 2017). Besides fingerprinting techniques based on web browser data there are similar approaches such as using time stamps of applications for de-anonymization (Bager 2016), or even the battery status of a technical device (Kleinz 2016). Social media in general process massive amounts of identity information, give deep insights into personal details and provide many options to gather quasi-identifiers. Integrative technologies such as social plugins (see Section 4.3.1) enable websites to gather and process identity information of a user from her social media profile. Hence, user details are also available to external services; e.g., name, profile photo, gender, networks, user ID, list of friends as well as status updates, comments, shared content etc. Moreover, different ways to re-identify individuals out of anonymized data ranging from analysing group associations, user preferences, online activities, social graphs, face recognition based on user photos, location data etc. were demonstrated by several security scholars (e.g. Narayanan/Shmatikov 2009; Wondracek et al. 2010; Nilizadeh et al. 2014; Gulyás et al. 2016). As social media extends is scope with social plugins and graphs, user details are also available for external sources outside the platforms. Further issues result from mobile computing. Compared to a smart card, mobile devices such as smart phones entail a number of additional information (e.g., phone no., device identifier, SIM card no., geo-location, IDs of Wi-Fi networks etc.). As smart phones can access the Internet, they also have according network identifiers, e.g., IP and MAC[92] address. Also ID's of a favourite Wi-Fi network can be used for user tracking (Vaas 2013).

A crucial part of the identity shadow problem results from identification mechanisms as integral parts of ICTs, i.e., identifiability by default, which can involve several sub-systems such as hardware devices, databases, software applications etc. The connecting function inherent to identification (as outlined in Section 3.1) is basically important to enable interactions between different entities or systems. When different entities (e.g., humans, technologies or applications) interact with each other (e.g., exchange of information, communicate, or co-operate), some forms of identification are typically involved. The processing of a piece of identifiable information (such as an identifier) is a precondition for networked systems to establish a connection. Therefore, technical devices are usually equipped with identifiers (e.g., a card number, network address, IP address etc.). This allows distinguishing one particular entity (a technical system or device) from another and enables interaction. Hence, identification creates an informational link between two or more entities. This is true for Internet connections[93] but in principal, some form of identification process occurs in every kind of network. Technical devices are thus usually equipped with an identifier enabling their identification in particular contexts. For instance, a personal computer connects to a network such as the Internet via an IP address, a website is accessible under a particular (unique) domain (e.g., www.internetsociety.org); e-mail communication requires a valid e-mail address identifying a client; a mobile phone has a particular serial number (IMEI[94]) and SIM[95] cards have a unique identifier (IMSI[96])

[92] The Media Access Control address identifies a technical component in a network such as a network interface.
[93] Such as TCP, see, e.g. Kozierok (2013).
[94] International Mobile Equipment Identity https://de.wikipedia.org/wiki/International_Mobile_Equipment_Identity
[95] Subscriber Identity Module https://en.wikipedia.org/wiki/Subscriber_identity_module
[96] International Mobile Subscriber Identity https://en.wikipedia.org/wiki/International_mobile_subscriber_identity

to interact with the phone network and a unique phone number to enable communication; even a simple Bluetooth device such as a headset is identified by the computer device it is linked to; RFID (Radio Frequency Identification) tags can be used to equip all kinds of objects such as clothes etc. with an identifier. Further examples can be found easily. Thus, as these common examples highlight, identification is an intrinsic process of ICTs or socio-technical systems shaping their connectivity and interactivity. Identification allows creating a sort of strong tie that links different systems in analogue or physical as well as digital environments. Indeed, technical forms of identification are different from human identification of individual persons. Hence, they do not necessarily lead to the identification of an individual person. However, as the identity shadow highlights, both forms can overlap as technical identifiers provide potential links to the persons using a technology. Identifiability by default mechanisms of technological design thus also affects the identifiability of the person. As identity is a relational concept, digital identities of individual persons can be expected to become increasingly networked by using technical devices or interacting with informational entities or agents (e.g., a device, a service, an application or any other kind of information system). These entities can then directly or indirectly refer to the individual. This may be a technical device such as a smartphone, but also a software application such as a social media or cloud computing account etc. Hence, put simply, the "identity" of a technological system can refer to the personal identity of its users. Through interactions with digital information systems, the properties of an individual's identity may be extended as the information representing her become virtually available. Moreover, technology enables the use of this digital identity representation also decoupled from its source. Therefore, the number of (sub-)systems involved in a user interaction or processing of user-related information has an impact on the identity shadow of this particular user. With a growing number of (sub-)systems the identity shadow is likely to grow as every usage context can potentially generate additional identity information. Hence, ICTs entail an incremental extension of the representation of our identities and affect identification. This identity representation is not always visible but, metaphorically speaking, mostly hiding in the shadows.

In total, the identity shadow can have many different shades. It may be seen as reflection of an individuals' identity, which concurrently morphs by enabling new space for de- and re-contextualisation beyond her control. This undermines the privacy principle of purpose binding and hampers contextual integrity (see, e.g., Section 3.2). Increasing amounts of digital environments and thus digital information bring further possibilities to (re-)identify and de-anonymize individuals even when they do not provide personal information directly. Hence, the boundaries between personal and non-personal information, public and private spheres blur as, e.g., observable in social media. Consequently, socio-technical contexts conflate which results in expanding identifiability and various forms of implicit identification.

5.3.2 Contextual identity layers

From a wider perspective, the problem highlighted with the identity shadow can be illustrated as a model consisting of expanding contextual identity layers. As outlined in section 3.1, identity does not shrink and thus identifiable information is likely to continuously grow with every usage in the course of time. At a particular moment in (space and) time, an individual may be represented by a particular number of identity attributes, i.e., a set of identifiable information. For instance, when a child is born, the initial information gathered about her includes a set of bodily characteristics (such as height, weight, gender, biometric features) as well as temporal, spatial and relational information (e.g., date of birth, place of birth, birth name and names of parents) that depicts birth. This is identifiable information referring to an individual and the amount of this kind of information usually expands over time as its socio-technical usage contexts grow as well. Given the dynamics of identifiable information and the multiplicity of contexts, the exact amount of information that identifies an individual in a particular context is mostly not strictly determinable. This aspect is crucial to grasp how identifiable information in general emerges and proceeds (for a more detailed discussion see Section 6). Metaphorically, an identity lifecycle can be seen as a sort of spiral where the individual is the very entity that is represented by a flow of identifiable information expanding across multiple, circular layers (contextual identity layers, as illustrated in Figure 20).

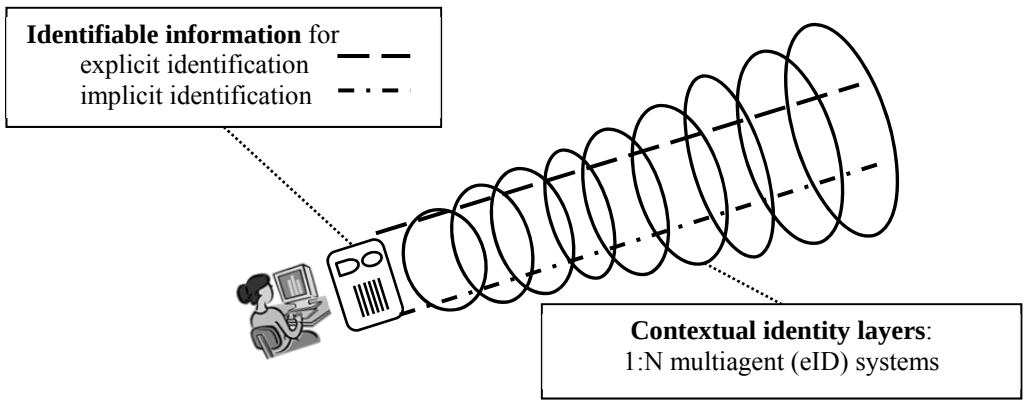


Figure 20: Spiral of digital identification. The circles of the spiral represent contextual (identity) layers and point out that identifiable information can be repeatedly processed in different contexts.

Multiple entities and can be involved in various application contexts. These entities can be other individuals as well as institutional entities (e.g., government agencies, employers, private companies, law enforcement, intelligence agencies etc.). Most of the institutional entities use applications with repositories (e.g., dossiers, registers, databases etc.) to store and process identifiable information. The involvement of ICTs complicates the identity spiral because due to the nature of digital information, contexts can easily overlap irrespective of their boundaries. Additional, multiple layers emerge that involve also

virtual, non-physical entities (e.g., services, applications, databases, repositories, technical hard- and software devices etc.). In this regard, the representation of a personal identity incrementally turns into a digitally networked identity. This digital identity can provide deep insights into individual privacy, i.e., who a person is, what she does, where she is and was, her personal and professional relationships, preferences, interests, behaviour etc.

On a meta-level, the totality of identifiable information referring to or representing a particular individual and the increasing conflation of usage contexts may be perceived as the virtual embodiment of a meta-ID. Norbert Wiener (1954: 96) stated that: "We are not stuff that abides, but patterns that perpetuate themselves". This aspect of (self-)perpetuation is particular relevant against the background of digital identification and identifiability. As shown, there are many forms of implicit and explicit identification mechanisms in a variety of contexts. Context can be understood as the condition of an application (cf. Nissenbaum 2010: 140ff.) or informational process. From an individuals' view, in many cases there may be only a one-dimensional context perceivable as he or she is only aware of the current application interacting with while in fact there may be a 1:n relation with multiple contexts in which her information is being processed. These additional, hidden contexts refer to the identity shadow problem as outlined previously. From a systemic perspective, a contextual layer can be seen as a relational entity (or sub-system) that may be linked to other subsystems which may further process identifiable information. For the individuals concerned, these multiple contextual identity layers mostly imply information asymmetries and lack of ISD. Firstly, because the individual can hardly avoid being subject to implicit identification (occurring in hidden, uncontrollable contexts). Secondly, because the thresholds between different contexts and related (sub-)systems may be unknown or beyond control as well.

It makes a difference whether identifiable information is bound to a particular context only or whether it is used for other contexts as well. The possibility for re-contextualization is thus a privacy risk which is in contradiction to a basic privacy principle, namely purpose limitation (or binding) (see also Section 3.2.1). The crux here is to clearly determine what counts as a context respectively a system that processes identifiable information. In general, a privacy relevant context involves the processing of identifiable information. In this regard, the processing of identifiable information is a necessary condition for a privacy-affecting application context. In many cases, a context may be a single application (e.g., an e-government or e-commerce transaction, the submission of a registration form etc.) in which identifiable information (e.g., name, date of birth, e-mail address) is processed with direct involvement or interaction of the individual. However, in addition to the primary context of the initial application, there may be other systems involved as well. Hence, other contextual layers respectively subsystems (e.g., technical devices, hard- and software systems) may affect privacy but remain undetected. Each system (or contextual layer) processes identifiers which can refer to an individual identity and thus can also be used for privacy intrusion. Thus even without direct involvement, a user can be identifiable as her information is associated with a technical device. As shown previously with the identity shadow, the same process can be found in other technologies and applications, where multiple types of technical identifiers may refer to a particular user. This information emerges from the information systems

(sub-systems) that are entailed to an application context. For instance, a standard web session at least involves three different information systems: a personal computing device, the operating system and the web browser. Via interfaces, information may also be transferred from one information system to another (as e.g., observable in social media platforms providing APIs to external entities for the use of social plugins, logins etc. as outlined in Section 4.3).

Hence, through the potentially unlimited flow of identifiable information, a digital identity can be involved in multiple processing contexts of multiple (individual, institutional as well as technical) entities, without the knowledge of the person concerned. In the case of explicit identification, the individual is involved in the emergence of the input information (e.g., by entering personal information into an information system or triggering an according event). Opposed to that, implicit identification happens without direct involvement of the individual whose information is processed. Implicit identification can occur during a user session regardless of whether the person is also explicitly identified by personal information. Instead of the person, an application, system etc. may gather or generate identifiable information about that person and pass it on to another system etc. For instance, when an application automatically gathers a user profile and transmits it to another system. These system-immanent interrelations can trigger a cascade of processes in which input and output of identifiable information can oscillate between multiple systems. Depending from the temporal availability of identifiable information, there may be longer time spans between input and output without any direct involvement or awareness of the person whose identity information is processed. In a worst case, the uncontrolled processing of identifiable information entails a cascade of identification processes where the individual is uncontrollably identified in many contexts without even noticing it.

5.3.3 Trends of expanding identifiability

Technology usage and technological progress entail a further expansion of the identity shadow and contextual identity layers in many respects. An increasing number of networked always-on devices reinforce permanent availability of digital information flows and thus identifiability. An important example concerns the constant growth in mobile computing with portable devices such as smart phones, or wearable technologies (e.g., smart watches etc.). These devices can serve as technological hubs to steer other technologies and provide additional applications such as location-based services (LBS) as well as information about geo-location and movements of individuals. Hence, mobile devices significantly extend user tracking which is not limited to digital spaces but now also includes a persons' movement in the real world. Consequently, it becomes increasingly difficult to avoid being tracked (cf. Clarke 2012a). Tracking information allows creating identity profiles including location and movements, which is another indicator for increasing overlaps between physical and digital environments. There already are a number of business models in the realm of big data aiming at monetizing telecom data (cf. IBM 2013; Leber 2013; Kannenberg 2016). Big data and datafication in general boost the trend to digitally gather and process maximum data from every-day-life contexts (Mayer-Schönberger/Cukier 2013; Lyon 2014; Strauß 2015a). The myriad of "smart"

technologies, networked devices towards trends such as the Internet of Things, ambient intelligence, pervasive computing etc. yield further ICTs that feed the big data paradigm. With these developments, the room of possibilities expands to gather unique patterns in digital (or digitised) information sets. Once a unique pattern is found, it can be used to create a quasi-identifier and apply de-anonymization techniques. Thus, the identity shadow is obviously of high value for the surveillant assemblage including the NSA and other security and surveillance actors (Section 5.1). The presented cases of ICT-related surveillance ranging from targeted advertising, various forms of profiling, up to global mass surveillance programs basically exploit the identity shadow of their targets by, e.g., creating and employing identity graphs for extensive profiling activities (as mentioned in Section 5.1.2). These identity graphs exploit all kinds of information about individuals from multiple sources and map them based on graph theory and computational models. That ICTs and networked technologies of various kinds are valuable for surveillance was not least highlighted by James Clapper, the chief of US intelligence: he mentioned that the NSA and other intelligence agencies will probably exploit the Internet of Things to spy on individuals: "In the future, intelligence services might use the internet of things for identification, surveillance, monitoring, location tracking, and targeting for recruitment, or to gain access to networks or user credentials" (cf. Ackerman/Thielman 2016).

Besides developments fostering the extension of technology-specific identifiable information, also the amount of person-specific information increases in many respects, not least as regards biometrics. The employment of biometric systems handling them is a high ranked issue in many security strategies (cf. Lyon 2009; Ajana 2013). For several years, there is a growth in the use of biometrics and face recognition observable. Biometrics in a technical sense is defined as "the automated recognition of individuals based on their behavioural and biological characteristics." (ISO 2010: 2). The scope of usage of biometric technologies ranges from access control in security areas to border and migration control, law enforcement, or different kinds of profiling activities. Biometrics is also gains in importance as an integrative feature of laptops or smart phones equipped with fingerprint scanners (e.g., Apple's "TouchID"). The proponents of biometrics see it as a sophisticated and highly secure means of identity verification or authentication. However, in fact, biometric identification is a form of pattern recognition that calculates a probability: A biometric system is basically an automated pattern recognition system that uses information about one's body, e.g., for identification, access control etc. It either makes identification or verifies an identity by establishing the probability that a specific physiological or behavioural characteristic is valid (Wilson 2010). Hence, biometrics approximates the validity of identity information based on a probability pattern. This also implies that a certain degree of false positives or errors is possible. Several studies pointed out that biometrics entail security and privacy risks (e.g., Prabhakar et al. 2003; Clarke 2001; Lyon 2009; Acquisti et al. 2014; Sharif et al. 2016). Besides that, biometric information can be copied and abused. For instance, security researchers demonstrated the insecurity of fingerprint sensors embedded in smartphones with relatively simple means: based on the photo of a fingerprint from a glass, the hackers created a synthetic digit (Arthur 2013). In other demonstrations, hackers show how to gather biometric information by using a standard digital camera. With relatively simple methods, even iris information

could be reproduced (Krempl 2015). It is thus questionable whether approaches to use biometric features for authentication are as secure as they are promoted.

While the employment of biometrics used to be rather limited to special security domains (e.g., for access control in high security areas), there are trends to include these technologies also in other domains. For example, authentication via fingerprint in mobile apps,[97] or novel forms of payment via facial image ("pay-by-selfie") (Leyden 2016) are promoted in the banking sector.[98] In particular facial recognition gains in importance in public as well as private sector (cf. Acquisti et al. 2014). Security and law enforcement agencies increasingly employ facial recognition for preventive surveillance activities. In 2011 the US Federal Bureau of Investigation (FBI) initiated the so-called "next generation identification" program (NGIP) to set-up the largest biometric database worldwide (Reardon 2012; EPIC 2016). Among the stored information will be, e.g., fingerprints, iris scans, DNA profiles, voice identification profiles, palm prints, and photographs. Facial recognition is among the core features of this surveillance system. The information is gathered from external sources such as police databases, CCTV systems, and similar systems from public as well as private entities (EPIC 2016). According to a study of the Georgetown University in Washington in 2016, about 117 million facial images are stored in governmental databases in the US. The collection inter alia consists of images gathered from driving licenses, passports as well as police databases. The majority of these images refer to normal citizens that have never been involved in criminal activity (Garvie et al. 2016). The US Government Accountability Organization (GAO) evaluated the facial recognition system of the NGIP and found several flaws as regards the accuracy of the technology. The GAO came to the conclusion that NGIP raises concerns about the protection of privacy and individual civil liberties (GAO 2016). Facebook introduced a facial recognition system already in 2010 (Isaac 2010). This system exploits the broad availability of images online (as e.g., boosted by digital cameras and smartphones). It enables recognition of a person based on its presence in photos uploaded to Facebook as well as automatic tagging which allows for, e.g., searching particular persons. The app "Facedeals" related to Facebook uses facial recognition "to provide a 'seamless method' for customers to announce their presence at restaurants, cafés, bars and other venues on Facebook" (De Andrade et al. 2013: 61). Also Google and Apple have similar software in their photo apps (Brandom 2016; Monckton 2017). Another application called SceneTap uses facial recognition for age and gender determination in bars etc. to provide people with information about suitable places to meet others (ibid). Facebook was confronted with heavy protest from privacy and data protection authorities in Europe but also in the US due to its facial recognition system. As reaction to the strong opposition and complaints of regulators in Europe, in 2012 Facebook decided to turn off the facial recognition feature in the EU for new users (De Andrade et al. 2013). With this strategic measure, Facebook avoided lawsuits with the EU. However, it does not really solve the manifold privacy problems as Facebook is a global service located in the US. In 2016, Facebook started to

[97] For instance, the fingerprint authentication of the federal credit union: https://www.gfafcu.com/fingerprint-authentication/ or the Unicredit group, see e.g. (Sayer 2013).

[98] A very drastic case exists in China, where female students were requested to provide nude photographs of themselves as condition to obtain loan (Connor 2016).

promote a new photo app which includes its facial recognition technology "DeepFace"[99] also in Europe, though, with limited functionality due to repeatedly heavy protest and legal issues with European regulation (Hern 2016). Users in other countries that do not want this feature automatically activated have to change their privacy settings O'Donnell 2017). According to media reports, Facebook's facial recognition software is more accurate than the FBI's NGIP (LaChance 2016). Further trends in the field of virtual and augmented reality (AR) can be expected to amplify the identity shadow and its privacy impacts. Smart devices such as wearable computers (e.g., "Google Glass" or other smart glasses) may indicate the next generation of mobile social media. Such devices may have the capability of bi-directional tracking via an integrated camera and thus could gather one's movements as well as what one looks at, including facial recognition etc. (Strauß/Nentwich 2013). But AR/VR could also be used to trick facial recognition: just recently, a technique to spoof facial recognition by employing a VR model was demonstrated (Cuthbertson 2016).

In total, these developments indicate a further convergence of physical and digital environments involving identification technologies. The more physical and digital worlds converge the more networked our identities become which further strains privacy protection. This is mainly because it becomes increasingly difficult to control the information flow from one system or environment to another, and thus, ISD suffers from this increasing convergence. Biometrics is a particular form of identification that underlines this issue. Biometrics demonstrate that "(…) the line between technology and the body is blurring; identities are defined in terms of bodily features as – a fingerprint or a facial image – captured in an algorithm" (Lyon 2009: 17). As a consequence, the body itself becomes a sort of code. Clarke (2001) early warned from the manifold risks of biometrics to privacy, as its extensive use can significantly reduce anonymity up to threats of dehumanization as individuals may be reduced to a set of biometric information. Although the latter risks may seem overstated, a scenario where biometric information is extensively processed indeed raises serious privacy issues. Considering different forms of biometrics being a core technology of every-day-life, then our bodily characteristics were the main representations of our identities. At the same time, identity representations where exposed to the risk of being permanently available for uncontrolled identification. This would reduce the options of individuals to autonomously express their identities and freely decide whom to reveal what informational parts of it in which contexts. The permanent possibility to be identified by biometric recognition drastically aggravates one's possibility to avoid being observable and traceable online as well as offline because one can hardly opt-out from her bodily characteristics. Consequently, the increasingly blurring contexts and boundaries of the public and the private sphere also affect the human body. These blurry boundaries are also an issue in contemporary art and activism which seek ways to deal with privacy threats. For example, an "identity prosthesis", i.e., a facial mask to undermine facial recognition systems, developed for an art project.[100] Similar approaches

[99] The technology was presented in 2014 by Facebook developers at a conference on computer vision and pattern recognition; see (Taigman et al. 2014).

[100] URME Personal Surveillance Identity Prosthetic, http://www.urmesurveillance.com/urme-prosthetic/ see also: Aktivist trickst Gesichtserkennung mit 3D-Drucker aus, *Der Standard Online*, August 16, http://derstandard.at/2000043002831/Aktivist-trickst-Gesichtserkennung-mit-3D-Drucker-aus

exist to trick fingerprint sensors with e.g., synthetic fingers[101] or to hide from surveillance systems by using specific clothing, i.e., stealth wear[102]. Facial recognition systems can also be undermined by using colourful glasses as security researchers recently demonstrated (Doctrow 2016; Sharif et al. 2016). These approaches indicate an ongoing societal discourse about the increase of digital identification and privacy; where among other things, civil society seeks for ways to deal with these issues.

As shown, there are manifold socio-technical developments that contribute to a further growth in the identity shadow and identifiability in general. ICTs or digital technology in general can affect identity representations, particularly when processing contexts are basically unrestricted. Given the possibilities of digital information processing, digital technology can transform initially non-persistent information produced during an interaction and enrich it with a certain property of continuity. The information is then digitally represented and can be made available also outside the context of the original interaction. While continuity characterizes an intrinsic part of identity as regards self-development and identity-building (Ricoeur 1992; Hildebrandt 2006, as outlined in Section 3), it is not explicitly mapped or documented in the analogue world, and thus not permanently available per default. Moreover, in an ontological sense, continuity is incompatible with identity as Sartre (2014: 263f.) stated. In this regard, permanent availability of identity information without any frictions or discontinuity may complicate self-determination. In line with privacy, identity-development needs some (ontological) frictions (Floridi 2013). Hence, identity also implies a certain solitude and seclusion so that it is unique and distinguishable from other entities. To some extent, the digital world strains this natural discontinuity, e.g., exemplified by social media platforms, where one's interests, actions, behaviour etc. in the past and in the presence can be observable. An example to highlight this issue is Facebook's "timeline" feature introduced in 2011, which visualizes how a user profile emerged and developed in the course of time (Panzarino 2011; Kupka 2012). The timeline can contain all details and stages of an individual user's life on Facebook, i.e., a sort of automatic autobiography including all events, photos, shared content, posts etc. This kind of information is not merely available for the particular user but can be accessed by external sources as well. Considering Ricoeur's (1992) distinction between the two basic types of identity – idem and ipse (see Section 3.1), the dynamic character of digital information enables further growth in both types while at the same time, the distinction increasingly blurs as the timeline example illustrates. The availability of information referring to ipse, i.e., social relations, associations, interactions, expressed thoughts and feelings, postings, likes, behaviour, shared thoughts and feelings etc. can be particularly problematic from a privacy perspective. This kind of information provides deep insights into one's identity and also bears risks of misinterpretation and prejudice. For instance, the idea to use Facebook profiles for national security (as mentioned in Section 4.3) entails a high risk that irrelevant information about a person is used to classify this person with a low-threshold to discrimination. ICTs significantly extended the scope of identity information and its

[101] Identity counter-surveillance kit, http://mian-wei.com/#/identity/
[102] Privacy Gift Shop - Stealth Wear collection, http://privacygiftshop.com/shop/stealth-wear/

naturally limited availability. Identity information can be processed regardless of whether the original context of an identification process ceases to exist. As a consequence, identity information from multiple contexts can be explicitly visible, accessible, observable, collectable, and to some extent also manipulable beyond the control of the individual concerned (as shown in the previous sections). This also has conceptual implications on the quality and scope of privacy, not least as regards is temporal and spatial dimensions. Privacy provides a space free from (factual and perceived) interference which implies a certain distance between the individual and society. In a traditional sense, privacy may be seen as a concept determined by certain zones or territories (cf. Kahn 2003), which were ever linked to social convention and norms and included spatial as well as virtual domains. Although this setting is still given, it is altered and challenged by increasingly pervasive ICTs. Hence, these "classical" privacy territories become digitised and thus informationally enriched. Spatial and temporal dimensions (e.g., the physical domains in which an individual is located at a certain point in time) of identity representations are more easily reproducible and accessible. Digital identities are exposed to the transparency reinforced by ICTs (see also Section 3.2.2). A technological device does not have to be at the same physical location of an individual at the same time to gather and process her information. This can paradoxically entail a form of permanent virtual proximity or closeness between technology and individual, a sort of quasi-persistent informational link (as e.g., already observable by smart phones). Besides other things, this virtual proximity enabled distant identification of an individual, irrespective of the spatial and temporal distance to the person, as her digital identity representation can be processed at multiple locations at the same time. This multidimensional, networked character of digital identity has many benefits. But the inherent co-referentiality between technology and identity (Floridi 2013) also complicates privacy protection as the informational boundaries of those contexts, in which an individuals' identity is processed (by personal as well as socio-technical entities) partially diminish. These diminishing boundaries further challenge the achievement of contextual integrity (see Section 3.3.2) as a digital identity can be subject to multiple processing contexts. An elementary characteristic and precondition of privacy and contextual integrity is its contextual separation or informational friction. This means that the flow of personal information is (ideally) not uncontrollably streaming from one context to another. The concept of unlinkability is a crucial requirement to technically implement these frictions (see Section 3.1.3). While this is mostly the case in the physical or analogue world, digital environments challenge unlinkability, and the individuals' capabilities of informational self-determination are hampered if not impeded. Against the background of a continuing convergence between physical and digital environments with expanding identifiability, a further reduction in ISD is very likely. Permanent identifiability thus repeals the boundary control function of privacy and hampers self-determined inclusion and seclusion in particular socio-technical contexts. An individual then cannot control whether to be identified or not. Therefore, enhancing the boundary control function of privacy by regaining control over identifiability and digital identity representations is among the core challenges of privacy. The next section presents and discusses the extent to which privacy by design provides suitable options in this regard.

5.4 Privacy controls – prospects and perils of privacy by design

The previous section explained the problem of the identity shadow and how technology extends identifiability and identification. As demonstrated, the design of technology significantly contributes to this problem with identifiability by default mechanisms. Therefore, it seems promising to foster privacy by design and privacy by default as integral features of technology to reduce identifiability. This section explores the prospects and perils of privacy by design, including a discussion of these issues.

5.4.1 Overview on scope and approaches

The rapid growth in ICTs and digitally networked technology in general reinforced the demand for privacy by design (PbD). PbD can be understood as an attempt to entrench privacy-friendly technology development and use in the informational ecosystem (Goldstein 2011; Cavoukian 2012a/2012b; Cavoukian et al. 2014; Danezis et al. 2014). Although approaches for so-called privacy-enhancing technologies (PETs) exist for several years, they are yet rather used in niches but not at a larger scale. PbD aims at stimulating PET development and usage as well as at fostering that technology is per default equipped with privacy-preserving mechanisms (Goldstein 2011). It concerns the implementation of mechanisms to safeguard privacy as a built-in feature of technology. However, it is less of a technical approach but a broad concept that embraces PETs and carries forward its approach to organizational handling of personal information.

According to Cavoukian[103] (2012a: 3ff.) there are seven foundational principles of PbD: (1) Proactive not reactive, and preventative not remedial, meaning that privacy protection is to be proactively considered and implemented in advance to reduce risks and not just when a risk or breach occurs. (2) Privacy as the default setting, i.e., the standard mode of technology should be to protect privacy and not to disclose personal information. (3) Privacy embedded into design, i.e., as activated, built-in feature of information systems and practices. (4) Full functionality – positive-sum, not zero-sum, i.e., dichotomized views and constructed trade-offs such as privacy versus security should be avoided as far as possible so that both issues are seen as equally relevant issues for system functionality. This also implies that PbD is not to be implemented in a way that hampers the functionality of a system. (5) End-to-end security life-cycle protection, i.e., a consistent protection of personal information at all processing stages from collection, use, storage to erasure of information. (6) Visibility and transparency, so that information processing is comprehensible and it is verifiable whether privacy is respected. (7) User-centricity, i.e., respect for individual user privacy so that privacy protection is designed in a user-friendly way whereas the user and her personal information is understood as central part of the system.

A diverse range of technical PbD-related approaches and privacy-enhancing tools can be found in many domains. For instance, web browser plugins or add-ons to prevent from

[103] Ann Cavoukian was the Information and Privacy Commissioner of Ontario, Canada and is among the first advocates of PbD.

advertising, user tracking and profiling[104], privacy-friendly search engines (e.g., Startpage.com, Duckduckgo.com), social network sites (e.g., Diaspora - diasporafoundation.org), e-mail encryption (e.g. Pretty Good Privacy – PGP, openpgp.org), anonymous mail clients that allow to send and receive e-mails without the need to proof one's identity to a mail provider (e.g., Hushmail.com) secure online messengers (e.g., Signal - whispersystems.org, crypto.cat), encrypted cloud services (e.g., Spideroak.com), anonymization services (e.g. proxy servers or anonymisation tools such as JonDonym - anonymous-proxy-servers.net, or the Tor network), privacy-friendly operating systems (e.g., Tails) etc. just to name a few. However, despite of the variety of tools, they often serve very specific purposes for privacy-aware users while a central aim of PbD is to make privacy protection a standard feature of every technology. Irrespective of its relevance, the PbD concept is a rather broad organizational guideline and therefore does not provide detailed technical requirements to improve privacy protection. Privacy engineers thus criticized this issue as it hampers the development of effective technical PbD approaches (cf. Gürses et al. 2011). Cavoukian (2012b) advocates this broadness because different organizations may have different specific requirements for privacy protection. Nevertheless, it seems natural to use common privacy principles (such as the fair information practices of the OECD (2013b) as point of reference (see also Section 3.2.1). Gürses et al. (2011) argue that especially the principle of data minimization is an essential requirement for PbD. This is particularly important in times where big data and similar developments boost the permanent availability and processing of information. Against this background, a counter-paradigm that highlights the positive effects of vanishing information is important to re-balance the privacy discourse. While most applications have privacy settings, their default values are mostly set to full and permanent disclosure of information which is highly counterproductive to protect privacy. A change of this common practice, i.e., a privacy-by-default standard contributes significantly to avoid unintended information disclosure and thus identifiability.

There are several technical options for privacy by design to achieve data minimization to reduce identifiability. This particularly concerns the implementation of anonymity and de-identification, pseudonymity and unlinkability (cf. Chaum 1985; Cavoukian 2012a/2012b; Pfitzmann/Hansen 2010; Danezis et al. 2014). For anonymization and de-identification techniques, the concept of k-anonymity is essential which means that a set of information is indistinguishable from at least k-1 other information sets to be anonymous (Sweeney 2002). Enhanced approaches are l-diversity and t-closeness which aim at increasing the degree of anonymity with greater diversity and less accuracy of data properties (Machanavajjhala et al. 2007; Li et al. 2007). A further, relatively novel concept is differential privacy which uses noisy data to normalize information and thus reduce identifiability (cf. Dwork/Roth 2014). An important basis of all these different approaches is the aim to decrease the applicability of information for identification; e.g., by erasing, masking or obfuscating identifiers or parts of identifiable information, grouping values, or adding noise to increase information ambiguity. One option is to reduce the temporal

[104] Common examples are Adblock Plus, Facebook Blocker, BetterPrivacy, Ghostery, TrackMeNot or the "do not track" feature of several web browsers such as Mozilla and others.

availability of information, for instance, by implementing the idea of an expiration date for digital information in technical systems[105]. This can be an in-built mechanism to partially or completely erase information which is not permanently required after a certain period of time (e.g., communication content), an automatic revocation of information access for external users, or a randomly changing identifier. In many contexts, though, a complete erasure of information is neither feasible nor of practical use. Therefore, mechanisms to remove parts of identifiable information are crucial. With such mechanisms, the functionality of a technology or application can be kept while the risk of unintended information disclosure is reduced, at least to some extent. A very simple example is a dynamic IP address which randomly changes. This can help to avoid that a person can be permanently traced on the Internet by the same address. Another simple, related example concerns the storage setting of web-cookies which are often persistently stored or with an expiration date lying decades in the future. Modern web browsers enable users to decide whether cookies are treated as session cookies which expire with the session they refer to and are then automatically deleted. This simple feature is a low level example for an expiration date. But there are also more sophisticated technical approaches with incorporated expiration dates available: an early approach was the software "Vanish" which combined bit-torrent technology with cryptographic techniques to equip data with a mechanism to self-destruction (Geambasu et al. 2009). Although Vanish had several flaws and was itself vulnerable to attacks as, e.g., Wolchok et al. (2010) demonstrated, the approach as such can be useful in many respects. A more recent example for a similar mechanism can be found in the instant messenger app "Snapchat", where messages are deleted within a defined period of time after being submitted and received. However, this does not sufficiently prevent from third party access to user content, as e.g., pointed out in a recent study on encryption and human rights protection (AI 2016). Another messenger with a similar approach is "Telegram", which offers more security and privacy features than Snapchat (ibid). A similar principle to limit temporal availability of information is applied in the concept of so-called "perfect forward secrecy"[106]. Systems with perfect forward secrecy frequently change the keys used to encrypt and decrypt information (cf. Krawczyk 2005). This improves information security and privacy, because even with a compromised key, an attacker can only gather a small piece of information, e.g., of a chat, phone conversation, or of an online service.

In general, the toolbox of cryptography offers a variety of promising concepts for PbD. Cryptographic functions can be used in manifold ways to improve information security, encrypt content or anonymize information (cf. Menezes et al. 1996). Encryption was ever essential for information security and privacy, but since the Snowden revelations, in generally gains in popularity (e.g. Kuchler 2014; Finley 2014). The encryption of content is a typical application of cryptography in order to keep information and communication confidential. This can be, e.g., achieved by cryptographic hash functions which are a very common security practice. A hash function inter alia enables confidential storage of information without knowing its meaning. This practice is basically used to

[105] For a detailed discussion about the idea of an expiration date see Mayer-Schönberger (2009).
[106] For a brief overview, see, e.g. (Greenberg 2016a)

protect passwords from abuse, which are not stored directly as plain text but only as hash value. The same principle can be used to protect personal information e.g., by avoiding direct use of identifiers (e.g., user ID) and storing only their hash values instead. This can help to reduce the identifying capacity of an information set and improve unlinkability. However, a hash value itself is mostly unique (corresponding to a pseudonym) and therefore, it can be used to breach security and infringe on privacy. For example, the hash value of a password and the username can be sufficient for an attacker to gain access to a user account. Besides that, hash functions can be vulnerable by so-called collision attacks (e.g., Klíma 2005; Preneel 2005), which attempt to find an identical hash value to undermine protection. It is thus crucial that PbD implementations are consistent and not just offer arbitrary protection. A prominent, classic example for content encryption is the software PGP (Pretty Good Privacy)[107] originally developed by Phil Zimmermann in order to empower individuals in their right to privacy.[108] PGP became relatively widespread in the crypto-community and today, there are several variants available (such as the open source software OpenPGP or GNUPG[109]) as well as add-ons for e-mail clients (e.g., enigmail for Thunderbird[110]). PGP is just one among many examples for content encryption. Similar features can be basically integrated into all kinds of ICTs. Examples are encrypted instant messengers or cloud services as mentioned above. A further example of an anonymization tool is the Tor project[111] which develops software to protect privacy and anonymity online. Tor is well-known among privacy and security experts and recommended by civil rights organizations such as the EFF (electronic frontier foundation). Its core piece is the Tor browser, making use of a dynamic network of servers which avoids direct connections by applying randomly changing, encrypted links. This approach called onion routing allows using private network paths to increase security of data transmission and to improve the degree of anonymity when accessing the Internet. The Tor network provides alternative servers to avoid direct connections between a user's computer and webservers, e.g., visited websites. This enables to use an alternative IP address than the original one when browsing through the web. Consequently, a user cannot be directly tracked via the IP address of her device. However, a user can be still identified and tracked via web browser data. To reduce this risk and protect from online tracking, the Tor browser provides private browsing as integrative feature (e.g., by blocking cookies and other tracking technologies). Besides these practical examples, there are further, yet experimental approaches to implement homomorphic[112] encryption schemes (such as Enigma[113], or Ethereum[114]) by employing blockchain[115] technology in order to protect

[107] https://philzimmermann.com/EN/findpgp/
[108] https://www.philzimmermann.com/EN/essays/WhyIWrotePGP.html
[109] www.openpgp.org , www.gnupg.org
[110] https://www.enigmail.net
[111] https://www.torproject.org/
[112] Homomorphism implies that the structure of an information set is kept. It counts as holy grail of cryptography, see, e.g., Micciancio (2010).
[113] Such as the Enigma project at the MIT https://www.media.mit.edu/projects/enigma/overview/
see also (Greenberg 2015)
[114] https://ethereum.org/
[115] Blockchain is the basis technology for the crypto-currency bitcoin but can also be used for other decentralized systems. (See, e.g., Pass et al. 2016):Analysis of the Blockchain Protocol in Asynchronous Networks: https://eprint.iacr.org/2016/454.pdf

personal information by applying decentralized, cryptographic data structures. The basic idea is to set up decentralized technical architectures to avoid a concentration of power and control by a single party. In this regard it aims at establishing a technical approach to foster checks and balances of a decentralized mode of governance (Zyskind et al. 2015). These and similar concepts are somewhat promising to ease technical privacy problems in the realm of big data and cloud computing.

As shown, there is a broad scope of technical approaches and developments in the realm of privacy by design. Irrespective of their peculiarities and differences, PbD approaches ideally share a common foundation as regards the mechanism of protection: they feature protection mechanisms so that identifiable information cannot easily and seamlessly flow from one application context or information system to another. As outlined in the previous sections, the aspect of a friction in the information flow is crucial in order to come towards effective technical protection mechanisms providing unlinkability. Hence, from a systemic perspective, PbD can be understood as attempt to foster the boundary control function of privacy by providing mechanisms, creating frictions in the information flow which can result in a decoupling of different information systems. Conceptually, cryptography provides the same mechanism and creates such informational frictions because encrypted information is not per se interpretable without decryption. In this regard, encrypted information is equipped with an inherent friction. The mentioned tools exemplify this: for example, Tor creates informational frictions by applying the onion-routing technique, where internet traffic is distributed across a randomized mix of cascading servers. One result is that direct connections between user clients and visited servers are avoided so that users cannot be easily traced by visited websites. A further example is the encrypted online messenger Signal. Besides encryption of communication content, it inter alia blocks the creation of screenshots (cf. Greenberg 2016b). The result is a friction protecting from visual privacy intrusion of a text-based application which could undermine the content encryption feature[116]. Put more generally, this approach reduces the options to gather and represent information processed by the application with alternative forms of representation (here visualisation). Or in other words: this feature creates a friction or boundary between the intended contextual layer (secure messaging) and a not intended further contextual layer (visual information gathering) to improve the security of the tool. This aspect of informational frictions and systemic boundaries between different application contexts is important to improve the conceptual understanding of privacy protection. Another approach to create frictions is sandboxing, as, e.g., employed in the recent version of the Tor browser.[117] Sandboxing is an approach in computer security to decouple different processes from each other so that, e.g., malware cannot easily intrude into a running application (cf. Prevelakis/Spinellis 2001). It is a form of virtualization which allows creating virtual hard- and software environments. Approaches of this kind are promising to achieve higher information security and establish informational frictions.

[116] Because a screenshot can be taken when the sender or the receiver reads a decrypted message.
[117] https://blog.torproject.org/blog/tor-browser-70-released

5.4.2 Can identity management foster privacy by design?

Theoretically, IDM (as outlined in Section 4.2) can contribute to PbD and is discussed as privacy-enhancing approach by some scholars (e.g., Hansen et al. 2004; Jøsang et al. 2007; Leenes et al. 2008; Naumann/Hobgen 2009; Danezis et al. 2014). This is reasonable because the uncontrolled handling of identity information is among the core problems of privacy. In this regard, IDM may contribute to ease this problem by incorporating PbD mechanisms (as outlined in the previous section). Although maximum anonymity and data minimization are the main aim, there are many cases where this is neither feasible nor practicable. For cases that require the processing of identifiable information, IDM can be used to standardize user identification and authentication. This can contribute to improve information security and thus raise the level of protection of an application or system. Depending on the IDM architecture and implementation, the handling of information can contribute to privacy protection. A general requirement for privacy-enhancing IDM is user centricity, i.e., the provision of some user control in handling her identity information. There are four general IDM architecture models: siloed (or isolated), centralized, federated and user-centric (cf. Jøsang et al. 2007; Bhargav-Spantzel et al. 2007; Rundle et al. 2008; Strauß/Aichholzer 2010). In a siloed system, information processing is separated from other systems and therefore, information cannot be easily linked over different domains. This separation provides unlinkability and is basically important for privacy protection (as outlined previously and in Section 3.1.3). However, a siloed approach aggravates information exchange with other systems, which may not satisfy the needs of efficient information processing. In contrast to a siloed approach, a centralized system stores all personal data in one repository (a database) handled by a central entity or provider. Service providers can integrate this central repository into their applications. Here, user authentication can be handled over one centralized account which offers more convenience than e.g., a siloed system. However, users are completely dependent from the central provider and a centralized system entails high risks for privacy and security, because all personal information is centrally accessible. The federated model combines the siloed and centralized approach. Identity information is managed by a central identity provider (IdP) that acts as intermediary between the user and different applications. This federation enables to include a single account for user authentication into multiple applications. The benefit is more convenience and lesser privacy risks than centralization, as an IdP usually provides identifiers but does not process all personal information. However, the IdP has the knowledge about all unique identifiers and usage contexts of a specific person which enables profiling and thus privacy infringement. Therefore, the reliability and trustworthiness of the IdP is a crucial issue in a federated model. The user-centric model tries to reduce the problems of a centralized IdP by offering the user more control. Users here do not have to fully rely on a central IdP but can select one or more different IdPs, and thus have at least some choice whom to trust in processing her identity information. Depending on the concrete implementation, users can manage their (encrypted) identity information with a technical device (e.g., a smart card) that can be used for applications where identification and authentication is needed. This should foster user control and informational self-determination. The flipside of the coin here is that users need high

technical skills to handle their identity information and understand its processing in order to make informed decisions about its usage. Moreover, privacy and security problems remain as the involved providers still have deep insights into one's identity information and thus abuse is possible. Although this approach seems somewhat promising, there are a number of open issues also regarding its technical implementation to provide a solid mix of convenience, usability, security and effective privacy protection (see e.g., Bhargav-Spantzel et al. 2007; Strauß/Aichholzer 2010). Neither federated nor user-centric IDM models provide sufficient mechanisms to safeguard information after it has been shared (among federation members or third parties). In each case, users have only marginal control over their information. The user-centric approach with the option to choose between multiple IdPs may overburden individual users and also bears risks of market concentration (De Andrade et al. 2013). Consequently, information asymmetries and issues of limited user control remain.

In practice, IDM approaches often occur as hybrid forms with several characteristics of these different models. But in general, there is lack of user-centric approaches that effectively improve user control and privacy protection. The implementation of privacy-enhancing IDM implies a combination of privacy protection and authentication (cf. Hansen et al. 2004); to foster the protection and controllability of identity information and thus reinforce informational self-determination. Crucial requirements in this regard concern the implementation of PbD features whereas the incorporation of unlinkability is of particular relevance to provide anonymous and pseudonymous use and appropriate user control (Strauß 2011). However, these requirements for a privacy-enhancing IDM are often not sufficiently implemented in practice. This is observable in public and private sector alike. The insufficiency of privacy protection of social media based IDM is relatively obvious as users factually have no effective option to protect their privacy (see next Section). But even in the government sector, where enormous efforts have been made to implement IDM systems for citizen identification there is a broad scope of different implementations with varying protection levels regarding privacy and security (cf. Naumann/Hobgen 2009; Kubicek/Noack 2010b; Strauß/Aichholzer 2010; Strauß 2011). A technical study on governmental eID systems in Europe pointed out that none of these approaches provided sufficient privacy safeguards (Naumann/Hobgen 2009). The study found that some systems even used unencrypted global identifiers (such as a social security number). But also more sophisticated approaches were criticised which serve as sort of best practices in e-government in Europe; such as the Austrian and the German eIDMS. Each of these systems applies unlinkability concepts. The Austrian eIDMS represents a sophisticated approach with a mix of federated and user-centric elements. It uses so-called sector-specific identifiers to provide unique identification in specific contexts or sectors (e.g., health or tax) but avoids that personal information, referring to the eID holder, can be easily linked across different sectors. This solution provides some protection and higher levels of security compared to other European solutions. However, the Austrian system was criticized for its high complexity, lack of usability and insufficient privacy protection, as serious privacy problems remain. One aspect concerns the sector-specific identification which in principle contributes to reduce the risk of linkability and profiling. But as one of the different sector-specific identifiers is involved in almost every usage context, this

undermines the unlinkability concept as personal information can then be aggregated via this identifier (Strauß/Aichholzer 2010; Strauß 2011). Also more recent studies found several privacy issues in European eID systems (cf. Sapelova/Jerman-Blažič 2014). Furthermore, besides design flaws as regards protection of personal information, there are other possibilities to link information across different contexts by using "semi-identifying" information which refers to the identity shadow problem (as discussed in Section 5.3.1). Hence, though it is a crucial requirement, unlinkability is ineffective when it is limited to reducing linkability of identifiers while privacy infringement via other types of information is not sufficiently prevented from.

A further critical issue concerns the notion of a core or universal identity incorporated in many IDM approaches aiming at facilitating identification based on quasi-centralisation. Accordingly, IDM is then often seen as means to provide "one identity per individual" (e.g. Waters 2004). This perception is common in technology development, where identity is seen as generalizable, abstract representation of an entity respectively person (e.g. Glässer/Vajihollahi 2010). The basic idea is that the "core identity" of a person can be simply represented by a specific set of identifiers used in multiple application contexts. This core identity is then represented by a digital artefact, a carrier device, an identity token etc. (see Section 3.1). The general aims of this quasi-universal digital identity are to standardise identification and improve security and efficiency of information processing. In general, it is a plausible intention to standardise identification procedures. However, the framing of a universal or core identity is reductionist and neglects the fact that identity is a relational and dynamic concept. This is in conflict with an (often neglected) aspect: the fact that there is no single universal identity. Because identity is context-sensitive, and thus every individual can have multiple (partial) identities in different contexts (see Section 3.1). IDM Approaches neglecting this aspect result in revealing real identities without options of anonymous or pseudonymous usage. Pfitzmann and Hansen (2010) highlight this aspect and define privacy-enhancing IDM as "managing various partial identities (usually denoted by pseudonyms) of an individual person, i.e., administration of identity attributes including the development and choice of the partial identity and pseudonym to be (re-)used in a specific context or role." This notion of IDM is essential to allow for unlinkablity. A lack of understanding and consideration of this issue aggravates effective privacy protection. An individual then has no real choice about which information to provide and no other option than revealing her (real) identity, which further reduces anonymity.

A core identity approach may be appropriate for some services (e.g., legally binding e-government and e-commerce transactions), where formal or hard identification is required and legally restricted. However, it is problematic for all kinds of services where identification is inappropriate or itself problematic from a privacy perspective (such as for information services). For example, unique identification and the disclosure of information such as full name, address, social security number etc. may be appropriate for a particular e-government service where a transaction is involved. But it is inappropriate for a common online application (e.g., social media, browsing or searching the web etc.). Identification is often employed for plausible reasons, but there are several tendencies to expand (digital) identification mechanisms, observable in various contexts (as shown in Sections 4, 5.1,

5.3). Trends to foster digital identification entail risks of function and mission creep, here meaning the incremental extension of identification resulting in a quasi-obligation for individuals to reveal their identities even in contexts where identification is not needed (cf. Lyon 2009). Examples include intentions to combine different identification systems such as governmental IDMS and social media, as mentioned in Section 4.3.3. The quasi-obligation to ID is partially conveyed by a "nothing to hide" logic (as discussed in Sections 3.3.2 and 5.2.2). Even though identification is necessary and important in many societal contexts, it is neither adequate nor necessary in every context. Moreover, anonymity used to be the standard case in most contexts of everyday life although there may be opposed trends. As shown in the previous sections (4.2.2, 5.1.2) IDM implementations are mainly driven by a complex mix of economic and security objectives. This securitization and economization of digital identification hampers the implementation and use of privacy-friendly IDM systems. In line with the principle of data minimization, identification and identifiability have to be avoided as far as possible in order to reduce privacy risks. Consequently, those contexts, where identification is needed (e.g., in formal transactions), the processing of identity information requires contextual integrity in line with the principle of purpose binding. Hence, privacy-friendly IDM approaches require integrated features enabling anonymity and pseudonymity. Otherwise, IDM bears certain risks to undermine privacy protection by reinforcing identifiability.

5.4.3 The limits of user control

Hence, IDM has several limits as regards PbD. Especially concepts following a core identity approach with a narrow focus on economic and security issues undermine privacy protection. Indeed, IDM with integrated PbD features can contribute to foster control by facilitating identification procedures and the handling of identity information. As outlined, a crucial part in this regard is to foster user centricity and enhance control. However, although this is by all means important, it is questionable whether it is sufficient to effectively improve privacy protection. The proper technical implementation of PbD mechanisms can be very challenging for developers. Technical safeguards yet often lack in protecting from unintended, privacy infringing identification and identity shadow exploitation. Some may perceive that better technical safeguards and more sophisticated IDM concepts such as in the governmental domain could ease the problem. However, as shown, IDM entails a number of privacy risks and a (quasi-)centralized IDM layer would reinforce the risk that IDM is abused as a surveillance infrastructure (Rundle et al. 2008; Lyon 2009; Strauß 2011). Furthermore, even if an effective privacy-enhancing IDM system would be available, its scope can be seriously limited when its proper usage mostly lies in the responsibility of the individual user alone. Certainly, fostering individuals in the self-controlled handling of their information is basically crucial. But it is insufficient to merely equip technology with some user control mechanisms, which reduces PbD to a sheer technical concept. One aspect concerns so-called security usability issues: when a system is too complex for the user and lacks in usability, this may undermine its protection concept. Several studies argue that users rather avoid IDM tools that are perceived as complex and inconvenient (cf. Jøsang et al. 2007; Bhargav-Spantzel et al. 2007;

Strauß/Aichholzer 2010). For instance, a comparative study on governmental IDM systems in Europe found that high system complexity is a great burden for citizens which results in rather low usage rates (Kubicek/Noack 2010b). Furthermore, the systems may improve the security of service providers but not for the individual users (ibid). Hence, insufficient usability of IDM may undermine information security and privacy of individual users. Especially, when they primary focus lies on the interests of the service providers. Consequently, individual users have limited options to protect their privacy. Besides the technical challenges, crucial problems result from insufficient or ineffective privacy protection among institutional actors processing personal information. This refers to issues regarding information asymmetries and agency problems (as discussed in the previous sections). These issues are inter alia observable in privacy settings and privacy policies which underline the limits of user control. Privacy settings basically enable users in customising the processing of their personal information in technologies and applications. In this respect, they may be seen as simple, practical PbD examples. However, in most cases, these settings are very limited in scope and often lack in effective privacy protection for several reasons. A prominent example is Facebook, where users can customize their profiles and define e.g., which information is available and which is private. While this is useful in general, it does not change the fact that Facebook itself as well as its many external partners exploit users' information and infringe upon their privacy. Figure 21 below shows a comparison of how Facebook's default privacy settings between 2005 and 2010:

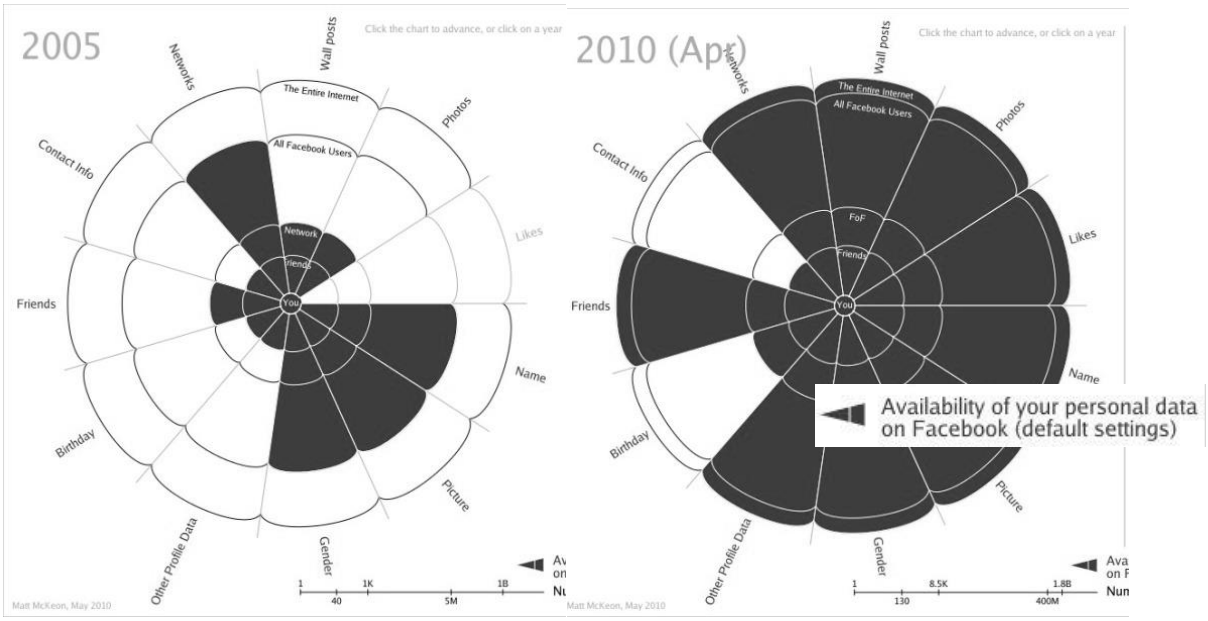


Figure 21: Facebook's privacy setting over time
Source: adapted from Matt McKeon 2010: http://mattmckeon.com/facebook-privacy

The aim of privacy settings is to enable users in controlling the disclosure of their personal information. However, already the standard setting undermines this as the visualization

demonstrates.[118] In its beginnings, Facebook at least provided a basic level of protection. With frequent updates of its privacy policy this has drastically changed. Access to personal information was limited in 2005 and mostly exclusive for direct contacts or members of the network (as the illustration shows with the inner circles). However, No such limits exist since 2010 anymore: since then, also entities outside the social media environment can access several parts of user information by default. Users who do not proactively change their privacy settings automatically disclose their profile information including contacts, photos, preferences etc. Besides the high complexity of the settings users are confronted with, the option to reduce profile visibility in general is rather "a quick fix (…) than a systematic approach to protecting privacy" (Debatin/Lovejoy 2009: 103). Since 2010, the situation did not improve from a privacy perspective as shown in the previous sections. Referring to another policy change, in 2013, the New York Times commented that, "Facebook's new policies make clear that users are required to grant the company wide permission to use their personal information in advertising as a condition of using the service" (Goel/Wyatt 2013). In fact, even privacy-aware users have little chances to prevent their profile information from being exploited by third parties such as marketers, security agencies or other entities of the surveillant assemblage (see Section 5.1). Facebook is only a prominent example of many for the numerous problems individuals are confronted with when attempting to protect their privacy online.

A core problem of insufficient privacy controls results from socio-technical identifiability as demonstrated in the previous sections with the identity shadow. The identifiability by default setting inherent to many ICTs offers various ways to gather identity information. Accordingly, privacy policies make use of the insufficiency of privacy settings and reinforce identifiability. In theory, the primary aim of a privacy policy is to inform users about how their data is treated and how their privacy is respected. This relates to the privacy principle of transparency to enable users in comprehending the processing of their information (see Section 3.2.1). However, in practice, this is mostly insufficient as privacy policies are often difficult to understand and rather provide general descriptions of the variety of ways, a maximum of personal information is collected but not how privacy is protected. Service or application providers usually justify their extensive collections of user information with user experience and high quality of services. This is highlighted by major players in the web like Google, Facebook and others, who gather large arrays of user information from their services as well as from all kinds of web content where their services are involved. An excerpt from Google's privacy policy[119], for instance, indicates the enormous magnitude of information collections:

"We collect information to provide better services to all of our users – from figuring out basic stuff like which language you speak, to more complex things like which ads you'll find most useful, the people who matter most to you online, or which YouTube videos you might like."

Information is collected in various ways, beginning with all information provided by users such as name, email address, phone no., credit card no., all information of user

[118] This visualization is only an excerpt and does not show the full scope of disclosed personal information.
[119] Google privacy policy: https://www.google.at/intl/en/policies/privacy/

profiles including name or nick, photos etc.; information from service usage including what is used, how, when etc., device-specific information including hardware model, software versions, unique identifiers, information from mobile usage such as phone no., device no. etc. which can be associated with a user account; Also "usage data and preferences, Gmail messages, G+ profile, photos, videos, browsing history, map searches, docs, or other Google-hosted content" is gathered and analysed by automated systems. Similar is the case for search queries or online activities such as searching a restaurant on Google Maps or watching a Youtube video; information about the activity or content (search term, video etc.) is processed as well as user location, device details etc. Furthermore, amongst others, besides cookies and log information also location information and unique application numbers are collected as well as tracking technologies (e.g., pixel tags) are used and may be stored on a users' computer or similar device. These technologies also enable partners of Google (such as online marketers) to process user information. Hence, put simply, individual identity shadows are exploited as much as possible. Similar information collection practices can be found in many other privacy policies as well. Users are confronted with complicated statements while at their core, privacy policies frequently declare that in many cases, maximum information is collected.

But how is this possible? The foundation for these extensive information collections lies in the concept of informed consent which is a basic part of privacy and data protection laws (as mentioned in Section 3.3). In most countries, collecting and processing of personal information for commercial self-purposes is legally prohibited and only allowed if the concerned individual agrees. Hence, service providers and other information processing entities have to ask their potential users for permission which is obtained by the informed consent. This idea of informed consent is basically crucial. However, the crux is that often, user acceptance is rather constructed to legalise personal information processing, e.g., for commercial purposes. But individual users mostly have no other choice than to fully accept the privacy conditions, and thus the processing of their information irrespective of whether third parties are involved or not. In case of dissent, the individual cannot use the service, application or technology. Options to partially agree are usually not foreseen such as to allow a service used directly to process information but prohibit third party access. Therefore, due to the lack of choice, informed consent often equals a "like it or lump it" approach as criticized by several privacy scholars (cf. Nissenbaum 2010; De Andrade et al. 2013; EGE 2014; Danezis et al. 2014; Zuiderveen-Borgeswius 2016). All in all, informed consent thus often entails an enforced agreement to accept privacy intrusion with an inherent quasi-obligation to explicit or implicit identification.

5.4.3.1 Privacy divides

Against this background, privacy settings as well as more sophisticated privacy control mechanisms are per se limited in scope. Further issues concern additional privacy tools (PETs). As outlined in Section 5.4.1, there is a number of PETs available to support individuals in protecting their privacy. However, a general issue is that in many cases, these tools are rather expert-centred and have a low familiarity among non-experts. Moreover, although usability increased, usage often requires high privacy awareness,

technical knowledge and entails several efforts. Classical problems of email encryption tools like PGP, for instance, include high complexity for standard users, lack of convenience, awareness, no "real" use cases and thus low usage rates, still exist today (Ruoti et al. 2016). Thus in fact, many privacy tools are still rather limited to a small group of experienced users. Furthermore, also skilled users may encounter several difficulties to achieve protection as the case of the anonymization software Tor highlights. Tor offers a relatively high level of protection but also has certain limits. One aspect concerns usability issues: basically, the Tor browser is easy to use, widely similar to a standard web browser. One of its peculiarities are default features restricting certain tracking features of websites (e.g., cookies, web bugs, embedded JavaScript, iframes etc.) to enhance privacy protection. However, an unintended side effect is that this can impair the functionality of websites. Though this is not the result of Tor design or usage but of insufficient web design, it can complicate individual usability. Another aspect is that protection ends at a user's "final" destination. Several security researchers demonstrated certain risks of de-anonymization of Tor users. A major issue is the problem that exit nodes of the Tor network can be abused to spy on individual users (cf. Le Blond et al. 2011; Sun et al. 2015). Nevertheless, Tor is a solid tool that offers some protection from profiling. But user data can still be gathered from the web content a user interacts with. For instance, Tor provides little options to prevent social media from exploiting or tracking its personal user profiles. As soon as e.g., a user accesses her profile at an online platform with her user credentials (e.g., username, password), she becomes identifiable. Thus even when using sophisticated privacy tools like Tor, situations can occur where users are confronted with the option to either accept privacy intrusion or avoid the use of platforms requiring identification. A further option is to use alternative platforms in addition to Tor such as the decentralized, privacy-friendly SNS Diaspora or Friendica[120], PGP-based e-mail encryption and an encrypted Cloud service (e.g., Spideroak). A combined approach provides additional protection but also increases the efforts of individual users. Moreover, privacy-friendly alternative services in general are rather niche applications with a very low number of users. For instance, compared to Facebook's about two billion users (see Section 4.3); Diaspora had less than 400,000 users in 2014[121]. This highlights the crux of PETs: low usage rates complicate interaction with persons that do not use these services (e.g., sending an encrypted e-mail is ineffective if the receiver cannot decrypt it; the circle of friends active on Facebook is significantly higher than in Diaspora). Consequently, the lack of a critical mass of privacy-aware and skilled users also limits the scope of privacy-friendly approaches and hampers unfolding of their potential to strengthen privacy protection.

Their limited outreach is a general problem of privacy tools. Among others, Papcharissi and Gibson (2011) detected a particular form of digital divide, a privacy divide between privacy haves and privacy have nots in the context of social network sites. Members of the former group are privacy-aware, have technical skills and are thus cautious in providing their personal information. Members of the latter though, lack in privacy awareness as well as in computer literacy and thus carelessly disclose personal

[120] Diaspora: https://diasporafoundation.org/ Friendica: http://friendi.ca/
[121] How many users are in the DIASPORA network? https://diasp.eu/stats.html

information. Moreover, with the predominating commercial processing of personal information "privacy gradually attains the characteristics of a luxury commodity (…)" as "it becomes a good inaccessible to most (…) is disproportionately costly to the average individual's ability to acquire and retain it" (Papcharissi/Gibson 2011: 85). Furthermore, a lack of privacy protection can create a social disadvantage. This privacy divide is not limited to social media but is generally observable between privacy-aware and skilled users; i.e., those with the knowledge and ability to e.g., use privacy tools and average users with a lack thereof. Moreover, another aspect is that privacy-aware individuals may experience social disadvantages precisely *because* they use PETs. There are tendencies among several providers to prevent users from employing privacy tools and creating user accounts without the provision of their real ID's. For instance, the registration of a free e-mail account without the need to provide valid identity information became rather difficult. A number of free mail providers request real IDs or the provision of a phone number to double check the validity of a user account (as illustrated on the example of Gmail in Figure 22 below).[122]

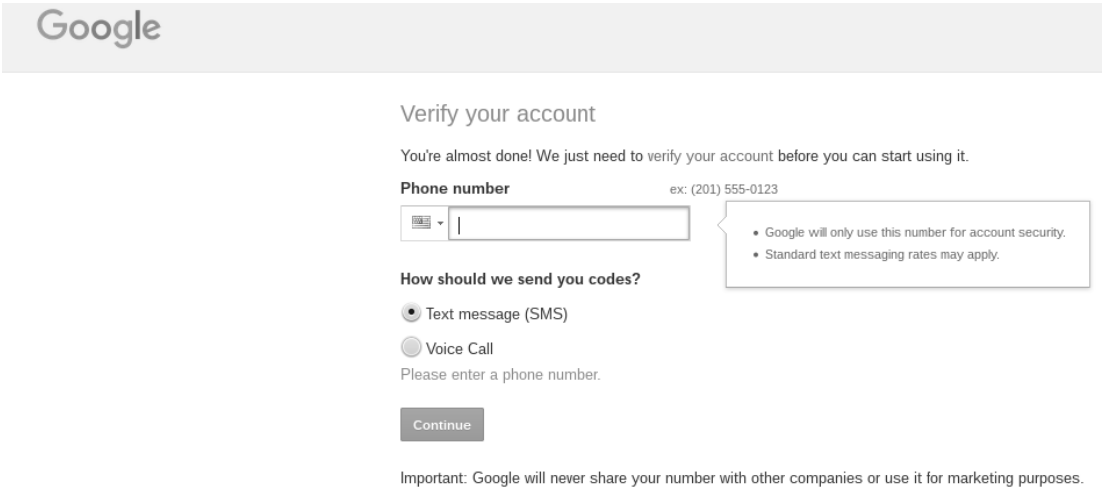


Figure 22: Example of e-mail registration phone number request

From the providers' perspective, this measure is partially plausible as long it aims at protecting from abuse, spam etc. and increase service security. However, from a privacy perspective, it is critical when the identity information processed by these measures is stored and used for additional purposes such commercial profiling etc., entailing further privacy intrusion. Basically, this information enriches the identity profiles a provider has about its users. Although some providers (e.g. Gmail) claim that the phone number is not shared for marketing purposes (as shown in the figure above), third party use is not per se excluded and may be a matter of purpose definition. Irrespective of their plausibility, measures like these also contribute to further constraints of online anonymity. That some providers perceive the use of privacy tools as burden to their business models is e.g., observable in the case of Facebook frequently trying to prevent the use of ad-blockers (e.g.

[122] As simple practical tests (conducted on September 22 2016) show. Among the tested provider where Gmail (google.com/gmail), Gmx.net, Web.de, and Outlook.com.

Tynan 2016). Also the real name policies of internet companies are not just a security measure but tied to economic interests as more detailed user profiles are more valuable for marketers. Privacy-aware persons who want to protect their personal information by e.g., avoid revealing more information than necessary for service usage (e.g., their phone number or other additional identity information) in order to avoid tracking, spam and to protect their privacy are excluded from usage.

Furthermore, some providers prevent users from accessing their services via anonymous networks such as the previously mentioned Tor network. According to study, the expression of political rights and prevention from political repression are strong drivers of Tor usage in repressive regimes and liberal democracies alike (Jardine 2016). It is evident that the use of privacy tools like Tor or the crypto-messenger Signal can be difficult in countries with problematic issues regarding human rights protection, where online activities are censored or constrained. For example, China, Saudi Arabia, Bahrain, Russia, Iran, Egypt, Turkey attempt to restrict or block these privacy tools (cf. BBC 2016; Osborne 2016; McCarthy 2017b). These blocking activities are particularly problematic in countries where regime critics, journalists, human rights activists etc. may encounter repression and thus have high demand for PETs and anonymization tools. However, also in other countries, users may encounter difficulties when using PETs. For instance, some IP addresses from the Tor network are occasionally blocked by some ISPs as well as online providers.[123] Some security agencies such as the FBI or the NSA claim that Tor and other anonymization networks are mostly used by criminals and terrorists (Ball et al. 2013; Froomkin 2015). This may be one reason why Tor users are classified as "extremists" by the NSA surveillance programs (as outlined in Section 5.1.3). While Tor is used for criminal activities as well (McGoogan 2016), this does not change the fact that anonymization services are essential tools for non-criminal, privacy-aware individuals, human rights activists, journalists, security researchers and other sensitive professions alike. In this regard, there is a certain "Dark Web dilemma", which may be resolved with cautious policing activities in the Dark Web to stem criminal activity (Jardine 2015). Surveillance of Tor and other privacy tools is highly critical from a human rights' perspective. Hence, more promising are cooperative approaches, e.g., between Tor developers and law enforcement to improve mutual understanding of the motives from privacy-aware users and law enforcement alike. For instance, developers of the Tor project cooperate with law enforcement to e.g., offer trainings to better comprehend the network.[124] This and similar measures can support law enforcement to target criminal activity without criminalizing the whole network and its users. A criminalization of Tor or other privacy tools as well as efforts to restrict them would harm human rights and (civil) society in many respects.

---

[123] For a list of services blocking tor see, e.g.,
https://trac.torproject.org/projects/tor/wiki/org/doc/ListOfServicesBlockingTor
Reasons for being blocked can be difficult to verify. For instance, during some practical tests of the Tor browser (with different IP addresses involved), conducted on September 24 2016, access to some websites via Tor such as Google search, Gmx.net, Facebook.com login, or Amazon.com was hampered and partially blocked. These problems did not occur in some further tests (September 30 2016), but again in some others. Reasons can be manifold but this further indicates problems privacy-aware users may encounter when using PETs.

[124] Trip report: Tor trainings for the Dutch and Belgian police. Tor Blog, February 05 2013, https://blog.torproject.org/blog/trip-report-tor-trainings-dutch-and-belgian-police ; see also (Dreged 2013)

5.4.3.2 The "privatisation" of privacy

The presented cases including tendencies to limit online anonymity and block privacy tools underline that there are many serious barriers for individuals to safeguard their privacy. A particular problem is the reduction of PbD to a sheer means of individual privacy control, while the responsibility of information processing entities (e.g., public and private institutions) to protect privacy is neglected. This relates to an exaggerated focus on technical issues of user centricity and control. Besides technical issues, several conflicting interests complicate effective protection. From a wider perspective, also here, agency problems between individuals and institutional entities become apparent. Hence, the uncontrolled processing of identity information is obviously not merely an issue of the individual but rather an institutional problem. The general increase in identification practices (as presented and discussed in the previous sections) results from a complex interplay of technological design and socio-political factors whereas securitisation and economisation of identification are among the core drivers of this development. The expansion of identification mechanisms is problematic because it entails an incremental increase in (quasi-)mandatory identification and thus a reduction in anonymity. Security agencies and law enforcement tend to see privacy tools providing anonymity as barriers to their work, and thus strengthen efforts to limit their use. Providers of online services aim at protecting their services from abuse as well as have additional economic interests to foster identification and identifiability. Every interest as well as their entailed activities is plausible to some extent. However, privacy-affecting activities are often opaque, lack in transparency and accountability. Therefore, it can be difficult to assess their factual plausibility as well as their legal and ethical justifications. Furthermore, there is a certain risk of what can be called "privatisation of privacy", i.e., where the responsibility to protect privacy is mainly shifted towards the individual concerned. As shown, there are several limits to avoid unwanted information processing and reduce identifiability. Basically, it is hardly possible for a single individual to manage her identity information in all the manifold contexts it is being processed.

As outlined in section 3, privacy is an individual human right but also a common good, i.e., a societal value of public interest, vital for democratic societies. This public or common value is at risk if privacy is reduced to a private value that is of individual concern only. As e.g., Kahn (2003: 388) pointed out, individual control over privacy cannot be fully detached from social norms because such a separation would result in a fragmented society. While the aspect of individual control over privacy is basically essential, this does not necessarily imply that an individual alone carries the full responsibility that her identity information is not misused. To the contrary, social reality and practices of information processing make clear that this is neither applicable nor desirable. The previously outlined privacy divides indicate a certain fragmentation in digital environments. However, the main problem is not a gap between privacy haves and -have nots but rather a gap as regards the responsibility to protect privacy between individuals and institutions. A sheer focus on user centricity and user control is not sufficient to ease this situation; especially not, when privacy is falsely reduced to a sheer private issue. Reasons for the reinforced tendency to "*privatise*" privacy can be derived

from the traditional notion of privacy as right to be let alone (Warren/Brandeis 1890) which emerged from ownership and property rights. This notion still plays an important role today. Warren and Brandeis (1890) inter alia argued that information about a person such as an image belongs to this person, respectively is in her property. They argued for privacy to be a stand-alone right disentangled from property rights (Kahn 2003). However, this classical framing of privacy lacks in considering it as a multidimensional concept and particularly its function as a public value (Nissenbaum 2010; Cohen 2012). Hence, privacy is not only a sort of "*materialistic*" good and matter of personal concern. However, if privacy is reduced in this regard, this may result in a paradox situation: individuals as original holders of their information can enable others to use their "property" by giving informed consent which often entails a loss of control over their information. At the same time, they are overburdened with the responsibility to control their information to protect from misuse.[125] This situation is inter alia observable in the dynamics of the digital economy where the free flow of (personal) information is framed as a core value which feeds into commercial interests. In line with this logic, the gathering and processing of personal information is a (semi-automated) standard procedure of most ICTs and online services. In this framing, privacy is rather perceived as a burden. Consequently, the responsibility to protect privacy is forwarded to the individual user. At the same time, there is a lack of transparency of information processing and accountability of the involved entities who neglect privacy responsibility.

Thus from an individuals' perspective, the privatization of privacy complicates effective protection of personal information and (digital) identity in many respects. A shown, there are several promising approaches in the realm of PbD. As economic interests are among the core drivers for privacy infringement, PbD bears potential to counterweigh this situation by creating and stimulating economic incentives to protect privacy. However, there is also a certain risk of a further privatisation of privacy. User-centric PbD approaches and PETs are no panacea even for experienced users. A general burden results from the strong focus on technology and the often high complexity of technical solutions. Moreover, the problem is less a technical but more a societal one. While average users per se have little chance to remain anonymous in online environments, experienced, privacy-aware users have to face eventual discrimination because of their privacy affinity. Indeed, individuals have the main responsibility to protect their privacy. However, against the background of imbalanced control over identity information, increasing identification, information asymmetries between individuals and institutions etc. there is very limited room for manoeuver in this regard. Therefore, to compensate insufficient privacy protection requires combined approaches to revitalize the public value of privacy. Crucial in this regard is shared responsibility between individuals and institutions, with particular focus on the latter. As regards technical approaches, this implies a need to improve PbD beyond some occasional control features. PbD is not limited to the provision of user control mechanisms but ideally is an approach to implement privacy safeguards with the

[125] This is a little bit as if a house holder entrusts a house keeper to take care of his home who then throws out the house holder and blames him for not taking care of his home.

main aim to minimize privacy intrusion by information processing institutions. Hence, a crucial requirement for effective PbD approaches is data minimization (cf. Gürses et al. (2011). More specifically this implies a reduction of identifiability wherever possible. But technical solutions alone are not sufficient to ease the problem. Rather there is a demand for socio-technical governance including organizational, regulatory and technical measures to compensate information asymmetries resulting from identifiability. While user control is without doubt essential too, it is not sufficient to merely shift the responsibility to protect privacy to individuals. Therefore, there is a need to foster the accountability and responsibility of information processing entities as regards the protection of privacy.

The crux is that the implementation of PbD can be challenging: engineers and developers need particular guidelines to design privacy-friendly technologies and applications; institutions taking their privacy responsibility serious need detailed knowledge about the functioning of the information processes in their organisations including the employed technologies and applications. Although PbD principles and guidelines are generally useful, several scholars criticized them for being too vague and thus difficult to consider for engineers (cf. Gürses et al. 2011). One issue is that PbD requires privacy awareness at organizational as well as individual level. Existing PbD guidelines are important to raise this awareness and stimulate the creation of a privacy-friendly culture which is a precondition for the implementation of PbD. However, these guidelines rather address general management issues than more concrete organizational and technical requirements. The factual implementation needs to consider the peculiarities of a technology or application as well as of the organization applying them. Hence, the crux is to identify the demand and requirements for PbD of a particular institution. Privacy impact assessment (PIA) is an instrument to identify these requirements which can support and stimulate the development of PbD. In this regard, it is a prerequisite for the informed implementation of PbD. In the long run, both concepts – PIA and PbD – can mutually reinforce each other and contribute to create a higher level of privacy protection standards. Therefore, a combined approach between both is essential. This also corresponds with forthcoming EU privacy regulation, where the complementarity of PIA and PbD is implicitly included. Article 25 of the GDPR regulates PbD (labeled data protection by design and by default): data controllers (entities who define the purpose of processing) are encouraged to "implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects." Amongst others, the implementation of these measures should take into account the "nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing" (Art. 25 GDPR). Hence, the implementation of PbD requires proper knowledge of the scope and impacts of data processing which refers to PIA as regulated in Article 35 GDPR, encouraging to "carry out an assessment of the impact of the envisaged processing operations on the protection of personal data" in advance to the processing. Therefore, PIA plays an important role to tackle many of the explored contemporary privacy problems, which is explored more in-depth in the following sections.

CHAPTER 6

Assessing and refining privacy impact assessment

The previous sections dealt with the emergence of digital identification, the complex mechanisms it is based upon as well as socio-technically amplified identifiability of individuals being a core problem of contemporary privacy protection. As shown, there is a privacy control dilemma of digital identification aggravating with increasing identifiability. This dilemma is accompanied by information asymmetries at the cost of individual privacy and informational self-determination. Hence, the boundary control function of privacy is significantly hampered by uncontrolled flows of identifiable information and diminishing informational boundaries between different socio-technical systems. Fostering privacy-enhancing technologies and PbD approaches is by all means important. As shown, there are some promising approaches to regain control over identity information, and thus over identifiability. However, there are also serious limitations for effective privacy protection. A central technical issue concerns the expanding identity shadow offering various ways of implicit identification which are often insufficiently addressed by PbD. Given the enormous complexity of the problem, including the entailed conflicting societal interests and practices, the effectiveness of technical solutions is seriously impaired. Particularly because of insufficient knowledge about the flows of identifiable information and control thereof, as information can be easily de- and re-contextualized. To improve privacy protection requires a combination of regulatory and socio-technical measures. Besides technical barriers of complex solutions, a main problem results from trends toward a further "privatisation" of privacy entailing a shift in responsibility to the individuals concerned. At the same time, agency problems remain and aggravate. In particular when there is a lack of transparency and accountability among the entities processing and controlling identifiable information. As a consequence, the effective implementation of essential privacy requirements such as unlinkability is impeded. Therefore, enhancing transparency and accountability is crucial to ease the privacy control dilemma by compensating, or at least reducing information asymmetries resulting from uncontrolled identifiability. From a wider, systemic perspective, this implies to make informational boundaries of socio-technical systems more transparent and verifiable. In this regard, privacy impact assessment (PIA) plays an essential role to improve the understanding of identity information processing conducted by institutions, and thus to raise their transparency and accountability. This is important for the development of more effective PbD concepts. PIA and PbD are thus complementary or on the same side of the coin.

The initial aim of PIA is to examine the privacy impacts of socio-technical practices and check whether they are in accordance with privacy and data protection regulation. This does not necessarily imply detection of privacy violations but of potential privacy risks thereof. However, technological progress challenges PIA in many respects as traditional approaches have a limited capability to grasp the privacy impacts of ICTs. Basically, PIA

used to be an instrument in the toolbox of data protection authorities (DPAs) and other privacy protecting institutions. However, these institutions often have very limited resources and capacities to effectively enforce privacy and set counteractions in case of privacy abuse or infringement (cf. FRA 2010; Strauß 2015b; Wright/Kreissl 2015). Furthermore, they need a legal mandate to become active and until recently, there was no legal obligation at all in Europe for companies to conduct PIA or proof privacy compliance. The problem of limited effectiveness of DPAs and other oversight bodies to enforce privacy rights is an issue at the EU level. The European data protection supervisor proposed the creation of a digital clearinghouse to stimulate the coherent enforcement of digital rights in the EU by fostering cooperation between national DPAs and other oversight bodies. The primary aim is to improve the protection of privacy in the age of big data. Among the planned activities of this institution is the assessment of impacts on digital rights and particularly on privacy (EDPS 2016b). This proposal is one among several indicators about PIA is about to gain more importance. Until recently, it played a rather marginal role particularly in the European context. This is about to change especially with the yet ongoing European privacy reform which, among other things, aims at institutionalizing PIA in the privacy and data protection regime. The recently enacted GDPR inter alia foresees a mandatory conduction of data protection impact assessments (DPIA) for public and private institutions under certain conditions, when specific risks occur as regards individual rights (Article 35 GDPR). As a consequence, particularly businesses in Europe have to deal more thoroughly with their information systems and their modalities of personal information processing. Against this background, an increasing relevance of PIA becomes apparent including the need for approaches facilitating the implementation of a PIA process.

As the requirements for PIA may vary in different institutions, it is difficult to develop a standard procedure applicable in multiple domains. However, irrespective of particular institutional settings, as shown, the emergence of a privacy impact is essentially affected by identifiability and the processing of identifiable information. Taking this aspect into consideration is thus basically important for every PIA process. An enhancement of PIA, with accordingly improved options to detect privacy risks and demand for protection can be supportive to develop more effective protection measures. More precisely, a perspective on privacy protection understood as protection of an information process, i.e., a flow of identifiable information could contribute to gain a more systematic framework of protection. In accordance with the boundary control function of privacy, such a systemic view on privacy enables to locate systemic boundaries. This allows better grasping the extent to which privacy intrusion might occur and thus can contribute to develop approaches for a decoupling of different but interwoven systems vital for more effective privacy protection. The basic premise here is that fostering the conceptual understanding of (socio-technical) identifiability contributes to enhance the assessment of privacy impacts. As a consequence, (more) appropriate strategies and socio-technical mechanisms can be developed to reduce certain privacy risks, comprehend and avoid unnecessary or unintended privacy intrusions.

The following sections present and discuss main aspects of privacy impact assessment, arguing for more emphasis on issues of identifiability understood as main determinant of

information processes with a privacy-intrusive capacity. The next section begins with a brief review on the role of privacy impact assessment and existing concepts in this regard. As will be shown, existing PIA frameworks have certain limits and there is a need for more systematic approaches in order to deal with the contemporary privacy challenges. Hence, based on these existing approaches, a novel typology of identifiability is suggested to come towards enhanced forms of PIA. The main part of this section (6.3) explores different types of identifiable information and proposes an identifiability-based PIA framework.

6.1 Overview on the functions and scope of PIA approaches

According to Clarke (2009: 123), PIA "is a systematic process for evaluating the potential effects on privacy of a project, initiative or proposed system or scheme". Wright and De Hert (2012a: 5) define PIA broadly as "a methodology for assessing the impacts on privacy of a project, policy, programme, service, product or other initiative which involves the processing of personal information and, in consultation with stakeholders, for taking remedial actions as necessary in order to avoid or minimize negative impacts". Ideally, this process is not conducted ex post but ex ante, and is already involved in the development of a technology, application etc. so that privacy and data protection issues can be detected early. PIA is thus a means to identify and assess potential effects and risks of information processing on privacy which contributes to reduce corresponding vulnerabilities. This latter aspect refers to the development and implementation of PbD. For organizations, conducting a PIA can serve multiple aims: a primary aim is to verify whether information processing of the organization is in compliance with privacy and data protection law. This verification can improve the trustworthiness of the organization in the public and the enhanced knowledge of its information processing can contribute to reduce costs and increase efficiency (Strauß 2017b). In total, PIA can help to foster transparency and accountability of an organization, serving the organisation, the individuals concerned of information processing, as well the wider public.

Since several years, PIA gains in importance on a global level. Interestingly, in countries like Canada, Australia, and New Zealand, PIA has some tradition. In Australia, the first approaches to assess privacy impacts in the public sector date back to the 1990s (Wright/De Hert 2012b; Clarke 2012b; Bayley/Bennett 2012). Also Canada started early and was among the first countries to consider PIA as instrument of privacy policy which is legally regulated and foreseen as mandatory for government agencies under certain conditions since 2002 (Bayley/Bennett 2012). In the US, the Internal Revenue Service (IRS) issued a PIA model in 1996 and today also has some legal obligation for government agencies to conduct PIA (Wright/De Hert 2012b). In Europe, PIA as such has a relatively short history so far. This does not imply that the assessment of privacy impacts is less important in Europe. But in contrast to the Anglo-Saxon countries, there were no guidelines or conceptual models available before 2007, when the first PIA handbook was published in the UK. A reason for this is that, until recently, the term PIA[126] was not

[126] Or its more common synonym in Europe DPIA – Data Protection Impact Assessment

explicitly used in legal frameworks. However, a related concept is prior checking which used to be more familiar in Europe. Article 20 of the EU Data Protection Directive 95/46/EC (DPD 1995) addresses prior checking and asks member states to determine processing operations that are likely to entail privacy risks and to examine these processing operations before their implementation. The task of prior checking is in the main responsibility of national DPAs who check the legality of applications reported by the data processing entities (e.g., companies, organizations etc.). In practice, as with PIA, this procedure varies significantly from country to country (Clarke 2009; Le Grand/Barrau 2012).

In 2009, PIA gained more attention in Europe as the EU Commission asked the Member States for input to the development of a PIA framework for the deployment of RFID technology. The Article 29 Working Party[127] endorsed the RFID PIA framework in 2011. The EU Commission also announced plans to include PIA approaches in new legal frameworks on data protection (Wright/De Hert 2012b). In specific cases, conducting a PIA should be obligatory, "for instance, when sensitive data are being processed, or when the type of processing otherwise involves specific risks, in particular, when using specific technologies, mechanisms or procedures, including profiling or video surveillance" (EU-C 2010b: 12). The EU Commission enforced these plans and inter alia induced the creation of a task force with experts from the industry to develop a template for PIA in the context of smart grid and smart metering. This template intends to give guidance to smart grid operators or other stakeholders on how to perform an impact assessment for corresponding systems (SGTF 2014). Neither the template nor the performing of PIA is mandatory but with an official recommendation[128] published in 2014, the EU Commission invited member states to encourage their stakeholders to perform a PIA on smart metering. The development of this template has to be seen against the background of the EU privacy and data protection reform. The smart metering PIA template is a showcase for a planned, further institutionalization of PIA in the European privacy and data protection regime. The new Regulation inter alia foresees a mandatory conduction of data protection impact assessments (DPIA) when specific risks occur as regards individual rights (Article 35 GDPR). A PIA is mandatory in the following cases: when "a systematic and extensive evaluation of personal aspects (…) based on automated processing, including profiling" occurs; when "special categories of data referred to in Article 9(1), or of personal data relating to criminal convictions and offences" are processed on a large scale; and when publicly accessible areas are systematically monitored on a large scale (GDPR Art. 35 (3)). The data categories listed in Article 9 are: "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation" (GDPR Art. 9(1)). Against the background of the GDPR coming into force in

[127] An independent advisory body established by Article 29 of the EU Data Protection Directive, consisting of different national and EU data protection representatives.
[128] Commission Recommendation of 10 October 2014 on the Data Protection Impact Assessment Template for Smart Grid and Smart Metering Systems http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32014H0724&from=EN

May 2018, it is likely that PIA will become an even more important issue in the European Union within the next few years.

The conditions defined in the GDPR help public and private institutions to determine when PIA is mandatory. However, the GDPR merely provides a legal minimum standard as regards PIA but does not include specific rules about its concrete implementation. This is obvious as the purpose of the law is to define legal rights and obligations but not to provide procedural guidelines. Nevertheless, the availability of frameworks and guidelines is important to enable organizations in carrying out PIA. There is a number of useful PIA guidelines and handbooks available which often vary in scope and approach. But there are some basic requirements which are relevant for every PIA approach, irrespective of its eventual peculiarities. De Hert (2012: 40) argues that from a human rights perspective there are three generic requirements for privacy affective information processing: "legality (Is there a legal basis in law for a technology that processes data?), legitimacy (Is the processing pursuing legitimate aims?) and necessity (Is the processing necessary in a democratic society?)." Although PIA is not limited to legal aspects and not to be misunderstood as sheer instrument of the law, these basic requirements for lawful data processing contribute to a vital fundament for PIA in general. But as De Hert (2012) notes, given the high complexity of legal issues, understanding and interpreting these requirements can be quite difficult. Equally important requirements are basic privacy principles (as outlined in Section 3.2.1) such as purpose und usage limitation of personal information collection, transparency and accountability, quality of the information, concepts for storage, access and protection to personal information etc. (OECD 2013b).

Irrespective of the variety of PIA approaches and guidelines in detail, there are some basic elements which can be found in most PIA models (e.g., Oetzel et al. 2011; Wright/De Hert 2012a; ICO 2014; CNIL 2015a; Bieker et al. 2016; Strauß 2017b). Hence, a PIA basically incorporates the following issues:

- Scope and objectives of the assessment (e.g., legal requirement, privacy compliance, foster transparency of information processing etc.)
- Description of the context, application or technology to be assessed
- Identification and assessment of personal information flows
    - What personal information is processed and how is it used (from creation/collection to processing, storage and deletion)?
- Identification and assessment of privacy risks
    - What are the risks and how likely are they?
- Identification of existing controls and recommendation of means to mitigate the risks (developing ways to avoid or reduce the detected privacy risks)
- Documentation of measures to resolve the risks and eventual remaining risks

These basic steps provide useful guidance on a general level. However, several issues aggravate the important role PIA could play to improve privacy protection. As mentioned, the function and meaning of PIA can vary significantly from country to country. This is

inter alia explainable by the complexity of the topic with varying national legal frameworks, many different national and international organizations, stakeholders, their interests and procedures etc. as well as a yet lacking legal obligation in most countries. Especially the strong role of policy and regulation is a double-edged sword. The development and employment of PIA approaches is closely linked to privacy policy, including legal and regulatory issues. PIA is often used as an instrument of policy with the motivation to create public trust (Clarke 2009). On the one hand, this is obvious as the law defines rules for privacy and data protection to differ legal and illegal intrusion. Public and private institutions conducting PIA have an interest to respect these rules in their workflows and processes to avoid legal problems. For them, PIA is mostly a tool of privacy compliance which is supportive for the organization itself as well as for the societal role of privacy protection. However, on the other hand, there is also a risk that PIA is misused as a fig leaf to distract eventual privacy concerns of the public about novel technologies. For instance, Bamberger and Mulligan (2012) analysed the role of PIA in the US. PIA is legally obliged for US government agencies when personally identifiable information is included in developing or procuring information technology systems. However, there are many problems and flaws of PIA in the US such as insufficient procedures, conflicting political interests, lack of independency of the agencies conducting the process etc. (ibid). Clarke (2009: 128) came to a similar conclusion and described the US as "a wasteland from the viewpoint of privacy policy". However, insufficient procedures and the like are not just a national problem of the US but also of other countries where similar problems as regards the effectiveness of PIA exist. Furthermore, as many ICT developments are located in the US, this has negative effects for privacy and data protection in Europe. Examples such as the case "Europe vs. Facebook"[129] or the regulatory issues in the EU concerning the dominance of Google[130] underline this aspect. These aspects are part of a regulatory problem on an international level[131] which is not further looked at in this research. Nevertheless, it highlights that the efficacy of privacy protection strongly depends from the agency of the privacy regime, i.e., its ability to act on a global scale. The upcoming new privacy regulation in Europe with the GDPR as a major instrument is a promising way to ease the situation in many respects. A common basic legal framework contributes to harmonize privacy and data protection regulation in the member states which can strengthen the European privacy regime in general. This development is promising to foster the role of PIA as a standardized instrument of privacy policy in Europe. On the longer run, this may have positive effects on privacy protection in the US as well.

However, the general issue remains that policy and regulation may lag behind technology which complicates privacy protection. It is rather obvious, that the law cannot cover all aspects of socio-technical reality. Because the main task of the law is to provide applicable rules that allow for a stable society, irrespective of particular technologies. In an ideal setting, the law enables society to react accordingly to new technologies. However, technology alters society and vice versa, and this interplay can create needs for additional

[129] www.europe-v-facebook.org
[130] See, e.g., Boffey (2017)
[131] This is inter alia related to the Privacy Shield regulation and Convention 108, as mentioned in Section 3.2.1

regulation and standardization. Nevertheless, it would be a great fallacy to occasionally adapt the law every time a new technology emerges. More important is to ensure that technology is developed and used in accordance with the law. The crux is to identify those issues resulting from technology and related socio-technical practices that impair the effectiveness of existing regulation. In such cases, adapting regulation is reasonable. As agency problems limit the effectiveness of privacy protection, reinforcing institutional transparency and accountability regarding information processing is an important measure. Here, new regulation can contribute to ease the situation. The European privacy reform, which inter alia led to the GDPR is an according example. Regardless of its controversial issues, the GDPR is about to ease some of the mentioned problems and strengthening privacy regulation, particularly by the obligation to conduct PIA under certain conditions (GDPR Art. 35) as well as the mandatory installation of a data protection officer for larger companies (GDPR Art. 37). At the moment, there are no legal obligations in this regard for companies in many European countries.[132] Consequently, many businesses do not even consider employing PIA, either due to unawareness of its purpose or in order to avoid efforts perceived as unnecessary. In this regard, the GDPR can improve the situation to some extent as institutions are encouraged to reflect on their privacy compliance. The GDPR fosters PIA in general. But for good reason, it cannot determine strict rules and detailed guidelines on how to conduct PIA covering every technology, application or practice, useful for every institution. This is not the task of the regulation but of the actors of the privacy regime. Although regulation is of fundamental importance, for PIA to be an effective tool, it is crucial to come toward approaches that allow grasping privacy issues not merely from a legal, but also from a wider, ethical and societal perspective in relation to technology. A narrow focus on legal issues can reduce the effectiveness of PIA as relevant privacy risks may be overseen (cf. Wright/De Hert 2012c). PIA can be particularly challenging in the case of ICTs which are a hybrid technology, involved in nearly every domain (Strauß 2017b). Raab and Wright (2012) analysed a number of different PIA methodologies and found several flaws. Besides other things, main issues are that most PIA concepts offer limited options to apply them to surveillance practices and ICTs. A basic reason for this is the relatively narrow focus of PIA on legal issues of data protection (Wright/De Hert 2012c). Although this focus is important and reasonable, it may be insufficient when the different types and dimensions of privacy are neglected then. Because socio-technical reality can bear privacy impacts that are not accordingly addressed by law.

Many factors and issues determine whether an individual's privacy is affected. As shown in the previous sections, privacy-intrusive activities can be conducted by many different actors. Some privacy impacts result from surveillance practices exploiting identity information, others as a by-product of common forms of information processing in various socio-technical contexts. Hence, an action does not have to be a modality of surveillance to be privacy intrusive. Socio-technical practices in general can affect privacy when they involve the collection, processing or storage personal information. The European Court of Human Rights thus declared that "[m]ere storing of data relating to the

[132] An exception is Germany where the installation of data protection officers is already required for most companies.

private life of an individual amounts to an interference within the meaning of Article 8 (right to respect for private life) of the European Convention on Human Rights" (EU-CHR 2017). As shown, technology design and usage entail and reinforce privacy challenges and risks. The progress in technology and privacy-intrusive techniques complicates to gather the extent to which a particular technology entails a privacy impact. Moreover, as with digital information, also the boundaries between different technologies blur. Hence, there are many socio-technical issues aggravating privacy impact assessment. Against this background there is demand for more systematic PIA approaches.

In total, PIA suffers from a lack of standards and theoretical concepts that address privacy impacts from a systemic perspective. Consequently, there is a certain risk that PIA processes vary in quality and from topic to topic, technology to technology etc. while at the same time, leave crucial privacy risks and issues unregarded. Although there are a number of useful PIA guidelines and handbooks available, they are mostly either very specific for a particular topic (e.g., RFID, smart metering), or to broad providing general organizational steps with emphasis on legal issues. For instance, the mentioned PIA template concerning smart grids and metering (SGTF 2014) offers useful guidelines for a specific domain. The intention of the EU Commission to use this template as a blueprint in order to stimulate PIA in other domains as well is reasonable. However, its specific character limits its scope and applicability. Other, basic PIA guidelines (e.g. ICO 2014; CNIL 2015a) are broader but with varying approaches. Some focus more on risk assessment others more on organizational issues and procedural steps etc. In each case, general obstacles result from relatively narrow, legally focused or diverse conceptualizations of privacy impacts. In this respect, the effectiveness of PIA suffers from a certain lack of common theoretical grounds about its key issues, i.e., the emergence of a privacy impact. This is particularly the case as regards ICTs. A critical issue is that the extent to which a socio-technical practice triggers a privacy impact varies with the involved institutions and technologies. When there is no common understanding of the emergence of a privacy impact, there is a certain risk that PIA arbitrarily considers some privacy threats but overlooks others. Therefore, in any case, PIA needs to be specified and tailored to domain in which it is applied to (including relevant organizational practices, technologies etc.). This necessary tailoring process could be alleviated by a more solid theoretical backing. As shown above, an integral element of every PIA approach concerns the analysis of personal information flows. However, the crux is often how to conduct this analysis. This can be a particular burden for organisations that may intend to conduct PIA (even without legal obligation) but shy away from the effort. Regardless of their usefulness, most PIA guidelines focus on analysing the explicit processing of personal information by a particular technology or application. Given the complexity of ICTs or digital technology in general, it can be challenging to assess what types of privacy are affected by what kind of activities processing personal information. The following section presents and discusses two approaches dealing with these issues.

6.2 Different types of privacy and privacy-affecting activities

There are different options to grasp privacy impacts and assess the risks privacy is exposed to. A precondition is a basic conceptual understanding of the (socio-technical) modalities that may affect privacy. This then enables to take a closer look at technology in particular. As outlined, the analysis of (personal) information processing is a core part of PIA. Irrespective of the technology, privacy impacts can result from many different activities and moreover, privacy as such can have multiple dimensions. One attempt to grasp these dimensions is to differ between different types of privacy.

6.2.1 The seven types of privacy[133]

Clarke (2006) provides a valuable classification of four major types of privacy: privacy of the person, privacy of personal behaviour, privacy of social communications and privacy of personal data. The first type of privacy makes reference to what is also known as bodily privacy, and aims at protecting the physical space and the body of a person. The second type of privacy aims at safeguarding the personal behaviour of individuals, such as religious practices and sexual activities. The third type covers some of the relationships and social ties that any individual builds and operates in. Finally, the privacy of personal data refers to the integrity and protection of all the sensitive data about an individual. Finn et al. (2013: 6ff.) complement Clarke's approach with additional dimensions and suggest seven types of privacy:

(1) Privacy of the person: addresses issues of keeping body functions and body characteristics (such as biometrics or genetic information) private. This type also refers to the strong cultural meaning of the physical body.
(2) Privacy of behaviour and action: one's "ability to behave in public, semi-public or one's private space without having actions monitored or controlled by others". This type includes "sensitive issues such as sexual preferences and habits, political activities and religious practices" (ibid: 6f.).
(3) Privacy of communication: the ability to communicate freely via different media without fear of interception, wiretapping or other forms of surveillance of communication.
(4) Privacy of data and image: this type includes issues about protecting personal data so that it is not automatically available to other individuals and organisations. Individuals should have the right to substantially control that data and its use. An image represents a particular form of personal data that can be used to identify and observe persons based on their visual characteristics.
(5) Privacy of thoughts and feelings: This type addresses one's freedom to think and feel whatever he or she likes without restriction. "Privacy of thought and feelings can be distinguished from privacy of the person, in the same way that the mind can be distinguished from the body. Similarly, we can (and do) distinguish between

[133] Parts of this section refer to Strauß (2017b) and are an extended version of the aspects discussed in this paper.

thought, feelings and behaviour. Thought does not automatically translate into behaviour". (ibid: 7)

(6) Privacy of location and space: addresses an individuals' ability to move freely in public or semi-public space without being identified, tracked or monitored. "This conception of privacy also includes a right to solitude and a right to privacy in spaces such as the home, the car or the office" (ibid: 7). Considering a growth in mobile computing devices such as smartphones and location data, the protection of locational privacy can be expected to increase in importance (cf. Blumberg/Eckersley 2009; Clarke 2012a).

(7) Privacy of association (including group privacy): affects one's right to associate with whomever one wishes without being monitored. Included are involvements or contributions in groups or profiles. This type of privacy is closely linked to other fundamental rights such as freedom of association, movement, and expression.

This typology can be useful to reflect on privacy impacts resulting from technologies or applications. Using these types as categories can be used to point out the extent to which a particular technology may have multiple privacy impacts. Table 4 exemplifies this with some common technologies. A distinction between different privacy types provides a more detailed picture of how privacy is affected, which may facilitate the creation of appropriate protection measures.

Table 4: Privacy types affected by different technologies. Own presentation (adapted from Strauß 2017b: 150) based on the typology of Finn et al. (2013). An "X" indicates that the privacy type is widely affected; "(X)" means that this privacy type is partially or potentially affected.

Technology / **Privacy of …**	(Smart) CCTV	Biometrics	Social Media	Smart phones
Person	(X)	X	(X)	(X)
Behaviour & action	(X)		(X)	(X)
Communication			X	X
Data & image	X	(X)	X	X
Thoughts & feelings			(X)	(X)
Location & space	X	(X)	(X)	X
Association	(X)	(X)	X	X

The most common, privacy-intrusive technology is CCTV. At first glance, CCTV mainly affects two types of privacy: those of the person and those of location and space. It may also affect privacy of association, namely if a group of persons is captured on CCTV. However, considering the combination of CCTV and a database for processing the images, this technology also touches privacy of data and image. When CCTV systems record activities, they also affect privacy of behaviour and action. New forms of "smart" CCTV

being able to recognize e.g., faces or patterns of behaviour (such as how a person moves) create additional impacts on this type of privacy. CCTV equipped with a microphone can also intrude into privacy of communication. Another example with increasing relevance is biometrics, i.e., technological means of recognizing individuals based on measurement of their physical characteristics (see also Section 5.3.2). In its broadest sense, biometrics focuses on (physical) information about the human body, most commonly the term refers to technologies such as fingerprint and iris scanners and increasingly face and vein pattern recognition and even DNA profiling. A biometric system is basically an automated pattern recognition system that uses information about one's body e.g., for identification and access control. Further processing of biometric information for profiling, pattern recognition and other forms of data mining is becoming more and more common in the security domain. While these technologies have obvious impact on privacy of the person or body, biometrics in a wider sense also affects other types of privacy such as privacy of data and image, as well as privacy of location and space (Strauß 2017b). A further example with multiple types affected is social media as e.g., demonstrated by Strauß and Nentwich (2013) in the case of SNS. In line with the core aims of SNS and other social media (see Section 4.3) to communicate, share information, create different forms of content and interact with others, privacy of communication, data and image, and association are affected by this technology. Depending on what information a user reveals, social media can potentially reveal many details of a person's life and thus also provide insight into that person's behaviour, shared thoughts or locations and places. The smartphone is another prominent technology that serves multiple purposes. It can be understood as a conglomerate of different intertwined technologies that entail a number of privacy impacts. The most obvious types are communication and association, as a phone primarily serves as a communication tool which then can also give insight into one's connections with other persons. But smartphones also serve many other purposes: they can reveal one's location and can be tracked, they can be equipped with a camera, used for online services to share information, e.g., via social networks or apps, or can even be used to scan biometrics etc. Hence, smartphone surveillance can give deep insights into the surveyed persons' behaviour, thoughts and feelings.

However, this mapping can only provide a general overview but not a strict assignment. Though the seven types of privacy provide a useful means to grasp the multiple dimensions of privacy, there are several limits regarding a more detailed impact assessment. The conceptual distinction between the different types allows gaining a quick, practical overview on the multiple ways a technology can be privacy-affecting. However, this benefit of a general overview is at the cost of a more detailed picture: as visible on the examples of social media and smart phones, it is rather difficult to grasp the extent to which ICTs affect privacy in depth. Contemporary and emerging technologies are increasingly interrelated and interoperable. As particularly the examples of social media and smartphones highlight, one socio-technical system can include a number of different technologies with accordingly broad privacy impacts. The nature of digital information and the convergence of technologies (e.g., a smart phone also has an integrated camera, SNS features include text-, audio- and video-based modes of interaction) make it likely, that in

several cases, every type of privacy might be affected. Hence, in such cases where a technology affects most or all types, there is little analytical gain without a further, more detailed analysis about how the different types are affected. Therefore, this typology merely represents a baseline to examine more in depth the extent to which a technology entails or reinforces privacy impacts (cf. Strauß 2017b).

Furthermore, a detailed view is aggravated because the approach does not include issues of identifiability and identifiable information. Each of these privacy types ultimately represents a form of identifiable information. However, the typology is basically a differentiation of personal information but it complicates to consider identifiable information resulting from technology usage. Therefore it provides limited options to analyse privacy impacts that result from implicit identification (as highlighted in Section 5.3). Consequently, less obvious or implicit forms of privacy intrusion may be unrecognized. This is problematic because, as shown in the previous sections, particularly implicit forms of identification can entail significant privacy impacts.

6.2.2 Privacy affecting activities

Another approach is to focus on privacy affecting activities. Privacy impacts can occur in manifold ways triggered by different activities. In his taxonomy on privacy, legal scholar Daniel Solove (2006) made an attempt to elaborate the various activities that affect privacy. He differs between four basic groups of activities that can be harmful to privacy: (1) information collection, (2) information processing, (3) information dissemination and (4) invasion. Each of these groups has related subgroups as shown in Figure 23 below:

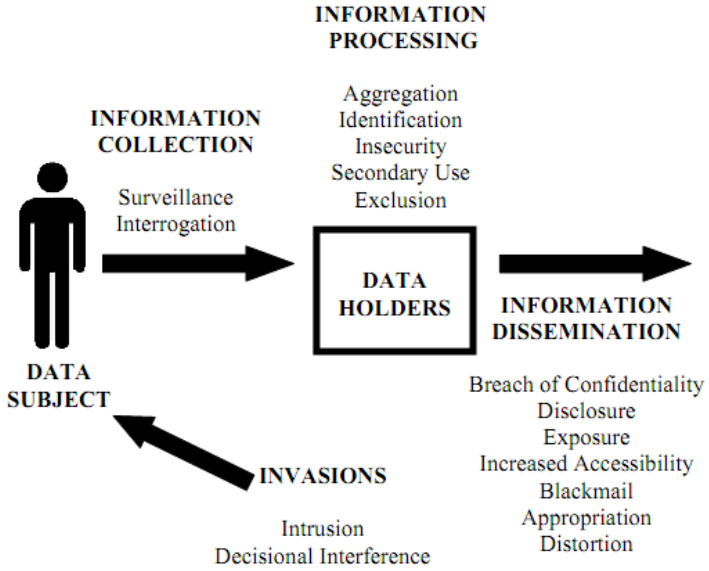


Figure 23: Groups of privacy affecting activities (Solove 2006: 490)

Solove's model begins with the individual (data subject) who is directly affected by the different activities. These activities can be conducted by various entities (e.g., other persons, government, and businesses) which are clustered as "data holders". These entities

collect information related to the individual which is then processed for further usage e.g., storage, combination, searching, manipulation etc. The next group, information dissemination, includes activities of data holders to transfer information to other entities or release it. The final group called invasions addresses activities that impinge directly on the concerned individual and her private affairs. In contrast to the first three groups which "pull" information away from the control of the individual, invasions challenge the individual directly and information is not necessarily involved (Solove 2006: 488ff.).
Each of these activities consists of multiple subgroups. The two subgroups of *information collection* are: surveillance which is defined as "watching, listening to, or recording of an individual's activities" and interrogation, i.e., "various forms of questioning or probing for information" (ibid: 490). *Information processing* includes aggregation, i.e., the combination of different data about a person; identification here is linking information to a particular individual; what Solove here calls insecurity means insufficient protection of the stored information so that it is vulnerable to leaks and improper access; secondary use is the use of information for other purposes without the concerned individuals consent; exclusion means that the individual is excluded from knowing and controlling how her information is processed by whom, for what purpose etc. The subgroups of *dissemination* describe privacy-affecting issues of transferring or spreading personal information: breach of confidentiality; disclosure, i.e., "the revelation of truthful information about a person that impacts the way others judge her character" (ibid: 491); exposure, i.e., "the exposing to others of certain physical and emotional attributes about a person" (ibid: 533); increased accessibility, here the possibility to access available information without the control of the individual; blackmail, i.e., the threating of an individual to disclosure her personal information; appropriation, i.e., the misuse of an individuals' identity to serve other interests such as identity theft; and distortion, i.e., "the dissemination of false or misleading information about individuals" (ibid: 491). The two subcategories of *invasions* are intrusion which "concerns invasive acts that disturb one's tranquillity of solitude"; and decisional interference, i.e., "governmental interference with people's decisions regarding certain matters of their lives" (ibid: 554).

Solove's model is particularly interesting as it sheds light on how the handling of information as part of different types of activities can affect privacy. It is thus another important contribution to grasp the multiple dimensions of privacy impacts. His taxonomy offers a detailed description of the different groups of activities he identified enriched with many legally relevant examples. However, the model remains relatively vague as regards whether the subgroups are particular activities, or consequences of the main groups they belong to. It is also not always clear, why an activity is part of a group and not of another. For instance, surveillance is seen as a part of information collection, which is reasonably separated from aggregation, being a subgroup of information processing. Surely, surveillance inevitably involves collecting information and the scope of surveillance can expand by further information processing. However, the aggregation of once separated information can also be part of a surveillance activity. Strictly speaking, this relation cannot be derived from Solove's model as surveillance is here presented as activity of information collection prior to processing. Further ambiguous is the framing of identification as a subgroup of information processing. On the one hand, Solove refers to

Clarke (1994a) and understands identification as "connecting information to individuals" whereas identification "enables us to attempt to verify identity" (Solove 2006: 510). On the other hand, in Solove's model, identification is merely included as one activity of information processing amongst others. Consequently, identification, i.e., the processing of identifiable information, here is not seen as a necessary condition for an impact on privacy or at least not presented as such. For Solove, identification "entails a link to the person in the flesh" while aggregation can, but not necessarily allow for identification (ibid: 510f.). Although this is basically the case, privacy-affecting aggregation of information concerning an individual inevitably involves some processing of identifiable information; because otherwise, she (and thus her privacy) cannot be affected directly. Therefore, as shown in the scope of this research, the processing of identifiable information is precisely the kind of process that triggers intrusion into the privacy of an individual. Identification certainly implies information processing, but it is more than just a privacy affecting activity amongst others: identification is at the core of privacy affecting activities. This is because it enables to link a piece of information to a particular individual so that this individual is distinguishable from others. This is the basis for eventual intrusions into the privacy of this very individual. However, this does not necessarily imply that information directly links to a person "in the flesh". It also does not mean that for privacy intrusion a person has to be uniquely identified. Because the real identity of person (e.g., her full name) is not necessary to have her privacy infringed (e.g., by observation or surveillance). Nevertheless, identifiable information is involved when informational privacy is affected. Ultimately, any information related to a person's identity may be used to intrude into her privacy. Because it can be sufficient that information related to a person's identity is gathered (e.g., against her will or awareness) for an activity to be privacy intrusive. Therefore, identifiability is a core problem of privacy protection.

In fact, Solove's privacy affecting activities basically involve the processing of some kind of identifiable information. Most obvious are identification and appropriation. But also surveillance and interrogation comprise the collection of information about a person; privacy-affecting information aggregation implies that personal information is involved; secondary use means processing of personal information without the consent of this person; breach of confidentiality implies information that a person perceives as confidential; disclosure involves true information about a person's character; exposure means that physical and emotional attributes about a person are revealed; increased accessibility refers to access to personal information; also blackmailing a person requires information about that person; and distortion because the dissemination of false or misleading information about a person implies to have correct information about that person. Hence, even though identification does not play a prominent role in Solove's model, the processing of identifiable information implicitly does.

The perspective on privacy-affecting activities is particularly important as it provides a more process-oriented view on privacy. However, the activities are closely interwoven and may be hard to distinguish when, e.g., ICTs are involved. In many cases, the sheer collection of digital information can lead to dissemination of this information more or less at the same time (e.g., in social media). This does not reduce the analytical relevance of

these categories. But there is demand for alternative approaches as well that incorporate the peculiarities of ICTs and digital information processing.

As shown in the previous sections, identifiable information can be gathered and processed in various ways, supported and reinforced by technology. In order to understand and conceptually grasp the modalities of information processing that may affect individual privacy, it is thus crucial to perceive identification as a core activity in this regard. Some form of identification is mostly involved when an individuals' privacy is affected, at least as regards informational privacy. The crux is that the individual is often not aware about how her identifiable information is being processed or involved in an information processing.

This does not mean that every form of identification entails a privacy intrusion. For instance, recognising the face of a by-passing person may also include some form of identification, though a volatile, non-persistent one. If technology is involved as e.g., one takes a photograph or a video of a person for a particular purpose, this has different implications as the identifiable information (the image or footage) becomes somewhat persistent/durable and available for further processing. As shown (Section 5.3), technology expands identifiability as the contexts of information processing expand as well. Consequently, there are more options to intrude into privacy. As identification is context-sensitive, it also depends on the context or purpose and degree of identification whether privacy is affected or not. Nevertheless, the processing of some identifiable information referring to a particular person is a precondition for the emergence of a privacy impact concerning that person. It thus is reasonable to focus on the multiple dimensions of identifiability and different types of identifiable information as the following section suggests.

6.3 Towards an identifiability-based framework for privacy impact assessment

As discussed in the previous sections, PIA is an important instrument to improve the effectiveness of privacy protection. However, current approaches are relatively diverse with several limits in scope. There is thus a certain demand to enhance the modalities of PIA. Yet there is no conceptual framework that takes the role of identifiability explicitly into account, which would be vital to come towards a common understanding of privacy impacts, applicable for multiple domains. On a more general level, this demand results from the classical problem of privacy protection with an inherent difficulty to comprehend the mechanisms of privacy intrusion, stemming from a relatively static view on privacy (as discussed in Section 5.4) including a relatively narrow focus on personal data or information. The dynamics of ICTs and expanding identifiability (Section 5.3) aggravate this problem.

A central part of every PIA concerns the analysis of personal information flows, which basically enable the emergence of a privacy impact. Existing PIA guidelines, irrespective of whether being very detailed or rather general, thus mainly focus on personal

information. This is obviously essential. However, a critical issue is the fact that due to ICTs it becomes increasingly difficult to determine what personal data or more precisely personally identifiable information (PII) *exactly* is. Legal definitions (Section 3.2) are reasonably kept broad and provide important minimum standards, but ICTs challenge their meaning in socio-technical practices. Consequently, the operationalisation of PII to properly implement PIA can be hampered. When analysing a personal information flow, the focus typically is on information directly related to the person, processed within the systemic boundaries of a single information system. Considering the complexity of digital technology, PIA models of this kind may be insufficient to grasp the scope of a privacy impact. Because in fact, personal information is often not processed by a single information system anymore, but rather of multiple information systems, which are increasingly networked and constantly process information in manifold ways. Moreover, as highlighted with the identity shadow (in Section 5.3), it is not always clear what kind of information counts as personal information or to what extent technology-generated information refers to the identity of person as well. Approaches that merely consider explicit types of PII (e.g., directly person related such as name, address, identification no., biometric information etc.) are likely to neglect other forms of identification and risks of identifiability (as presented in the previous sections). Technological design and socio-technical practices further complicate these issues.

Therefore, there is demand for a deeper conceptual understanding of the emergence of a privacy impact. A more systematic PIA approach with emphasis on identifiability and different types of identifiable information could contribute to overcome some of the existing limits. I thus suggest a basic PIA framework which comprises identifiability as core determinant of the emergence of a privacy impact. This allows for a more systemic perspective on privacy impacts which is vital to improve the quality of PIA as well as of privacy protection concepts.

In general, identifiability is the condition which results in an individual being identified (directly or indirectly) based on a set of personally identifiable information (ISO 2011: 2). Therefore, it is here understood as the initial condition of a privacy impact. At the core, the processing of identifiable information can be understood as a main trigger for the occurrence of a privacy impact. In this regard, a privacy impact can be seen as (possible but not inevitable) result of a process (or a chain of processes) in which identifiable information is employed, for one or several application contexts. When an individual is directly or indirectly involved in the information processing chain, her identifiability is enabled. This process can comprise one or several socio-technical systems. The extent to which privacy is affected by a socio-technical system thus depends from the capability of this system to process identifiable information. A narrow view on a technology or application as an isolated information system leaves opaque whether and how the information may flow into another context/system. Therefore, in line with the boundary control function of privacy, it is important to consider the amount of systems and application contexts involved. The complexity of information processing is likely to increase with the number of information systems, which may intensify the risks of privacy intrusion as ISD further decreases. A major reason is that each system may entail the creation of additional identifiable information and thus extend the identity shadow, which

can be exploited for further purposes. Furthermore, each system can be intruded or misused and the more systems are available the higher the room of possibilities for intrusion. To assess privacy risks and the according demand for protection mechanisms thus requires transparency and verifiability of information processing and of the technological systems involved. A perspective on the meta-architecture of a socio-technical system in which identifiable information is processed contributes to this and alleviates to analyse the extent to which privacy is affected. This is vital to come toward more systematic conceptualisations of privacy impact assessment.

Figure 24 sketches a general PIA framework with an emphasis on identifiability and identifiable information. This simplified model illustrates the interplay of basic elements of a socio-technical system determining a privacy impact, such as (1) the amount and type of personal identifiable information being processed, (2) the entities gathering and using this information, (3) the context(s) of processing (varying in space and time), (4) the amount and type of information (sub-)systems and how they process the identifiable information. These factors are interwoven and primarily affect identifiability, entailing additional privacy risks.

At the core of this framework is the flow of identifiable information which is determined by the modalities of information processing (i.e., how the information is collected, used, transmitted, stored etc.). These modalities affect the lifecycle of identifiable information, shaped by the interplay of entities, the contexts for which they use the information as well as of the design of the socio-technical system including its sub-systems (e.g., technologies, applications, databases and services involved). A variable number of entities can be involved in the socio-technical system with different functions and roles: e.g., system operator, provider of technologies, applications or services, information processing software agents etc. Besides these internal entities there can also be external entities such as third parties with information access or to whom information is transferred for additional purposes (e.g., contractors of the operating entities, advertising companies, law enforcement agencies external systems etc.). These entities employ technologies, applications, services etc. as well, which may automatically gather and process identifiable information. Hence, among the involved entities are not necessarily institutions or individuals only, but also technical systems or informational agents (e.g., algorithms gathering information, software bots and the like). Considering a general increase in machine learning, (semi-)automated systems and autonomic computing with progress in the field of artificial intelligence, such machine entities are likely to gain in importance within in the next years.

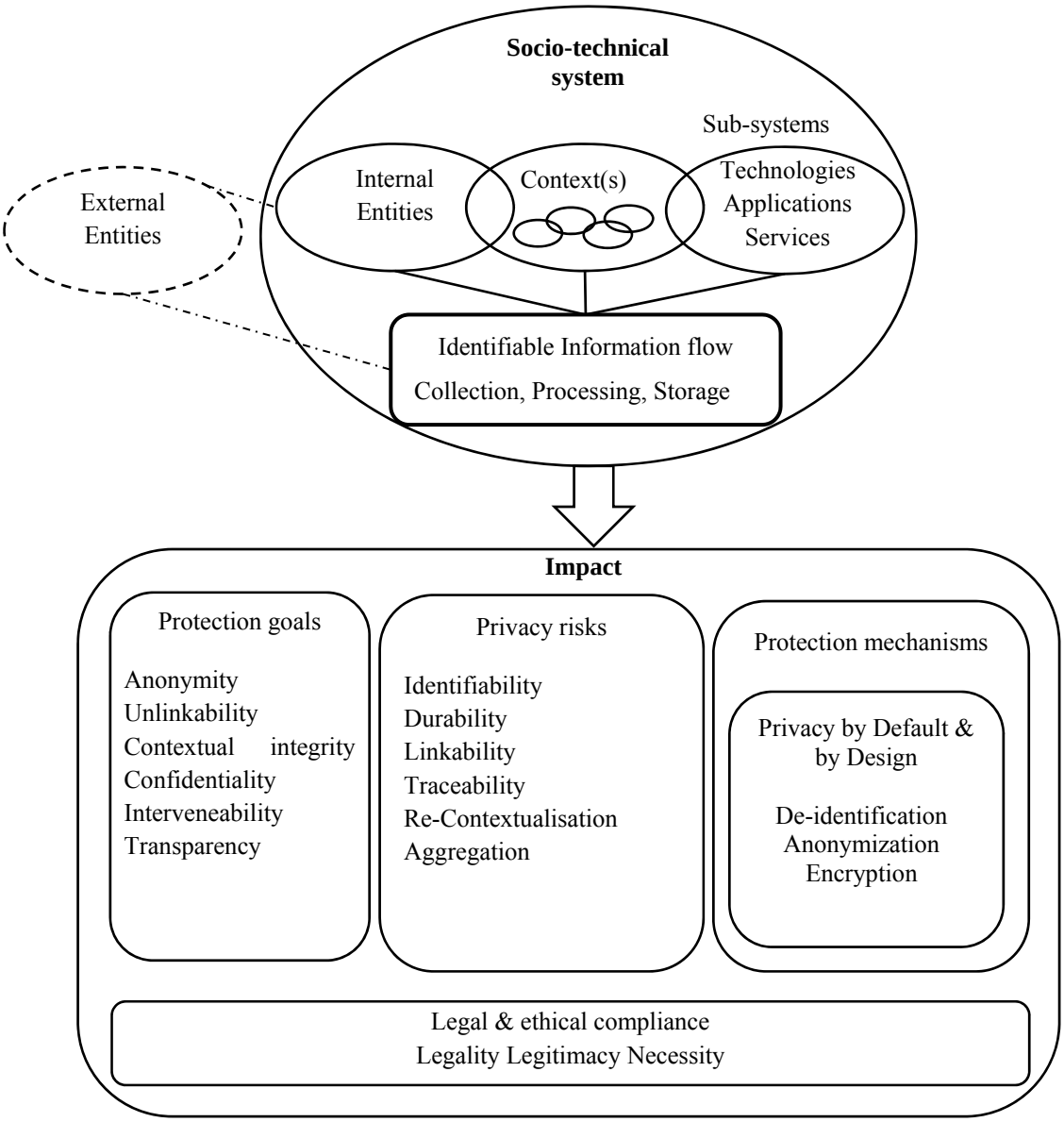


Figure 24: Identifiability-based framework for privacy impact assessment

This setting as whole affects the occurrence of a privacy impact including the extent to which privacy is exposed to different risks. Above all, the primary risk is identifiability which entails a number of additional risks that can reinforce identifiability. The occurrence of a privacy risk resulting from identifiability does not necessarily imply that a person's identity is unambiguously revealed. Already the fact that identifiable information about a person is gathered can be a sufficient condition for a privacy impact. The basic reason is that this information links to an individual's identity and thus may be used to take actions or decisions affecting the person and interfering with her privacy. The extent to which a socio-technical system provides identifiability of an individual is at least determined by the following core factors: a) the amount of (personally and technically) identifiable information; b) the durability (or persistence) of identifiable information, determining its temporal availability; c) the linkability of identifiable information, enabling to cross-link information from separated contexts; and d) traceability allowing for reconstructing trails

of an individual's identity (e.g., to trace an individual's movements, behaviour, actions, relationships etc.). Durability fosters traceability and linkability because the longer identifiable information is available, the more traces and the more options for cross-linking information may occur; linkability may extend durability because combined information may undermine temporal limits. Similar for traceability as e.g., even deleted information may leave reproducible traces[134]. Hence, these factors are interrelated can affect each other mutually. These interwoven factors shape the condition of identifiable information. They also represent risks as they enable privacy intrusion in many respects. Not least the storage and retention modalities can be critical as risks are likely to increase when identifiable information is accessible for longer periods of time. It makes a difference whether identifiable information is available only for a limited timespan (such as a dynamic IP address changes after a certain period of time), or whether this information is stored, temporarily or permanently available, separated or linked with other information, as well as whether information is bound to a context or traceable without any limits. Two major risks emerging from these conditions are: re-contextualization or secondary use, i.e., the use of information for additional privacy-affecting purposes; and aggregation, i.e., the combination or linkage of information from different sources. Aggregation also enables de-anonymisation or re-identification by combining different types of information to create (quasi-)identifiers (as highlighted in Section 5.3). This can result in profiling, e.g., the use of information to recognise patterns about particular persons and create identity profiles.

Thus a privacy impact is shaped by a variety of issues resulting from the modalities of information processing. It can make a qualitative difference what type of information that is, how this information is gathered and processed, whether its availability is limited in time, or it is stored and used for other purposes, what technologies are involved etc. In practice, the purposes for collecting and using information often differ or feed into additional purposes beyond the initial processing context. For instance, the primary purpose for collecting identifiable information may be user registration to enable access to a service including an identification/authentication procedure for securing a transaction, to fulfil a legal obligation or for CRM. But there can be other, additional purposes involved such as third party access for commercial purposes, profiling or surveillance activities etc. (as shown in Section 5). Secondary use is particularly problematic when it is without the concerned individuals consent. A person might prefer to avoid information provision to a service if secondary use is foreseen, such as for targeted advertising, profiling etc. Hence, secondary use can lead to breach of confidentiality as individual users may not be aware of all contexts in which their information is used for. Moreover, (as discussed in Section 5.4.3.), informed consent is often insufficient to prevent from unintended further use; particularly when the individuals concerned find themselves in an accept-or-leave situation. But insufficient privacy protection is not merely a risk for the individuals concerned but also for the institution responsible for the information processing. Lacking protection of information processes can reinforce security risks such as unintended

[134] The secure erasure of data on digital storage components (e.g., hard disk drives) is an issue in computer science. Typical, secure approaches are overwriting information with random values as a simple software deletion is insufficient. Data on modern storage devices such as solid state disks are more difficult to erase. For more details see, e.g., Wei et al. (2011).

disclosure to external entities. From a wider perspective, this may lead to a reduction in system stability, when e.g., information systems, applications etc. are vulnerable to abuse, attacks, data loss, data breaches, or identity theft.[135] This is another reason, why privacy protection is not to be misinterpreted and falsely reduced to an issue of individual concern only. This aspect also highlights that privacy and security are often complementary and on the same side of the coin.

In order to address privacy risks and stem uncontrolled privacy impact, protection goals are crucial. A general precondition for the processing of identifiable information is its compliance with legal regulations (see also Sections 3.2.1 and 6.1). As privacy is not merely a legal but an ethical issue, ethical compliance is an essential requirement as well. It is thus vital that the protection goals are informed by a combination of these basic requirements. A PIA process can serve many different objectives of the organization conducting it. However, irrespective of specific strategic objectives, there are some fundamental protection goals to reduce privacy risks. As security is an important, related issue as well, it is crucial to find a proper balance. In information security there are three typical security goals, i.e., confidentiality, integrity and availability aiming at ensuring that information is authentic, comprehensible, confidential, and protected from unauthorized access (cf. Hansen et al. 2015; Bieker et al. 2016). These objectives largely correspond with goals of privacy protection in cases where the processing of identifiable information is necessary. However, a sheer focus on these three is not sufficient to reduce privacy risks regarding identifiability. To some extent, there can be conflicts with privacy protection. For instance, availability of information to authorized entities may be in conflict with minimum disclosure to achieve a high level of confidentiality. To extent the scope of protection goals with respect to privacy and data protection issues, Hansen et al. (2015) suggest unlinkability, transparency and interveneability as three additional protection goals for privacy. This six goal approach is promising to enhance the development of protection mechanisms in accordance with privacy by design. A crux is that a full achievement of each goal at the same time is often not possible. Thus it can be challenging to properly balance different goals (ibid; Bieker et al. 2016). However, to some extent, tensions can be avoided by a stronger emphasis on identifiability and the introduction of contextual integrity (Nissenbaum 2010, see Section 3.3.2) instead of availability. Availability is thus not considered as a protection goal in the proposed framework because it misleadingly suggests information disclosure. In fact, it means that information is available for authorized use only. Contextual integrity covers this aspect as it implies information is properly processed for a particular purpose accepted by the concerned individual only and not for others. This includes secure handling of this information within this application context.

The primary protection goal in this framework is to process only a minimum amount of identifiable information, i.e., non-identifiability which equals anonymity. This is in line

[135] There is growing number of data breaches observable in the last years at a global scale. Among others, prominent examples are: comprised information of 77 million customers of the Sony play station network in 2011; the so far biggest case of a security breach in 2014, affecting over 500 million user accounts of internet company Yahoo (in 2016, Yahoo reported that about one billion accounts may be compromised already in 2013); theft of four million customer records of a British telecom provider in 2015; to name just a few (cf. Newman 2016; Dunn 2017).

with the basic privacy principles (see Section 3.2.1) of data minimization, purpose and usage limitation. The basic aim is thus to minimize identifiability, reduce according risks and allow for anonymity and pseudonymity wherever possible. Even though this goal is not fully achievable in many contexts, it makes sense to use it as an ideal condition or best case scenario serving as a fundamental reference point for the further objectives. These are unlinkability, which is a crucial requirement to avoid that identifiable information is aggregated and cross-linked from multiple contexts. Depending on its implementation, unlinkability is the basis to allow for anonymous and pseudonymous information processing (see also section 3.1.3). Integrity in a technical sense means to ensure that information is reliable, authentic and correct (cf. Bieker et al. 2016). As already mentioned, here it is meant in a broader sense in line with Nissenbaum's (2010) concept of contextual integrity which is preserved when informational norms are respected so that information is not used for other purposes than the individual has given her consent. Confidentiality means that identifiable information is kept confidential and not disclosed to non-authorized entities. Confidentiality thus includes unobservability, i.e., to avoid that an individual is traceable by her information (cf. Solove 2006; McCallister et al. 2010; Hansen et al. 2015; Bieker et al. 2016). Interveneability primarily incorporates informational self-determination and control of the individual concerned so that she can intervene when necessary; enforce changes and corrections of her information. Practical examples are privacy settings, deletion of information, or revocation of consent. In addition, interveneability is crucial for supervisory authorities such as DPAs to intervene, e.g., in case of privacy violations. Finally, transparency means that all privacy-affecting information processing activities are understandable and open to scrutiny. Transparency is thus a precondition for accountability so that it is comprehensible and verifiable whether the responsible entities appropriately handle and process the information (cf. Hansen et al. 2015; Bieker et al. 2016).

In order to achieve these goals and reduce privacy risks, protection mechanisms have to be set up. This basically refers to implementations of privacy by design and by default (as presented and discussed in Section 5.4). Useful guidance about the implementation of common safeguards can be found in IT security standards and frameworks (e.g. BSI 2008; ISO 2011; McCallister 2010; EuroPriSe 2017). Typical controls include organizational as well as technical measures such as physical controls, organizational norms, operational access restriction procedures, guidelines, privacy and security policies, authentication, role and access management, and encryption methods. The concrete requirements of protection mechanisms and their usefulness obviously depend from the particular application context, the organizational setting etc. Furthermore, also social practices and the privacy culture in an organization affect the level of protection. However, irrespective of details, control and protection measures should incorporate basic privacy principles (e.g. ISO 2011; OECD 2013b; EuroPrisE 2017) which can be seen as core requirements to achieve protection goals and address the risks. In particular the commonly accepted OECD privacy principles for fair information practices are important guidelines such as data minimization and avoidance, purpose limitation, minimum retention and storage duration and deletion of unnecessary data etc. (OECD 2013b). These principles are mentioned in most privacy

frameworks and also of the European Privacy Seal EuroPriSe which provides detailed descriptions about privacy controls in accordance with EU legislation (EuroPriSe 2017). Altogether, effective protection mechanisms contribute to privacy-preserving processing, retention and storage modalities to limit unnecessary collection, use, retention, sharing and disclosure of identifiable information. For the achievement of technical protection the implementation of privacy by design and by default is essential. Wherever possible, information should be de-identified. De-identified information means that its capacity to identify a person is removed, e.g., by deletion, masking or obscuring identifiable parts of an information set (e.g., of an identifier). Techniques to reduce the risk of re-identification can be inter alia reduced by generalization so that information is grouped and made less accurate, replacing selected values with average values, by erasing parts of identifiable information, or adding noise to the information (cf. McCallister et al. 2010). Technically, this can be achieved with cryptography which provides several methods for de-identification and anonymization. With the use of pseudonyms (see Pfitzmann/Hansen 2010 as outlined in Section 3.1.3) different levels of linkability can be established so that application contexts that may require identification can be protected from misuse as identifiers are not easily linkable. With methods of encryption, the risk of unlimited information disclosure can be reduced (for more details about technical PbD concepts see Section 5.4).

Hence, in general, there are many protection mechanisms available to improve privacy protection. However, a crux is that it is often unclear how and what kind of identifiable information is being processed. This hampers to analyse the mechanisms and practices that may induce a privacy impact as well as the development and deployment of effective safeguards. The proposed framework for privacy impact assessment with a focus on identifiability can contribute to improve this situation. A core piece of this framework is the flow of identifiable information. In order to determine this flow, the following section suggests a typology of identifiability with different basic types of identifiable information. For an overview of the basic steps of a PIA process with respect to the presented identifiability-based framework, see Section 6.3.2.

6.3.1 A (draft) typology of identifiable information

As outlined, the analysis of personally identifiable information flows is a core task of every PIA process. However, technology and a lack common understanding about the basic types of this information impede this task, and thus the assessment of privacy impacts. The crux is that ICTs aggravate to determine what counts as PII because as shown, digital information offers myriads of ways to identify a person with and without her involvement. There is yet no generally valid typology that appropriately considers the role technology has on identifiability and identifiable information. As discussed previously (in Section 6.2), typologies such as the seven types of privacy (Finn et al. 2013) or Solove's (2006) description of privacy-affecting activities rather focus on personal information and do not sufficiently address identifiability in a broader sense. The term personal information implies information originating from a person while identifiable (or identity as synonym)

information is conceptually broader, even though it refers to the identity of a person as well. The relevance of this distinction for PIA, particularly as regards technology, will be discussed in the following. While every kind of personal information is a type of identifiable information, not every type of identifiable information necessarily results directly from a person. As shown in Section 5.3.1, technology-specific information may enable various forms of implicit identification. Technology usage may automatically generate identifiable information which can be used to identify a person, even though this person did not provide the information directly. It is often sufficient to have some piece of information relating to particular person gathered from technologies. Identifiable information is generic and context-dependent. In one particular context, information might not refer to one's identity. However, linked to another, it can then become personal identifiable information due to this linkage of contexts. Information aggregation is thus is a potentially privacy intrusive practice. The more data is available, the easier it is to find patterns out of them which can be used for identification purposes and thus privacy intrusion. Through aggregation, identity patterns can be refined, which is particularly fostered by ICTs. Hence, the identifiability of the individual grows with the amount of identifiable information and its usage contexts. Taking these dynamic characteristics of identity and (personal) information more into account may contribute to improve privacy protection in general.

Current privacy standards mainly focus on personal data or personally identifiable information (PII). Standard guides to protect PII such as of the US National Institute of Standards and Technology (NIST) define PII broadly as "any information about an individual maintained by an agency, including (1) any information that can be used to distinguish or trace an individual's identity, such as name, social security number, date and place of birth, mother's maiden name, or biometric records; and (2) any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information" (McCallister et al. 2010: 1). The ISO/IEC 29100:2011 privacy framework[136] offers a similar definition of PII as "any information that (a) can be used to identify" a natural person "to whom the information relates, or (b) is or might be directly or indirectly linked to" (ISO 2011: 2). According to this framework, there are at least four instances where information can be seen as identifiable information: "if it contains or is associated with

- an identifier which refers to a natural person (e.g., a social security number);
- an identifier which can be related to a natural person (e.g., a passport number, an account number)
- an identifier which can be used to establish a communication with an identified natural person (e.g., a precise geographical location, a telephone number);
- or if it contains a reference which links the data to any of the identifiers above." (ISO 2011: 7).

[136] This ISO framework is an international privacy standard that supports persons and organizations in dealing with organizational and technical privacy issues of ICT systems.

Hence, information does not necessarily consist of an identifier to be understood as PII. Also a specific characteristic which distinguishes a natural person from others (e.g., biometric data) can be considered as PII (ibid). An identifier can be used directly for identification while a characteristic can be used for identification in combination with other information. For example, a name is an identifier; eye colour is a characteristic which alone does not identify a person, except she is the only person in a crowd with this eye colour. The combination of eye colour and other body-specific characteristics may allow finding a particular pattern that makes an individual uniquely identifiable in a certain context. Moreover, if characteristics are linked to an identifier, then the degree of identifiability increases. But also other information, which contains neither an identifier nor a specific characteristic, can relate to a person and enable identification. The quality of identification inter alia depends from the uniqueness of the identifiable information in relation to its environment. For instance, an identifier consisting of a two-digit number (ranging from one to ten) is obviously insufficient to uniquely identify a person in a group of hundred people. This is a matter of statistics and mathematics and concepts like k-anonymity (Sweeney 2002), t-closeness (e.g. Li et al. 2007) or differential privacy (Dwork/Roth 2014) deal with this issue from a technical perspective (see also Section 5.4.1). Irrespective of the technical aspects, this dependency from environmental factors affects identifiability. Therefore, the quality of identifiable information may alter with a changing socio-technical environment. In general, identification is closely related to pattern-recognition techniques, as any type of identifiable information can be suitable to explore (unique) patterns which may then serve identification purposes. Given these issues, it can be challenging to grasp the exact amount of identifiable information, particularly when ICTs are involved in the processing. Although the outlined definitions and standards are essential, there is a general lack of coherent guidelines on how to assess PII and in particular as regards technology-induced identifiable information. Existing typologies of PII suffer from reductionist perspectives which only frame those information as PII which "actually identifies a specific individual" in a specific context (Schwartz/Solove 2011: 49). Hence, they are mostly limited in scope and do not mirror contemporary technology and socio-technical practices. As a consequence, risks of de-anonymisation or re-identification and use of technology-specific identifiable information often remain unrecognized. A further issue is that the types of identifiable information are not conceptually grasped but mainly derived from legal frameworks only. With different national legal frameworks, also the understanding of personal information can differ. Among other scholars, De Andrade et al. (2013: 21) argued that there is a lack of a common legal understanding of electronic identification and EU data protection legislation "is not sufficient to cover all the aspects involved in the protection of and management of electronic identities". Nevertheless, the GDPR (Article 4(1) offers an important baseline as its definition of personal data includes direct and indirect identification and considers identifiable information as privacy relevant (see also Section 3.2). However, for conducting privacy impact assessments this baseline is of limited use, especially when aiming for ethical compliance that goes beyond legal requirements. Most PIA approaches provide some guidance on how to assess personal data such as CNIL (2015b) differs between common and sensitive personal data. Although this distinction is legally important it is limited in scope and can be ambiguous. For instance,

the mentioned examples for common data are among others civil status, identity, identification data, financial situation, connection data such as IP addresses, location data etc. Data listed as perceived to be sensitive are, e.g., biometric data and bank data. And sensitive data in a legal sense are, e.g., philosophical, political, religious and trade-union views, sex life, health data, racial or ethnic origin (CNIL 2015b). Also additional distinctions between content and non-content respectively metadata (data about data), referring to the peculiarities of digital information processing are not appropriately considered in PIA. Such a distinction was inter alia used in the currently abandoned EU Directive on the retention of telecommunications data (as mentioned in Section 5.1.2). Proponents argued that as no content of communications was gathered but only metadata, the individuals' right to privacy would not be violated. However, critics highlighted that metadata provides deep insights into personal communications and a separation between content and non-content is not feasible in socio-technical reality. Not least as e.g., NSA surveillance practices make heavy use of metadata (cf. Greenwald 2014, Schneier 2014). In his rulings the EU Court of Justice made clear that the retention of metadata is deeply privacy intrusive (CJEU 2014/2016). This case underlines that ICTs reinforce diminishing boundaries between different types of personal and non-personal information, which challenges to determine the extent to which privacy is affected. It generally makes sense to consider metadata in privacy impact assessment. However, metadata are either unconsidered or vaguely associated with PII. In the ISO privacy framework (ISO 2011), for instance, metadata are only incidentally mentioned as example for information not easily visible to the user, such as properties of a document. In total, the role of metadata for PIA is rather ambiguous. A stronger consideration of metadata in PIA is relevant but not sufficient as it is conceptually too broad. In total, existing typologies of identifiable information for PIA are relatively erratic and not systematic. Furthermore, there is a lack of approaches explicitly considering identifiable information resulting from technology usage. There is thus demand for updated concepts of PII. Accordingly, the following section suggests four basic dimensions of identifiability to grasp the different types of identifiable information.

6.3.1.1 Four basic dimensions of identifiability

As shown in the previous sections, there is a broad range of possibilities to directly or indirectly, explicitly or implicitly identify individuals. ICTs reinforced identifiability and stimulate a further expansion thereof which may span across a multiplicity of socio-technical contexts. These conflating contexts further complicate privacy protection. Therefore, making the boundaries between different contextual identity layers or subsystems more visible, i.e., increasing the transparency of explicit and implicit identification processes is an important step toward more effective privacy concepts. To achieve this requires a deeper understanding of identifiability and the emergence of identifiable information. Identifiable information not necessarily originates from the individual concerned but can also result from technology usage without direct involvement of that individual. Hence, it can be difficult to even detect such types of information due to the widely lacking options to categorise it as type of PII. Consequently, potential privacy

impacts of the corresponding information processing may be unrecognised. Thus I argue for a more detailed consideration of these types of information in PIA processes and a distinction between personal and technical identifiability (although both are intrinsically linked). Therefore, a typology of identifiable information should not merely comprise PII (as discussed above) but also information related to the technologies that are involved in the processing of PII, i.e., technically identifiable information – TII. Although both types are strongly interrelated it is reasonable to explicitly consider identifiable information emerging from technology. TII is defined as any information resulting from the employment or use of a technology that refers or relates to an individual and can be used to (directly or indirectly) identify the individual. Technology primarily means ICTs. As will be shown, this distinction between PII and TII can be of practical use, supportive to improve privacy impact assessments. However, these two types alone are not sufficient to improve the general understanding of how identifiability and thus how identifiable information emerges. From a theoretical stance, the dynamics of identifiable information processing in general may be characterized by a set of basic dimensions which may be valid for personal and technological entities alike. Such basic dimensions could improve the theoretical understanding of identifiability, which is supportive to enhance PIA and PbD approaches. I thus further suggest the following four basic dimensions or meta-categories of identifiable information. It has to be noted that these dimensions are far from being comprehensive but are an attempt to theoretically grasp the emergence of identifiability with respect to privacy impact assessment.

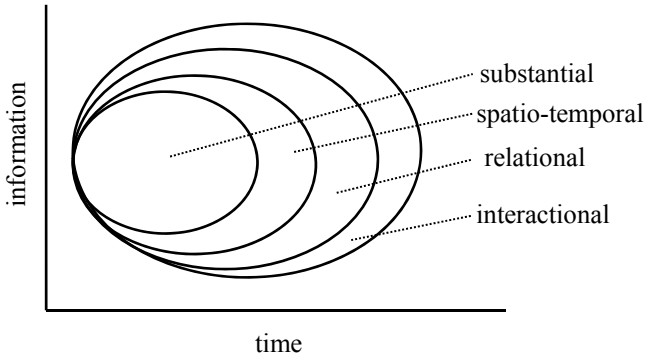


Figure 25: Four basic dimensions of identifiable information

As illustrated in Figure 25, the dimensions are not isolated but interrelated and thus may also be seen as layers. The basic rationale for these dimensions is informed by the dialectical character of identity with its relatively static or stable items as well as its dynamics (what Ricoeur called idem and ipse, see Section 3.1). However, given the high dynamics of ICTs, a sheer distinction between stable and dynamic, is of limited analytical value. In particular the dynamic issues of (digital) identities complicate the analysis of identifiable information. More reasonable is thus to explore dimensions that determine the composition of identifiable information with respect to these dynamic issues. Of main interest is what the reference point of identifiable information is and what it describes (e.g., substantial details of a particular person, a spatio-temporal setting, relations or interactions

the person is linked to), how identifiable information occurs, how technology may refer to the individual entity (a person) represented by that information, and how the dynamics of technology may alter this information or create additional one. As a result, four basic dimensions of identifiable information, namely substantial, spatio-temporal, relational and interactional can be detected.

These dimensions can be explained from a system theoretical perspective. Similar to a system, which is dynamic by nature, characterized by the interplay and relations of its elements, also an individual can be understood as a system that consists of some substance or core, has relations to others, and interacts with its environment in different spatio-temporal contexts. The same is given for the technologies and applications that are used to process information about the individual. These items can be seen as sub-systems of a socio-technical system related to the individual. As far as these sub-systems process information related to the individual, for each of these systems, their substantial, spatio-temporal, relational, and interactional information they use or produce is of interest, if it links to the individuals' identity. The benefit of these categories is that they allow considering both types – PII and TII alike as well as the dynamics and relations in-between. The following Sections present options on how to categorize PII and TII with these dimensions. Indeed, given the dynamics of digital information, a distinct, completely unambiguous mapping providing a comprehensive list of identifiable information is hardly feasible. Nevertheless, the proposed typology contributes to facilitate the analysis of identifiable information. It represents a concretization of the identity shadow problem as presented in section 5.3.This typology can also be useful to detect types of information which may be unrecognized yet though privacy relevant.

6.3.1.1.1 Personal identifiability and corresponding identifiable information

As regards PII, these dimensions can be used to structure how a person can be represented by her identifiable information. Briefly speaking, a person has a substantial, unique identity which is involved in multiple different contexts. Identifiable information thus can reveal who a person is, where a person is located at a certain time, with whom a person is related to, what a person does respectively how she interacts and so on. The substantial dimension corresponds with the relatively durable/stable[137] type of identity – idem. The other three types can be more variable, and thus rather refer to the dynamic type of identity – ipse. During her lifetime, an individuals' identity is involved and thus represented in many different spatio-temporal, relational and interactional contexts where she can be identified by information referring to her; e.g., being at a certain place at a certain time; being related to other persons, associated with different organizations etc.; information about individual behaviour and actions etc. In combination, these four dimensions are basic determinants of the representation of an individual identity. Hence, they shape the identifiability of a person.

(1) Substantial includes information about the "substance" of an individual person. Regardless of whether identifiable information is dynamic, every individual person can be

[137] It is described as relatively stable, because also these types of information can change, though rather occasionally and over longer periods of time.

described by a set of information which can be used to represent her unique identity in a certain context. Or in other words: which can be used to substantially identify her. As a member of society, a person is identified in a variety of domains and thus has a social, economic and political identity. These identities are usually represented by some kind of unique identifier to substantially identify the person. The first dimension is thus called substantial and comprises all information that allows for unique identification of an individual person (at least in a certain context). Consequently, basic characteristics about a person but also different types of identifiers fall into this category. This includes information that directly refers to this very person such as (a) *body-specific characteristics* (eye colour, hair colour, height, size, weight, gender etc.), biometric features (e.g., facial and bodily appearance, fingerprint, iris structure, DNA, voice[138]); but also (b) *person-specific* information used to represent one's social, economic or political identity such as full name, social security number, credit card number, passport number, driving license ID, bank account no., customer ID etc.[139] (2) The second dimension is called spatio-temporal. It comprises all information that refers to the spatial and/or temporal occurrence of an individual such as age, date of birth, place of birth, home address, postal code, nationality, ethnic origin, location of living, location of work, current location; (3) The third dimension is called relational and addresses all information about relationships of an individual such as personal status (married, single), employment status, employer, family and relatives, friends, personal contacts and associations etc. (4) The fourth dimension, interactional comprises all information about personal interests, behaviour and actions, communications, expression etc. including sensitive information such as political activities, religious beliefs, sexual preferences etc., resulting from or relevant for interactions with others.

6.3.1.1.2 Technical identifiability and corresponding identifiable information

All the mentioned types of PII can be represented and processed by technical means. The processing of identifiable information by ICTs makes this information reproducible which leads to the emergence of digital identity. Technology can extend or enrich identifiable information in manifold ways. Therefore, a digital identity representation is likely to expand (as shown e.g., in Section 5.3). It could be argued that some of the mentioned types of PII involve technologies as well, such as social security, passport, credit card no. etc. This is true, but this kind of identifiers basically serves formal identification purposes, directly related to the person. Therefore, these forms are assigned to PII and not to TII (although in some practical contexts, a clear distinction may not always be achievable). Technical identifiability and TII address information of virtual nature and/or which have virtual processing contexts. TII typically serves to identify a technical device in the first place, which refers to a person. While identifiable information in the physical or analogue world refers to a kind of physical object that has a matter or substance (a natural person, a document representing this person etc.) this is not necessarily the case in digital environments which process information about virtual representations of the original

[138] Against the background of increasing applications with embedded voice recognition (such as in digital assistents), voice patterns gain in importance.
[139] Although these identifiers involve technology, they serve formal identification purposes directly related to the person. Therefore, these forms are assigned to PII and not to TII.

objects. Hence, the technology (or set of interrelated technologies) applied to process identifiable information can entail the creation of additional identifiable information. For example, a typical online user session may request some kind of information for user authentication. At the same time, it involves at least a computing device (e.g., PC, laptop, smartphone) and a web browser to access a website, service, application etc. Each of these systems bears some kind of identifiable information which can refer to the individual user. Hence, in this example, three technical systems are involved whereas each may provide identifiable information. With the number of sub-systems involved, the amount of TII is likely to increase. This aspect is crucial for the understanding of technical identifiability. The virtual, non-physical processing of identifiable information complicates the conduction of PIA. Metaphorically speaking, every ICT usage can throw an identity shadow which may expand, as shown in Section 5.3.1. Besides PII, also technology-specific identifiable information can be used in various ways for privacy intrusion, e.g., by advanced techniques of de-anonymization and re-identification such as digital fingerprinting. It is thus important to consider these technology-specific types of identifiable information as well. It has to be noted that just as with PII, the following description cannot be a comprehensive list of all types of TII either. The types and amount of TII can vary with technologies, applications etc. But these basic dimensions allow to look from the same analytical lens at different applications, technologies etc. to gain a more detailed picture of identifiable information and its impact on privacy. Against the background of a growth in converging or hybrid technologies, conglomerates of interrelated applications etc. is likely that the complexity of PIA further increases. This typology can be supportive to deal with this complexity.

TII can be categorized with the same four basic types – substantial, spatio-temporal, relational and interactional. (1) Substantial here means identifiable information that originates from those technologies, devices or applications that are primarily involved in the processing of PII. Basically, this includes information applicable to substantially identify an individual based on a technical artefact (an application and/or technical device) she makes use of. A general guiding question to explore this information is e.g.: what kind of technologies and applications (hard- and software) are employed and how they identify a particular user? Typical are (predetermined or generated) unique identifiers. In some cases, it may be useful to distinguish between *(a) application-specific and (b) device-specific* information. Basic examples of *application- (or service-)specific* information are user credentials (usernames, pseudonyms, e-mail address, phone number etc.), as well as particular identification numbers (e.g., Google or Facebook ID, user session ID etc. but also other unique identifiers of a digital identity). Particular cases are external IDM services such as social logins (Section 4.3). They process substantial identifiable information (e.g., a Facebook ID) but originate from and refer to an external entity, i.e., the social media platform they originate from (they are thus also part of relational TII, see below). *Device-specific information* typically includes identifiers of technical devices, (e.g., IP address, MAC address, SIM card ID, IMEI of mobile phone, smart card number, what kind of device is used, whatever identifiers used to establish a user session).

(2) Spatio-temporal means temporal and spatial information about the (primary) usage context of a technology, application or service, e.g., about where and when a service was

used. Typical examples are geo-location, date, time and duration of usage, (timestamps), time zone, last login, duration of user session, date and time of user activity (e.g., postings), time or similar information about when and from which device a person used a particular application etc. Information of this kind may be e.g., stored in log files, usage protocols and the like. Depending on the amount of additional technologies or applications involved in the usage context, various forms of spatio-temporal information may be gathered. These types are described as relational.

(3) Relational basically means information (or metadata) about technologies or applications (respectively sub-systems) that are additionally related to a usage context; either directly or indirectly. Typical examples are the employed computing device, databases and other repositories processing and storing information; or technologies which predetermine an application such a web browser in case of an online service, or integrated social media plugins or logins, or the social graph (see Section 4.3). An example of increasing relevance concerns "*apps*", i.e., micro-programs, typically used to enrich smartphones (but also other computing devices) with additional features. Basically, apps can extend the functionality of a system, and thus its relations to other systems. They may also process PII and TII and share them with external sources (e.g., username, phone no., geo-location etc.); In some contexts in can make sense to differ further between internal and external relations: internal includes all features and applications that are directly involved in a usage context. External may comprise features resulting from external sources or applications with interfaces to other external systems for third party services. Relational TII comprises information available from the related sub-systems. Depending on the number of sub-systems, there can be myriads of ways to gather additional TII and use fingerprinting techniques to create quasi-identifiers.[140] Therefore, configuration details of, e.g., a user's computing device can be assigned to this type, which can be read out to gather identity patterns. For instance, details about the operating system (e.g., type and version), language settings, particular software installations, screen resolution and colour depth, installed fonts, plugins, information about web camera or microphone etc. In case of an online service, a variety of information can be gathered (e.g., http header information[141]), web browser data (e.g., bookmarks, referrer, history of visited sites, configuration, information from cookies, browser settings and user preferences such as cookie settings, adblocker settings, list of fonts, list of installed plugins, storage settings[142]); further examples are metadata of digital objects such as documents, specific settings for image representation (e.g., pictures rendered with HTML), and so on. Even the list of favourite Wi-Fi networks as well as the list of installed apps can be exploited in this regard (see also Section 5.3.1).

(4) Interactional refers to information that occurs during an application context or results from a user interaction. This can be content-specific information, i.e., information that represents the content of a communication or interaction; such as a typical information

[140] There are some awareness raising tools such as "am I unique?" (https://amiunique.org) or https://panopticlick.eff.org which calculate a user's browser fingerprint based on a number of user client information.
[141] Details about http header fields can be found in https://www.w3.org/Protocols/rfc2616/rfc2616-sec14.html
[142] Even information with temporal limits such as stored in the cache can be used for fingerprinting, see e.g., https://kazuho.github.io/http-cache-fingerprint/

occurring in social media usage, ranging from comments, postings, photos, videos or audio, a digital voice pattern, textual messages, e-mail content, contacts and interaction partners, social media content shared links and "*likes*" etc. But also metadata about communications such as information about involved communication parties, time and amount of messages, timely duration of calls or chats, location of the involved parties etc. Moreover, even hardware information e.g., generated while a user interacts with a computing component (e.g., keyboard, mouse, touchpad/screen etc.) can be used to gather unique patterns for fingerprinting[143] to identify a particular user.

6.3.1.2 Discussion

The description of the different types of TII highlights that there are numerous forms of additional identifiable information. ICT usage often involves a conglomerate of many interrelated technologies, entailing enormous complexity. Therefore, the different types of TII can overlap, and a clear-cut distinction is not feasible in many cases. This is particular given in online services, where usually, multiple different technologies and applications are involved. A detection of the exact amounts of TII can thus be enormously challenging. The depth of information gathered by these types strongly depends from the technology or application etc. explored. Also the assignment of information may vary with the primary focus of the analysis. Nevertheless, a basic distinction can be useful to gain a more detailed picture of how identifiability emerges from technology and the socio-technical practices in which they are applied. Not least because it can help to reveal how a person might be identifiable by TII even though she does not directly provide personal information when using a technology etc. This can be supportive to explore what types of identifiable information are necessary and what types could be avoided with respect to data minimisation. The distinction between PII and TII facilitates to reveal what information types emerge directly from an individual's identity attributes and what types result from technologies or applications. For instance, in several PIA models (e.g. CNIL 2015), IP addresses are deemed as personal data; in others (e.g. ICO 2014), they are unconsidered. In each case, it remains rather vague to what extent they are privacy relevant. The suggested typology allows specifying it as a device-specific type of TII and eases its recognition in PIA.

To provide a more practical notion of the typology, the following examples sketch a brief mapping against the types of PII and TII: A common e-mail application may process substantial PII – (a) gender, (b) personal name; spatio-temporal PII – contact details (affiliation, address, phone no. etc., as e.g., provided by the signature); relational PII – associated institution, message receivers/communication partner(s). interactional PII – content of communication; Substantial TII – (a) e-mail account ID, (b) – IP address, MAC address (if access via phone, in addition eventually also IMEI, IMSI); spatio-temporal TII – timestamp, geolocation; relational TII – e-mail header information (e.g., IP address an domain of involved e-mail servers, eventual additional digital fingerprinting info (e.g., cookies, http header information, type of submitted document etc.); interactional TII –

[143] Norte, J., C. (2016): Advanced Tor Browser Fingerprinting., Blogpost, March 6, http://jcarlosnorte.com/security/2016/03/06/advanced-tor-browser-fingerprinting.html

message content, textual structure and patterns (e.g., writing style). Hence, even a typical e-mail application contains a number of identifiable information.

The most prominent example with extensive arrays of PII and TII is social media: substantial PII – (a) gender, photo (facial appearance, eye colour), (b) personal name; spatio-temporal PII - Date of birth/age, place of birth, home address, postal code, nationality, language, location of living, location of work, current location; relational PII – friends, family, personal status, employment status; interactional PII – personal interests, hobbies, education; substantial TII – (a) Facebook username/ID, pseudonyms, e-mail, phone no.; (b) – IP address, MAC address (if access via phone eventually IMEI, SIM ID); spatio-temporal TII – date and time of logins; relational TII – internal: mapping of social networks, groups involved or associated with (social graph), external: social plugins: websites visited and signed in with Facebook ID; interactional TII – content produced, shared or "liked", such as uploaded documents, images etc., posts, comments, search terms, videos watched, music listened, favourite films, music, books etc.

Similar, though much more complex mappings may be gathered e.g., for a smart metering application: substantial PII – (a) gender; (b) personal name of consumer, energy contract and billing details (e.g., customer ID, payment method); spatio-temporal PII – home address, phone number, details about energy demand and power connection; relational PII – number of persons in household; interactional PII – energy consumption behaviour patterns; substantial TII – (a) account ID, eventual user credentials (e.g., username, password) for an online application, (b) – ID of energy device/smart meter, IP address, MAC address (if access via phone eventually IMEI, IMSI); spatio-temporal TII – usage data, e.g., date and time stamps of energy demand, frequency of usage; amount of provided energy and power etc.; relational TII – data and function of individual consumers; interactional TII – Usage patterns (e.g., hours of use, number and type of occupants/devices –how many computers, TVs, washing machines etc.

These examples demonstrate that the amount of PII and (particularly) TII can vary significantly, obviously depending on the assessed technology/application. But there are also several types relevant in each case, when online applications are involved, which may be an indicator for PbD demand.

In practical use, when conducting a PIA, mappings are likely to be less detailed because there is no need to explore all potential types of TII. For PIA, only those types of identifiable information are relevant which are factually gathered and processed by the analysed application or system; used for explicit or implicit identification purposes. For instance, when a user's IP address is not gathered by an application or linked to the user, it is less problematic in this regard. Ideally, with respect to data minimization, the amount of TII processing is kept to a minimum, so that only information inevitably required for the application is gathered. From a wider view, especially the categories of TII may contribute to technology development with respect to PbD and privacy by default mechanisms, as those types of information requiring particular protection may be easier to detect. In this regard, the typology also corresponds with the notions of PIA and PbD as proactive tools to early detect privacy risks and define corresponding protection measures. As it impossible to completely map all types of identifiable information, a basic typology can support to consider (e.g., during development) the extent to which additional identifiable

information may emerge; for instance, from the development of new technologies or the integration of additional sub-systems. Considering further progress of ICTs, recent trends such as "smart" technologies, ambient intelligence etc. the emergence of additional identifiable information and further expanding identifiability is likely. To name just one example, smart home apps may create detailed profiles of their users and their embedded environments (e.g., heating temperature, energy consumption preferences, number of persons in household etc.). As a consequence of these developments, also privacy impacts may further intensify.

The primary aim of the suggested typology is to improve the theoretical understanding of identifiability, which can facilitate the analysis of identifiable information flows crucial for PIA. The practical implementation and applicability strongly depends from the particular system and contexts a PIA is carried out for. A detailed empirical testing is beyond the scope of this research. The practicability of the typology thus needs to be explored in future work and further research may be necessary to sharpen the categories. Nevertheless, there are some basic steps relevant in each PIA process. The next Section provides a general overview on a PIA process based on the presented identifiability-based framework, which takes the proposed typology into account.

6.3.2 Overview on an identifiability-based PIA process

The previous sections presented and discussed a general framework for PIA with an emphasis on identifiability and different types of identifiable information. The presented framework can improve the conceptual and theoretical backing of PIA as it is not limited to specific technologies. The typology of identifiability is an additional value as it allows to shed light on different types of identifiable information, and thus to get a more detailed picture of a privacy impact. This Section briefly outlines how this framework may be practically applied in a prototypical PIA process, sketched as follows:

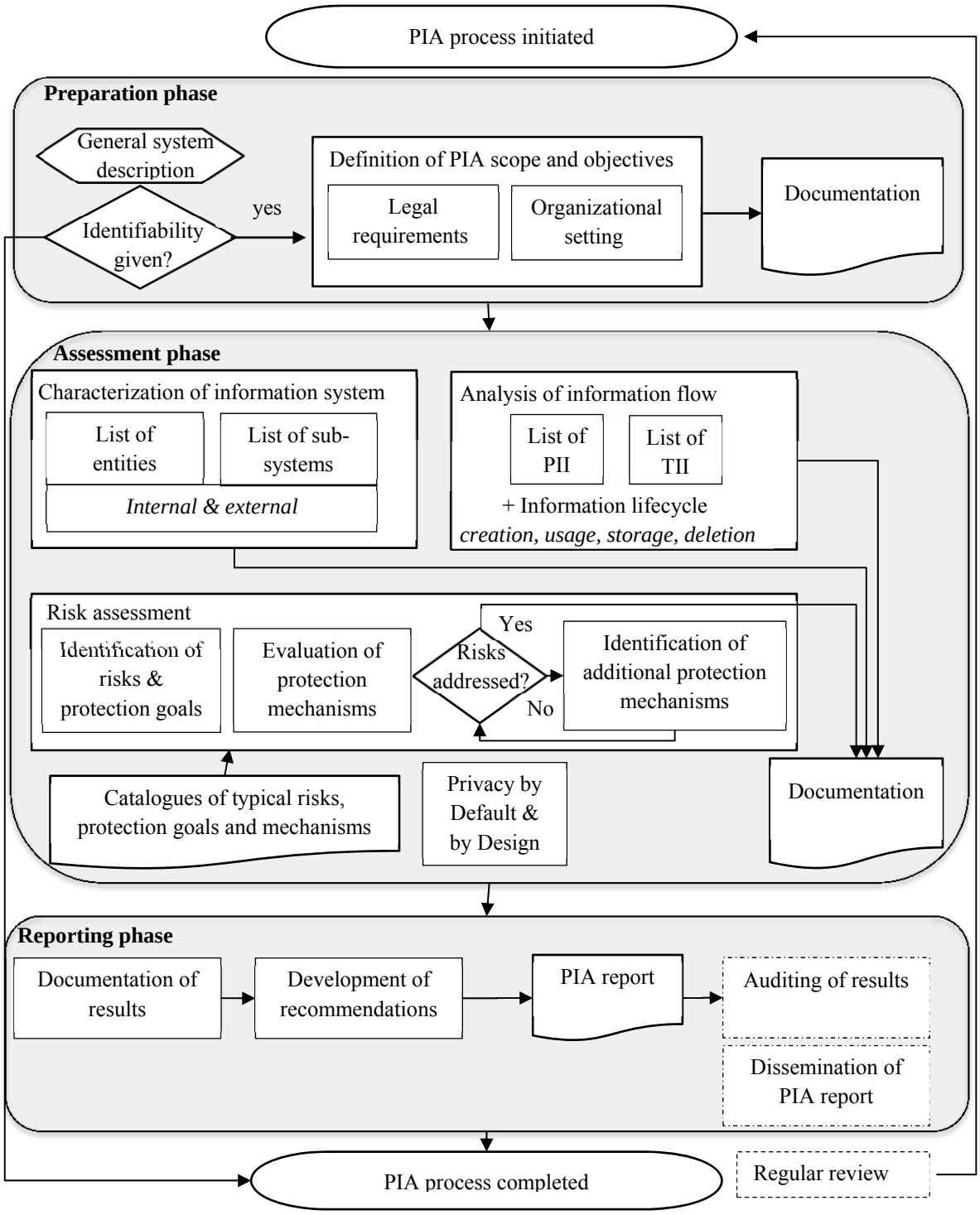


Figure 26: Prototypical PIA process

The basic steps of this process (as illustrated in Figure 26) are informed by existing PIA approaches (as discussed in Section 6.1; e.g. Raab/Wright 2012; ICO 2014; CNIL 2015a; Bieker et al. 2016; EuroPrise 2017) and adapted to the presented general identifiability-based framework. These phases built on each other, although the particular steps in the assessment phase are not necessarily iterative. Some tasks may overlap or complement each other. Documentation is important in each phase to feed the subsequent phases and prepare the PIA report, which is the final outcome of the process.

6.3.2.1 Preparation phase

Initially, the system, application etc. of interest needs to be briefly described including its general function and purpose. This short description is the starting point to determine the general requirements and circumstances for the PIA process. The basis for the PIA can have several reasons, e.g., the assessment of an existing service or product, the development of a new or the improvement of an existing socio-technical system etc. At this stage, it is also useful to know whether a PIA, a privacy audit or a similar procedure was conducted before. Ideally, there is according documentation available such as a previous PIA report, an auditing document etc.

As identifiability is the primary risk, the initial task is to determine whether identifiability is given, i.e., is a person identifiable by any of the information being processed. A guiding question in this regard is for instance: Does the system/application/program etc. process information that refers or relates to an individual person? If there is no clear negation of this question possible, then a general precondition for PIA is given. In the next step, the scope and objectives of the assessment should be defined. First of all, this means to check whether a PIA is legally required (e.g., by the GDPR and/or national laws) or carried out on a voluntary basis.[144] This is followed by a clarification of the organizational setting including the objectives to conduct the PIA such as: ensuring privacy compliance, improving the level of privacy protection, fostering transparency and accountability of the product/service/system etc. Furthermore, it needs to be determined who is responsible for the PIA process (e.g., name and role of team members). Ideally, there is a particular person in the function of a data protection officer entrusted (as e.g., intended in Art. 37 GDPR) with the task and the according resources (such as team members from the IT department, product development and quality management) to conduct a PIA process.

6.3.2.2 Assessment phase

6.3.2.2.1 System characterization

In this main phase, the system that processes identifiable information is described more in detail. This includes a description of the purpose(s) and characteristics of information processing, of the main internal and external entities as well as the sub-systems involved in the processing. The result of this description is an overview on the function and purpose of the system, relevant actors and their roles as well as basic components (e.g., technologies, applications, databases, or software interfaces) that determine how and in what domains identifiable information is processed. This should include information about integrated data repositories (e.g., databases, registers etc.) to outline the general system architecture and processing modalities. Already the number of entities and sub-systems involved can be an indicator for eventual risks of re-contextualization and secondary use (e.g., secondary use might be more likely with a high number of external entities and sub-systems involved). Typical internal entities are the departments that process personal information (e.g., finance, contracting, billing, CRM, product development etc.); typical external entities are service providers or third party contractors with access rights etc. (e.g.,

[144] Detailed guidance for legal requirements with respect to the GDPR can e.g., be found in (EuroPriSe 2017).

marketers, advertising companies, data analysts, security authorities etc.). An entity description may contain e.g., name, short description of function or role (e.g., provider, 3rd party contractor, data controller, data processor[145] as specified in GDPR Art. 4(7 and 8)), relation (internal or external) and purpose of involvement in the information processing system. As privacy and data protection regulation is yet not harmonized internationally, it is also relevant in which country an entity is located and whether EU law is applicable or not. As entities are mostly involved by the provision or use of an integrated sub-system (e.g., a database, an external service, a cloud computing infrastructure etc.) in many cases, entities and sub-systems may be grasped in one list. The table below briefly exemplifies a description of entities involved (on the example of a web service):

Table 5: Example list of entities involved in the information processing

Entity name	***Role***	***Relation (internal/external)***	***Purpose***	***Related sub-system or interface***
Controlling department	Data processor	Internal	Quality assurance	Controlling software
Customer care company	Data processor	External	CRM	Customer database
Facebook	Data controller	External	Identity management, user authentication	Facebook Connect

6.3.2.2.2 Analysis of information flow

A core task of the assessment phase is the analysis of the flow of identifiable information. This task should ideally also include the information lifecycle and the processing modalities from collection/creation, usage, storage to deletion. A useful starting point is a general but not detailed overview on how information is processed and enters the system. The general aim of this system description is to show how the system processes identifiable information: i.e., what are the origins, locations and destinations of the information, to what extent the individual provides this information, which entities are involved in the processing. It is thus important to describe the modalities of identification or authentication; i.e., whether and how an individual is identified. In case of online systems or applications, it is particularly important to consider third party services in this description. For instance, external IDM services such as social logins and social plugins (see Section 4). Is there some kind of IDM approach, e.g., a standardized user registration and login procedure, is this an internal part of the system or an external service (e.g., an integrated social media profile such as of Facebook or Google+), is there a centralized or decentralised user profile etc. The utilization of typical tools such as basic use case descriptions, data flowcharts, workflow diagrams or similar can be supportive to illustrate how personal information flows within the system and its components. Such an overview

[145] Basically, the controller determines the purpose of processing or adds additional purpose, while a processor acts on behalf of the controller. In practice, the distinction between controller and processor can be difficult. For a discussion on these issues, see e.g., (WP29 2010).

is also useful to show the interplay of entities, system components and personal information (process information of this kind is generally important for IT management and thus may be available in at least in larger companies). It should also be checked and described whether a service serves the basic systems' purpose or an additional purpose and whether a purpose is necessary and legitimate; because every additional purpose may cause additional privacy impacts.

This overview can support the main task to reveal and list the different types of identifiable information. This means basically to apply the typology of identifiability by mapping which types of PII and TII are factually processed. This is particularly important as most PIA approaches merely focus on personal information and leave other relevant types of information unregarded (as discussed in Section 6.3.1). Practically, both types of information (PII and TII) are usually gathered in one or several databases or similar repositories. It is thus important to consider these repositories in the analysis. It can be supportive to check the structure of user forms and user interfaces, because these determine e.g., what information is prompted from the user or stored in her data record. For PII and TII alike, it is generally important whether the processing is necessary and legitimate. Particularly relevant is to explore how and for what purpose(s) the information is gathered and stored, whether third parties can use the information etc. As TII often is a general by-product of technology usage it is relevant to grasp how the information is generated and stored as well as whether this is required for the processing context, for technical reasons or avoidable. This is a crucial difference because eventual technical requirements may indicate a demand for privacy by design. The four basic categories of identifiable information (substantial, spatio-temporal/contextual, relational and interactional) and their sub-categories (as described in Section 6.3.1) can be used as reference point to detect the amount of identifiable information. For practical use, PII and TII may be coded as follows:

Table 6: PII and TII categorisation

PII	**TII**
P1: substantial PII	T1: substantial TII
P1.1. body-specific	T1.1. application-specific
P1.2. person-specific	T1.2. device-specific
P2: spatio-temporal PII	T2: spatio-temporal TII
P3: relational PII	T3: relational TII
P4: interactional PII	T4: interactional TII

The codes of PII and TII can then be used to categorize the amount of information. The necessity of a detailed mapping depends from the concrete system being examined. In practice, a strict and detailed mapping may often be difficult and also not necessary for the objective of a PIA. For practical reasons, it can thus be sufficient in many cases to differ between PII and TII without the extra work of a more detailed additional categorization. Nevertheless, these basic categories can be used as guiding questions to explore the composition of identifiable information; e.g., what information is used to substantially identify a person? Is any body-specific or biometric information gathered? What identifiable information is gathered from technical hardware devices, applications, sub-

systems involved? What information is stored that refers to a particular application context? In total, the typology can contribute to gain a more detailed picture of where identifiable information originates from, how it is processed etc. This can also be supportive to identify eventual demand for additional protection mechanisms as well to early detect information with a yet marginal but potentially higher relevance; For instance: Are there any further types of identifiable information being processed which are expected to have a future privacy impact? This can be useful for example, in case of a planned additional system feature (e.g., an integration of biometrics, or an additional technical user device such as a smart watch, a body badge etc.).

The mapping of the amount of identifiable information and its processing modalities is an important indicator for the risks individual privacy may be exposed to. Creating a list of identifiable information is thus crucial for the assessment of risks. The following table exemplifies a possible, practical way to gather PII and TII:

Table 7: Description of identifiable information processing

Description of information			**Source/ processing modality**	**Storage duration**	**3rd party access? secondary use**	**Usage purpose**
Type	**Category**	**Item**				
PII	P1.2	Full name	Provided by user	Unlimited	Yes	Billing, technical requirement
	P2	Date of birth			Yes	
	P2	Address			No	
	P1.2	Social security no.			Yes	
TII	T1.1	User ID no.	Generated by application	Unlimited	No	technical requirement
TII	T1.1	e-mail address	Provided by user	Unlimited	Yes	CRM
TII	T1.1	Username	Provided by user	Unlimited	No	Technical requirement
TII	T1.2	IP address	Automatically gathered from device	Temporary; after session expires	No	undefined
PII	P1.1 (T4)	Facial image	External source (e.g., social media)	unlimited	Yes	CRM
TII	T2	Geo-location	Automatically gathered from application (e.g., web browser)	temporal	No	undefined

The example may be a fictional online shop. The mapping raises some privacy questions, e.g., why the social security number is involved which is probably not necessary for the purpose. As it is shared with third parties, this may be a critical issue regarding legal privacy compliance. The gathering of an IP address may be unproblematic as it is neither permanently stored nor used for a particular purpose. The processing of a facial image can be problematic as it is gathered from an external source as well as shared with third parties. This example also indicates eventual difficulties as regards informed consent as a user may

have accepted the terms of use on Facebook but is not aware, that a completely different service uses her photo for CRM including third party access. Geo-location indicates that this information here may be a sole by-product. A processing could be avoided when the application does not automatically gather this type of information by default.

6.3.2.2.3 Risk assessment

Based on the system characteristics and the analysis of the information flow, risks and protection goals can be identified as well as measures to mitigate these risks. This includes a target-performance comparison, i.e., an evaluation of the existing protection mechanisms and their suitability to address the risks and goals. This phase can be backed by existing standards and catalogues of typical risks and protection goals (cf. BSI 2008; McCallister et al. 2010; ISO 2011; Hansen et al. 2015; Bieker et al. 2016). An integral part of this phase is also an evaluation of eventual additional protection mechanisms, when existing protection is lacking or insufficient to address the risks. Particularly relevant issues are protection mechanisms that provide privacy by default and privacy by design.

It has to be noted that risk here basically means that individual privacy is affected. This goes beyond the legal requirement to conduct PIA, where a risk may be seen as potential violation of legal compliance only. However, as this PIA approach is broader, it is crucial to view on the information processing from the individuals' perspective. As shown in the previous Section, the primary risk of identifiability entails a number of further basic risks and threats, such as unlimited availability, linkability and re-contextualization. General protection goals help to tackle these risks. These goals are to be understood as ideal standard setting to achieve privacy protection. Ideally, anonymity and pseudonymity are provided as far as possible. However, full anonymity is often not feasible without challenging the main purpose of information processing. Therefore, the protection goals aim at minimizing the risks by providing basic conditions for secure information processing so that identifiable information is largely protected from misuse. Although there is a number of basic risks and protection goals that are of general relevance, each information system differs in functioning and purpose. Hence, the implementation of the risk assessment process, the effectiveness of protection mechanisms and eventual need for additional safeguards strongly depends from the system and its processes as a whole. It is thus important that the risk assessment takes the system characterization and the flow of identifiable information into account. The list of PII and TII and the processing and storage modalities provide several indications for privacy risks and threats. For instance, the risk of unlimited availability is shaped by storage modalities. The longer identifiable information is available and stored, the higher the according risk. Limited storage and retention duration thus contribute to reduce risks of this kind. A related issue concerns access to information and processing modalities. A centralized database containing a full record of personal information bears a higher risk than de-centralized storage concept with separated pieces of information. The direct use of identifiers in multiple datasets (e.g., in separated databases etc.) amplifies the risk of linkability and data aggregation across multiple contexts. Secondary use and access to information by third parties affect the risk of re-contextualisation. All these aspects are relevant to identify the risks and protection mechanisms. General guiding questions are

e.g.: What kind of risks is identifiable information exposed to? How and with what protection mechanisms are these risks addressed? Is the processing of identifiable information in accordance with common privacy and data protection principles (data minimization, purpose limitation etc.)? What are the existing protection mechanisms and how suitable are they to mitigate the risks? How is data protected from illegal or unintended access? Is privacy by design considered in the information processing context and in what form? Is identifiable information encrypted, pseudonymised, anonymised, or deleted? etc.

There are many options to assess the severity of risks, e.g., with scales from high, medium to low or similar. A further option is to assess the protection level related to a risk e.g., with categories (4) high, (3) appropriate, (2) sufficient, (1) insufficient, (0) missing; as shown in the example below. This mapping can indicate to what extent the protection mechanisms contribute to mitigate the risks. The results of the risk assessment indicate eventual need for the implementation of additional safeguards. A simple example to compare risks and controls may look as follows:

Table 8: Example description of risks and protection mechanisms

Risk type	Description of risk	Protection mechanism	Current protection level
Linkability	Identifiers are directly used and refer to full data records also to external entities	Unlinkability, pseudonymization, encryption	3
Durability	Data access is not restricted	Access management	2
Durability	Data storage is unlimited	-	0
Aggregation	Data of multiple sources is aggregated and centrally stored	Anonymization	4
Re-contextualization/ purpose extension	Secondary use, e.g., TII to track users geo-location without informed consent	Informed consent or usage limitation	1
Traceability	Individual user behaviour is monitored for profiling activity	-	0

The concrete realization of a risk assessment procedure strongly depends from the scope and aims of the PIA. A standard application with a low amount of PII and TII has different requirements than a large scale application that processes sensitive information (e.g., in the health sector). But in any case, for the evaluation of protection mechanisms it is vital to create use cases. The involvement of legal as well as IT experts is important in this regard. Moreover, to avoid organizational blindness, including standard users can be vital to gain the perspective of an individual concerned from privacy intrusion. A potential side-effect of user involvement is a potential usability evaluation, which can be supportive for service provision.

6.3.2.3 Reporting phase

Finally, all assessment results are documented in the PIA report. This report ideally also provides recommendations on risk mitigation and improvement or implementation of

protection mechanisms. For quality insurance, an optional auditing to evaluate the PIA report by an independent authority (e.g., DPA or external data protection officer) can be useful to detect and handle eventual conflicting interests and facilitate the implementation of protection mechanisms. This audit can also be linked to a certification procedure such as the European privacy seal[146]. A public dissemination of the PIA report contributes to improve accountability and transparency of information processing, enables public scrutiny, and may be supportive for reputation management. Depending on the particular function of the PIA, this document serves as reference guide for privacy compliance, a discussion paper for development (e.g., to improve product or service quality and security by integrating privacy by design) as well as input information for a further PIA process. Ideally, a PIA process is continuously revised and conducted in a defined period of time, e.g., every 5 years or if the system or its purpose has significantly changed by, e.g., new features or technologies.

[146] https://www.european-privacy-seal.eu

CHAPTER 7

Summary and conclusions

This research shed light on the complex interplay between privacy, identity and (digital) identification from a systemic perspective. A system-theoretical approach was applied as a research heuristic to grasp the socio-technical transformations related to ICTs that shape this interplay. The analysis focussed on the role of identification, i.e., the processing of identifiable information which is a basic condition for the emergence of a privacy impact. This general nexus is given in analogue as well as in digital contexts. Already the possibility to process information that refers or relates to the identity of a person can entail privacy risks. Therefore, as argued, socio-technical identifiability is a crucial determinant of a privacy impact and lacking control thereof is a core problem of contemporary privacy protection. This basic issue of privacy aggravates as technology development and socio-technical practices significantly extended the degree of identifiability. A stronger consideration of identifiability is particularly important against the background of further informatisation and growing amounts of digitally networked environments.

As shown, identification practices altered with technology which also has an effect on the scope and functioning of privacy. Given the peculiarities of ICTs and essentially, of digital information, technology usage extends the representation of personal identities. These developments increasingly challenge traditional notions of privacy and of personal (identity) information. In contrast to analogue forms of identification, digital identification can comprise multiple dynamic contexts in which identifiable information is being processed. These multiple contexts or (identity) layers are often beyond the control of the individual, exposed to privacy intrusion. Hence, there are myriads of ways to gather, re-contextualise or reproduce personal information as well as options to use other types of information for identification as well. The already very limited options for individuals to control the processing of their information can further erode. As a consequence, threats to privacy can aggravate and the effectiveness of protection mechanisms further decreases. This is not merely the result of technology but of usage practices as well as of the complexity of privacy as such. Apparently, privacy is a relatively abstract concept with various roles and meanings, which is one reason for complications as regards its protection. Therefore, it is relevant to reconsider what protecting privacy essentially means in order to achieve appropriate levels of protection with respect to socio-technical change.

7.1 Privacy versus … – a contradiction in function?

As shown in Section 3, privacy, identity and identification – are intrinsically linked. As a fundamental human right, privacy has an inherent boundary control function, regulating the interplay between private and public spheres. Privacy protection includes regulating

informational relations and boundaries between individuals and other entities. Ideally, privacy provides a domain (one's private sphere) in which the individual can act self-determined and free from interference. Privacy thus represents a constitutive framework for autonomy enabling self-determination, vital for identity development (Section 3.2.3). At the same time, identity constitutes the private sphere of a particular individual, because otherwise, privacy has no subject or benchmark. In the same manner, identity is the socio-technical construct shaping the interactions of an individual with others and its environment. Privacy enables self-determined, free action and participation of individuals in society. Therefore, privacy is an enabler of other rights such as freedom of expression and thought, of movement, association etc. Hence, the private and the public sphere are no opponents but they complement each other. Privacy enables to regulate the interactions between individual identities and society. This is not a decoupling of the individual from its environment but contributes to a self-determined involvement in society. Consequently, privacy is not merely a private but also a public value essential for democratic processes.
However, there are certain tensions with partially conflicting concepts which complicate the protection of privacy and challenge its public value (Section 3.3). Particular tensions result from the misleading trade-off between privacy and security. Further controversies exist between privacy and transparency, whereas notions of post-privacy question the necessity of privacy due to increasing technical and societal transparency. In these controversies, privacy is framed as concept in contradiction to security as well as transparency. Narrow framings of privacy assumed to be in permanent conflict with these concepts jeopardise the public value of privacy. In fact, there is no permanent contradiction in each case. As argued (in Section 3.3.1 and 5.2.2), the assumed perpetual trade-off between privacy and security is a common fallacy which reinforces securitisation and privacy-intrusive surveillance practices supported by corresponding technologies. Indeed, privacy intrusions are foreseen by the law; however, as exceptional options to protect democratic principles in the interest of the public, but not as permanent necessity as suggested by the trade-off. In the trade-off logic, security is falsely presented as a dominant value frequently endangered by privacy. This misleadingly justifies a reinforcement of privacy-intrusive security and surveillance practices in the sense of a proceeding securitisation. Metaphorically speaking, a constructed security continuum generates an expanding privacy vacuum reinforced by technology. To some extent, the rationales of post-privacy and of securitisation overlap: both misleadingly reduce privacy as a form of secrecy aiming at hiding information. As argued, this logic is a fallacy because privacy comprises more than personal secrecy or confidentiality. Privacy is not least a public value requiring contextual integrity, so that personal information is processed for a particular purpose in accordance with legal and ethical norms. This refers to responsible handling of personal information by the processing entities. However, a reductionist framing of privacy as secrecy widely neglects this responsibility. In fact, security and transparency of information processing are major requirements of privacy protection. Accordingly, opacity undermines accountability and scrutiny of (personal) information processing and thus effective privacy protection.

The boundaries privacy regulates are basically informational ones. Informational self-determination (ISD) is thus an essential concept of the boundary control function of

privacy. ISD is the individuals' capacity to control the processing of information about her. It contributes to self-determined maintenance, disclosure and performance of personal identity. For ISD to be viable, the individual needs to know, e.g., what personal information is collected, stored and processed for what purpose, and by whom. Hence, ISD implies transparency and control of information processing. However, the informational boundaries privacy protects and thus ISD are heavily under pressure due to socio-technical practices amplified by the dynamics of ICTs.

There is a functional difference between privacy protection and identification which can be critical in this regard: privacy implies the existence of informational frictions so that personal information is ideally not disclosed or accessible to others without intention and control of the individual concerned. Identification implies to establish an informational link between different entities and thus to cross informational boundaries. Digital technology and identification practices foster seamless information flows and complicate the provision of self-controlled informational frictions, i.e., unlinkability (being a crucial concept of technical privacy protection). Moreover, these dynamics of ICTs and digital identification benefit the dynamics of securitisation which reinforce privacy intrusions. While our identities become increasingly transparent in socio-technical systems, their usage purposes, i.e., the processing of identity information, is increasingly opaque.

As shown in Section 4, digital identification emerged within a wider socio-technical transition, including various transformations in social, economic, political and technical domains. In a relatively short period of time, ICTs became increasingly interactive, interconnected and deeply integrated in society. They are not merely tools of information and communication anymore but deeply embedded *in* and have substantial impact *on* societal structures. Visions of pervasive computing and similar ideas suggesting hyper connectivity became more concrete in recent years. Analogue and digital environments increasingly converge, whereas ICTs represent socio-technical artefacts connecting both worlds. They foster networking structures, connectivity and further growth in digital information processing. This includes extended representations of our digital identities and changing identification practices. Identification basically has a connecting function enabling links between different entities serving social, economic, political as well as technical purposes. This connecting function is embedded in and reinforced by these developments in many respects. Different forms of identification are involved to establish and maintain these socio-technical networking structures. Socio-technical systems generally include technical identification mechanisms as two or more entities require some processing of identifiable information about each other to establish a connection. Consequently, individuals and institutions are increasingly networked by the technologies which process their information. The growth in networking structures also affects the handling of digital identities: formerly rather isolated user profiles and other forms of identity representations of different socio-applications or socio-technical systems are increasingly networked as well. Social media platforms prominently highlight how interactive and interconnected online identities became. They are thus a blueprint for the networking dynamics of digital identities.

These socio-technical transformations entail different modes of personal as well as technical identification and boost the amount of identity information. To deal with the

growing complexity and foster control of digital information processing, concepts of IDM gained in importance (Section 4.2). Basic aims include improving efficiency and security of identification processes to handle digital identities. IDM is widespread with different forms being integrated in ICTs and online services serving various purposes e.g., to conduct online transactions, to provide services in e-commerce and e-government, to manage user profiles of online platforms etc. Hence, IDM affects the relationship between individuals and institutions in the public and private sector. IDM, digital identity representations, personalisation and thus different forms of identification generally increase entailing network dynamics. Online platforms providing social plugins and logins highlight how far-reaching digital identity information can be cross-linked and aggregated over multiple application contexts.

While technological progress triggered a general demand for IDM, its implementation is mainly driven by a number of interrelated economic and political interests. Policy makers in Europe as well as in other countries highlight IDM as a tool of governance to improve administrative procedures and especially to stimulate the digital economy. Digital identification serves a variety of economic purposes including e.g., service efficiency, personalisation, CRM, targeted advertising, profiling as well as service-for-profile business models. Social media platforms highlight the commercial exploitation of identity information which has an enormous economic value. Furthermore, IDM is closely related to a number of security objectives ranging from securing online services, issues of cyber security, fighting identity fraud, crime and terrorism and thus national security. Hence, regimes of the digital economy and of the security domain are strong drivers of digital identification. The trend of a further expansion of digital identification purposes results from a complex interplay of technological, economic and socio-political factors: ICTs generally extend the representation of our (digital) identities, reinforced by a convergence between analogue and digital environments. Social media is a prominent showcase for this expansion of digital identities serving various commercial interests (Section 4.3). But also besides social media, identity information is used for a number of economic, political and security purposes. Initially, IDM was used for formal identification to conduct e-transactions: However, its scope extended with ICT diffusion and usage. Today, formal and informal, explicit and implicit identification overlap in many respects. Trends to further expand identification, such as plans to integrate social media profiles into formal identification procedures (e.g., for national security purposes such as border control or law enforcement) highlight that digital identities increasingly enter “real world” contexts, closely related to governance and control.

7.2 The privacy control dilemma and the quest for (identification) control

Control over (digital) identity information is a crucial issue for privacy as well as for identification, though for different reasons. Protecting privacy aims at shielding individuals from unintended and uncontrolled identification. This implies protecting information

which directly or indirectly represents the identities of individuals. Identification includes the processing of this information to determine the identities of individuals distinct from others. Uncontrolled processing of digital information challenges privacy protection as well as identification. The increasing importance of IDM can be seen as an attempt to regain control over digital information, mainly to improve security of the latter. However, as argued in Section 5, this can lead to a further loss of control from a privacy perspective. This is particularly the case, when IDM is designed without privacy features. There is thus a certain privacy control dilemma of digital identification.

(Digital) identification serves many vital societal functions, is a basic instrument of social, economic and political governance. But it also represents a control mechanism. In general, the striving for control involves a quest for security and stability of a matter. This applies to political and administrative power, national security as well as economic growth. Identification is a means towards this quest with the basic aim to reduce uncertainty and improve security in particular settings by gaining knowledge about individuals' identities. Compared to security and economic interests, privacy protection plays a rather marginal role in the implementation and usage of IDM or related digital identification practices. Given the strong influence of the digital economy and the security domain, including their self-dynamics, we can speak of economization and securitization of digital identification (Section 5.1.2). On the one hand, several empirical examples highlight that identity information is treated as a valuable economic factor for digital markets. On the other hand, identity information feeds into a wide array of security and surveillance practices driven by the logic of securitization: digital identification is also framed as a tool of security governance increasingly linked to forms of preventive risk detection in various contexts. Economic and political actors of surveillance here often overlap, together shaping the surveillant assemblage. The Snowden files bear prominent examples for this complex interplay of private enterprises and security authorities gathering identity information from ICTs. However, the nexus between surveillance and digital identification is not limited to this case. Irrespective of the usage purposes in particular, the various examples ranging from commercial exploitation to different forms of surveillance underline that the broad availability of digital identity information stimulates desires to use this information. Hence, the basically unlimited options to gather this information intensifies risks of function and mission creep, i.e., the incremental extension of usage purposes. The often indistinct mix of economic and security objectives digital identification practices relate to, underlines this aspect.

Hence, to some extent, there is a rather thin line between surveillance and (digital) identification (as argued in Section 5.1). Apparently, its close relationship with surveillance does not imply that identification is a means of surveillance and control in any case. As argued, in brief, surveillance is privacy-intrusive when it involves the processing of identifiable information. Identification is privacy-intrusive when it breaches legal or ethical privacy norms such as a violation of contextual integrity. Also the processing of identity information beyond control of the individual concerned affects her privacy. The crux of identification lies in imbalanced control over identity information and lacking ISD, which hampers privacy protection. In this regard, there are certain overlaps between panopticism and identification observable as regards their basic functioning (Section

5.1.3). Considering these overlaps is also relevant to understand the modalities of "new" technology-aided forms surveillance as discussed in the field of surveillance studies. A central functionality of panoptic power (mostly inherent to surveillance) is the creation and maintenance of information asymmetries for the benefit of the entity exercising this power. Also identification can create information asymmetries, when the individual person lacks control over being identified and has little or no knowledge about the use of her information. Identity information is generally used by a variety of different actors for various reasonable purposes ranging from fostering administrative procedures, stimulating economic development as well as for political objectives. In many cases, the processing of identity information is primarily controlled by institutional/organizational actors (e.g., public sector institutions, security authorities, social media providers, businesses etc.). This particularly the case in digital environments as individuals often lack in control over the technologies and systems processing their information, provided or employed by institutional entities.

Thus at its core, the privacy control dilemma is determined by information asymmetries and agency problems resulting from imbalanced control over identity information. The perceptions of citizens on privacy, security and surveillance (as presented and discussed in Section 5.2) further confirm that extensive institutional power and information asymmetries challenge privacy. More precisely, the processing of identifiable information beyond individual control is a basic trigger of privacy-affecting information asymmetries. Agency problems hamper ISD and bear the risk of moral hazard. In a privacy context this means that information is used at the cost of the individual. As argued in Section 5.1.3, information asymmetries concerning digital identities can entail many forms of social control, discrimination as well as manipulation. Plans such as the so-called citizen score in China based on the national identification system are a drastic example for the misuse of identity information for panoptic forms of power. In line with the risk of automated power inherent to the panopticon, there is a certain risk that individual identities are reduced to a quantifiable pattern of information. Semi-automated algorithms processing identity patterns for profiling, price discrimination, scoring, risk calculation etc. demonstrate that this risk is not merely theoretical, also in Western countries. There is already evidence for algorithms reinforcing social disparities including stereotyping and social sorting and the big data paradigm supports tendencies to extend scoring and other forms of automated risk calculation.

Given the dynamic features of digital identity representations and the incremental extension of identification practices, information asymmetries can increase. As identity likely grows over time, it has narrative characteristics which are naturally volatile in analogue environments. However, technology reduces this volatility as identity information and the aggregation thereof can make the narrative of identity explicitly visible and reproducible. Consequently, digital identity representations can be reproduced and processed in multiple contexts, decoupled from its originating individual persons. ICTs demonstrate this in many respects with social media platforms as prominent showcase. While digital identification has various benefits and serves legitimate purposes of governance, there are trends of extending explicit and implicit forms of identification which are critical from a privacy perspective. There is thus a general increase in

identifiability which reinforces privacy-affecting information asymmetries. Therefore, uncontrolled socio-technical identifiability was identified a core issue of contemporary privacy protection (Section 5.3). As demonstrated with the identity shadow, digital identities can be exposed to several uncontrolled contexts of information processing. Hence, besides explicit forms of identification, technology usage generates various types of information suitable for implicit identification. Explicit forms of identification can entail a quasi-obligation for individuals to provide identity information; implicit forms identification benefit from the broad availability of digital information referring to individuals. The basic design of ICTs facilitates this by identifiability by default mechanism which provides various ways to gather quasi-identifiers and create digital "fingerprints". This type of information is inter alia used for large-scale profiling and the creation of identity graphs. The crux is that digital information processing can entail multiple contextual identity layers. Moreover, there are several trends of expanding identifiability (Section 5.3.3) as basically, every technology usage can lead to an extended representation of an individuals' identity. The increasing use of biometric technologies including fingerprint scanners and facial recognition systems exemplifies that diminishing informational boundaries even affect privacy of the human body. Technological progress makes a further expansion of the identity shadow and thus of identifiability very likely.

7.3 Revitalising the public value of privacy with PbD and PIA

To revitalise privacy protection and tackle the problem of expanding socio-technical identifiability requires actions at several fronts. Among other things, there is need to enhance ISD and individual privacy controls as well as more effective safeguards, implemented by information processing entities. Fostering the combination of PbD and PIA is vital to raise the effectiveness of privacy protection in total. Both concepts can also stimulate economic incentives to protect privacy which is, besides mass surveillance and other things, yet among the core barriers to effective protection. The GDPR is an important step stone to strengthen the European privacy regime in this regard. To unfold its potential requires effective implementation of PbD and PIA.

As argued, improving privacy protection implies means to compensate information asymmetries resulting from socio-technical identifiability. The fact that identifiability triggers privacy impacts does not imply that every form of identification is privacy intrusive or harmful. Identification practices are vital for the functioning of society. But an explicit consideration of identifiability when assessing privacy impacts is fruitful to improve the theoretical understanding of privacy protection as well to improve the effectiveness of safeguards. Section 5.4 explored the prospects and perils of privacy by design (PbD) in this regard. The analysis revealed that there are several technical concepts to improve the protection of identifiable information and thus of privacy: encryption ever was a backbone of technical privacy protection ranging from content encryption, providing unlinkability, to different anonymisation techniques; novel approaches such as differential privacy, or concepts employing blockchain technology (e.g., Etherum) to foster

decentralised information processing can improve PbD, e.g., in the realm of pervasive computing. From a systemic perspective, PbD represents an approach to foster the boundary control function of privacy. It aims at providing mechanisms to create informational frictions so that different socio-technical systems can be decoupled with respect to privacy protection. Encryption technology is an important means to achieve this. Basically, there are several promising approaches for PbD. Even IDM can contribute to enhance privacy when it provides unlinkability and features for anonymous and pseudonymous usage (which is actually not the case).

However, there are several barriers to the effective implementation of PbD. Firstly, there are problems of so-called informed consent which often enforces individuals to fully accept the privacy conditions of e.g., a service. PbD is then of limited effect. Secondly, the effectiveness of PbD suffers from a still relatively high complexity of technical tools (PETs) which complicates individual handling and thus ISD. Privacy controls are often limited in scope as only skilled users can properly handle them. In this regard, there are certain privacy divides between skilled and standard users. Furthermore, there are tendencies to discriminate users of privacy tools (e.g., by blocking access to online services or complicating registration without real ID). Thirdly, barriers of PbD result from a certain tendency of what can be called privatisation of privacy which shifts the responsibility to protect privacy mainly to the individual. As a consequence of a further privatisation of privacy, it then becomes are sort of luxury good while institutional responsibility is neglected. As an individual can hardly control all processing contexts of her identity information, privacy protection reaches its limit. Therefore, PbD aiming at enhancing individual privacy controls is essential but not enough to relativize the privacy control dilemma and improve the level of protection. Hence, also here, agency problems become apparent. To reduce these problems and foster PbD requires a revitalisation of the public value of privacy. This implies a shared responsibility between individuals and institutions. For PbD this means that the processing of identifiable information is to be avoided as far as possible. However, technical solutions are not enough to ease this problem. There is thus demand for regulatory, organisational, and technical measures to reduce privacy-affecting information asymmetries. This requires more transparency and accountability of information processing entities to achieve a shared responsibility regarding privacy protection. Basically, the GDPR is a promising regulatory approach in this regard as it fosters the relevance of PbD as well as of privacy impact assessment (PIA). This new regulation partially fulfils the claim of privacy advocates to make PbD and PIA a legal requirement. But apparently, the law cannot give procedural guidance on their implementation. To effectively implement PbD requires knowledge about the functioning of information processes including the involved technologies and applications. PIA is essential in this regard as it can support institutions in identifying their demand and requirements to implement PbD. In the long term, PIA and PbD can complement and reinforce each other, leading to more effective privacy standards.

As PIA is an important precondition for PbD, its functions and scope were examined in Chapter 6 in order to develop a refined approach. An evaluation of existing PIA concepts revealed that they are often either tailored to a specific issue or offer rather general organisational steps with limited guidance to explore privacy impacts; some focus

more on legal compliance, others more on risk assessment. This great diversity of PIA concepts is partially plausible as a PIA process needs to be specified with respect to particular institutional settings. However, a main barrier to the effective implementation of PIA results from a lack of common understanding of the emergence of a privacy impact. Consequently, the implementation of appropriate privacy safeguards is hampered as well. As argued in this research, identifiability is a core determinant of a privacy impact. Therefore, a general, identifiability-based PIA framework was proposed (Section 6.3). This framework can support the analysis of those information flows which are relevant to grasp privacy risks and develop corresponding protection.

The framework has no legal focus, although it can be supportive for compliance checks as well, as this requires knowledge about the processing of identifiable information by all means. In this framework, identifiability represents the initial privacy risk from which further risks can emerge (i.e., durability, linkability, traceability, re-contextualisation and aggregation of information). These risks can then be addressed by basic protection goals (anonymity, unlinkability, contextual integrity, confidentiality, interveneability, transparency) and suitable PbD approaches. To grasp privacy risks requires a deeper understanding of the amount of identifiable information being processed. Privacy intrusion due to ICTs is not limited to personal information anymore. Technology altered the role and generation of personal information so that the boundary between personal and non-personal, or personally and technically identifiable information blurs. Neglecting this fact can significantly hamper privacy protection. Existing privacy standards often focus on personally identifiable information (PII) only and rather neglect what I call technically identifiable information (TII).

As highlighted with the identity shadow, technology offers many options of implicit identification (including de-anonymisation, re-identification etc.) based on information which is not perceived as PII. Hence, there is a growing demand for PIA as well as PbD concepts incorporating TII. Contemporary privacy protection requires a deeper, process-oriented understanding of (digitally networked) information. Therefore, an alternative typology of identifiable information (Section 6.3.1) was suggested, which explicitly takes PII as well as TII into account. This can improve PIA of a particular technology or application. The typology is based on four basic dimensions of identifiable information: substantial, spatio-temporal, relational and interactional. These layers allow considering not merely PII but also TII when analysing privacy-relevant information flows. The rationale of these dimensions is that information processing can involve multiple socio-technical (sub-)systems. Awareness and knowledge about these systems and their dynamics is relevant to assess and reduce the risks of implicit identification (e.g., for unintended third party access, hidden profiling etc.).

In the light of the GDPR, PIA can be expected to gain in importance within the next years. Although PIA is only mandatory under certain conditions (as regulated in Art. 35 GDPR), public and private institutions have to evaluate and document their privacy-relevant information processes in order to act in compliance with the law and to avoid penalties. A PIA approach enabling more knowledge about identifiable information has several benefits in this regard: it contributes to improve transparency, accountability and legitimacy of information processing which supports institutions in providing privacy

compliance; it helps to respect the privacy principles of data minimization and purpose binding, because it eases to evaluate what types of information are necessary for a particular purpose and what types can be avoided; processing entities can proactively protect their information and improve security of their processes; the implementation of PbD is fostered as more effective safeguards can be developed which contributes to improve information security as well as raise the general level of privacy protection. This is also relevant in the light of increasing security threats of cyber-attacks which demonstrate the vulnerability of socio-technical systems. More transparency and protection of identifiable information also corresponds with the partial complementarity between privacy and security.

The proposed framework is primarily a contribution to improve the theoretical understanding of privacy impacts and their assessment, which is of practical relevance as well. Accordingly, Section 6.3.2. sketched a prototypical PIA process incorporating major steps of the proposed framework. The typology of identifiable information is an attempt to support a more systematic analysis of privacy-relevant information processes, which is an integral part of PIA. Given the high complexity of digital information processing, it can be challenging to detect the different types of PII and TII. A clear assignment may not always be achievable in practice. The practical employment of this typology may thus require refinement or simplification. Nevertheless, a more systematic incorporation of PII and TII supports to raise awareness and gain a more detailed picture of privacy-relevant types of information. This can support the development of PbD and related technical privacy concepts, ideally in cooperation between technology vendors, providers and operators. An additional value can be to detect hidden impacts inherent to technology which may imply security risks as well. In the longer run, there is thus potential to stimulate the creation of better standards for PIA and PbD.

The proposed framework is only a small contribution to tackle some of the various challenges privacy protection encounters. Further research is needed to evaluate and test the practicability of the proposed framework and the typology of identifiable information. Besides issues of transparency and accountability (as addressed with the PIA framework), there are several other issues such as remaining problems of informed consent, third party usage of identifiable information, general increase in biometric identification including facial recognition, extensive profiling and preventive surveillance practices etc. to name just a few. Not least, the outlined problem of an increasing privatisation of privacy needs a wider paradigm shift to revitalise the public value of privacy beyond PIA.

Although TII is mostly not of legal relevance yet, a stronger, systematic consideration of TII is particularly important in the light of further technological progress entailing further digital networking where visions of pervasive computing take more concrete shape. Furthermore, considering an incremental increase in (semi-)automated systems, machine learning algorithms etc. additional privacy problems concerning the processing of identifiable information by machine entities can be expected. Further challenges on privacy and increasingly also human autonomy are likely. Already today, there are algorithms capable of (semi-)autonomous identification based on a set of identity criteria for scoring, profiling, surveillance etc. As a consequence, information about human identities may be increasingly processed by technological agents on an automated basis. This bears further

risks of social sorting, discrimination as well as conflicts between (semi-)autonomous systems and human autonomy; especially when decisions based on automatically gathered identity information affect the individual. Regulation to prevent from automated decision-making is already strained and it is an open question whether it offers sufficient protection. Hence, there is a number of issues suggesting further research. To ensure that privacy protection has a stable, future proof foundation not least requires a solid privacy regime including a mix of regulatory, technical and political measures to rebalance the relationship between privacy and security with respect to liberty being their defining value. Because while the fundamental role privacy fulfils in society is of enduring value, its continuity is at stake when protection mechanisms lack effectiveness to deal with socio-technical practices.

Bibliography

All URLs in this document were last accessed on July 23 2017.

Abelson, H., Lessig, L. (1998): Digital identity in cyberspace. White Paper Submitted for *6.805/Law of Cyberspace: Social Protocols*. Available online at: http://groups.csail.mit.edu/mac/classes/6.805/student-papers/fall98-papers/identity/linked-white-paper.html

Ackerman, S., Thielman, S. (2016): US intelligence chief: we might use the internet of things to spy on you. *The Guardian*, February 9, http://www.theguardian.com/technology/2016/feb/09/internet-of-things-smart-home-devices-government-surveillance-james-clapper

Ackoff, R. L. (1971): Towards a system of systems concept. *Management Science,* 17(11): 661–671. Available online at: http://ackoffcenter.blogs.com/ackoff_center_weblog/files/ackoffsystemofsystems.pdf

Acquisti, A. Gross, R. (2006): Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook. In Danezis, G., Golle, P. (eds.), Privacy Enhancing Technologies. PET 2006. LNCS Vol. 4258. Berlin/Heidelberg: Springer: pages 36-58.

Acquisti, A., Gross, R., Stutzman, F. (2014): Face Recognition and Privacy in the Age of Augmented Reality. *Journal of Privacy and Confidentiality*, 6(2): 1-20. Available online at: http://repository.cmu.edu/jpc/vol6/iss2/1

Acquisti, A., Taylor, C., Wagman, L. (2016): The Economics of Privacy. *Journal of Economic Literature*, 54(2): 442-492. Available online at: https://www.aeaweb.org/articles?id=10.1257/jel.54.2.442

Agre, P.E., Rotenberg, M. (eds.) (1998): *Technology and Privacy: the new landscape*. Cambridge: MIT Press.

AI – Amnesty International (2016): For your eyes only? Ranking 11 technology companies on encryption and human rights. London: Amnesty International. Available online at: https://www.amnesty.org/en/documents/POL40/4985/2016/en/

Aichholzer, G., Strauß, S. (2010a): Electronic Identity Management in e-Government 2.0: Exploring a System Innovation exemplified by Austria. *Information Polity – An International Journal of Government and Democracy in the Information Age*, 15(1-2): 139-152.

Aichholzer, G., Strauß, S. (2010b): The Austrian case: multi-card concept and the relationship between citizen ID and social security cards. *Identity in the Information Society (IDIS)*, (3)1: 65-85.

Ajana, B. (2013): *Governing through biometrics. The biopolitics of identity*. Palgrave Mcmillan: Basingstoke/New York.

Akerlof, G. (1970): The market for 'lemons': Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3): 488-500. Available online at: http://www.econ.yale.edu/~dirkb/teach/pdf/akerlof/themarketforlemons.pdf

Alba, D. (2016): Pentagon Taps Eric Schmidt to Make Itself More Google-ish. *Wired*, February 3, https://www.wired.com/2016/03/ex-google-ceo-eric-schmidt-head-pentagon-innovation-board/

Albrechtslund, A. (2008): Online social networking as participatory surveillance. *First Monday*, 13(3): March. Available online at: http://firstmonday.org/article/view/2142/1949

Allkott, H., Gentzkow, M. (2017): Social Media and Fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2): 211-236. Available online at: https://web.stanford.edu/~gentzkow/research/fakenews.pdf

Altman, I. (1975): *the environment and social behavior: privacy, personal space, territory, crowding*. Montery: Brooks/Cole.

Amoore, L. (2008): Governing by identity. In Bennett, C. J., Lyon, D. (eds.), *Playing the identity card - surveillance, security and identification in global perspective*, London/New York: Routledge: pages 21-36.

Angwin, J., Parris Jr., T. (2016): Facebook Lets Advertisers Exclude Users by Race. *ProPublica*, October 28, https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race

Annan, Kofi (2000): Secretary-General Salutes. *International Workshop on Human Security in Mongolia. Two-Day Session in Ulaanbaatar, May 8-10*, Press Release SG/SM/7382. Available online at: http://www.gdrc.org/sustdev/husec/Definitions.pdf

Arora, S. (2008): National e-ID card schemes: a European overview. *Information Security Technical Report*, 13(2): 46-53.

Arthur, B. W. (1989): Competing Technologies, Increasing Returns, and Lock-in by historical events. *Economic Journal*, 99: 116-131. Available online at: http://www.haas.berkeley.edu/Courses/Spring2000/BA269D/Arthur89.pdf

Arthur, C. (2013): iPhone 5S fingerprint sensor hacked by Germany's Chaos Computer Club. *The Guardian*, September 23, https://www.theguardian.com/technology/2013/sep/22/apple-iphone-fingerprint-scanner-hacked

Bager, J. (2016): Timing-Attacke deanonymisiert Website-Besucher teilweise. *Heise Online*, September 13, http://www.heise.de/newsticker/meldung/Timing-Attacke-deanonymisiert-Website-Besucher-teilweise-3319599.html

Ball, J., Schneier, B., Greenwald, G. (2013): NSA and GCHQ target Tor network that protects anonymity of web users. *The Guardian*, October 4,

https://www.theguardian.com/world/2013/oct/04/nsa-gchq-attack-tor-network-encryption

Ball, K. and Webster, F. (2003): *The intensification of surveillance: crime, terrorism and warfare in the information age*. London: Pluto.

Ball, K., Haggerty, K. and Lyon, D. (eds.) (2012): *Handbook on surveillance studies*. Abingdon/New York: Routledge.

Balzacq, T. (2005): The three faces of securitization: Political agency, audience and context. *European Journal of International Relations*, 11(2): 171-201.

Bamberger, K. A., Mulligan, D. K. (2012): PIA requirements and privacy decision-making in US Government agencies. In Wright, D., De Hert, P. (eds.), *Privacy Impact Assessment*, Dordrecht: Springer: pages 225-250.

Baringhorst, S. (2009): Introduction. Political Campaigning in Changing Media Cultures - Typological and Historical Approaches. In Baringhorst, S., Kneip, V. and Niesyto, J. (eds.), *Political Campaigning on the Web*, Bielefeld: transcript: pages 9-30.

Baum, G. R., Kurz, C., Schantz, P. (2013): Das vergessene Grundrecht. *Frankfurter Allgemein Zeitung*, 27. Februar, http://www.faz.net/aktuell/feuilleton/debatten/datenschutz-das-vergessene-grundrecht-12095331.html

Bayley, R. M., Bennett, C. J. (2012): Privacy Impact Assessments in Canada. In Wright, D., De Hert, P. (eds.), *Privacy Impact Assessment*, Dordrecht: Springer: pages 161-185.

BBC (2016): Turkey blocks access to Tor anonymizing network, *BBC Online*, December 16, http://www.bbc.co.uk/news/technology-38365564

Belbey, J. (2016): Crime and social media: law enforcement is watching. *Forbes Magazine*, August 31, https://www.forbes.com/sites/joannabelbey/2016/08/31/crime-and-social-media-law-enforcement-is-watching/#6cf1d7cd541d

Benkirane, R. (2012): The Alchemy of Revolution: The Role of Social Networks and New Media in the Arab Spring. *GCSP Policy Paper*, No. 2012/7, edited by Geneva Center for Security Policy.

Bennett, C. J. (2011): In Defence of Privacy: The concept and the regime. *Surveillance & Society*, 8(4): 485-496.

Bennett, C. J., Haggerty, K. D. (eds.) (2011): *Security Games: Surveillance and Control at Mega-Events*, New York: Glasshouse.

Bennett, C. J., Lyon, D. (2008): *Playing the identity card - surveillance, security and identification in global perspective*, London/New York: Routledge.

Bertalanffy, L. (1950): An Outline of General System Theory. *The British Journal for the Philosophy of Science*, 1(2): 134-165. Available online at:

http://www.isnature.org/Events/2009/Summer/r/Bertalanffy1950-GST_Outline_SELECT.pdf

Bertalanffy, L. (1969): *General System Theory: Foundations, Development, Applications*. Revised Edition, New York: George Braziller. 18th Paperback Printing 2015.

Bertalanffy, L. (1972): The History and Status of General Systems Theory. In Klir G. (ed.), *Trends in General Systems Theory*, New York: Wiley: pages 21–41; reprinted in: *Academy of Management Journal*, 15(4): 407-426. Available online at: http://amj.aom.org/content/15/4/407

Best, S. (2016): Will you soon need a government ID to log into Facebook? Europe Commission proposes controversial scheme to access social media sites. *Daily Mail Online*, June 3, http://www.dailymail.co.uk/sciencetech/article-3623396/Will-soon-need-government-ID-log-Facebook-Europe-Commission-proposes-controversial-scheme-access-social-media-sites.html#ixzz4JedgLpcu

Beuth, P. (2017): Die Luftpumpen von Cambridge Analytica. *Die Zeit*, 7. März, http://www.zeit.de/digital/internet/2017-03/us-wahl-cambridge-analytica-donald-trump-widerspruch

Beyreuther, T., Eismann, C., Hornung, S., Kleemann. F. (2013): Prosumption of Social Context in Web 2.0. In Dunkel, W. Kleemann, F. (eds.), *Customers at Work -New Perspectives on Interactive Service Work*. London: Palgrave McMillan/AbeBooks: pages 223–252.

Bhargav-Spantzel, A., Camenisch, J., Gross, T., Sommer, D. (2007): User Centricity: A Taxonomy and Open Issues. *Journal of Computer Security*, 15 (5): 493-527.

Biddle, S. (2017): How Peter Thiels Palantir helps the NSA to spy on the whole world. *The Intercept*, February 22, https://theintercept.com/2017/02/22/how-peter-thiels-palantir-helped-the-nsa-spy-on-the-whole-world/

Bieker, F., Friedewald, M. Hansen, M., Obersteller, H., Rost, M. (2016): A Process for Data Protection Impact Assessment under the European General Data Protection Regulation. In Rannenberg, K., Ikonomou, D. (eds.), *Privacy Technologies and Policy. Fourth Annual Privacy Forum (APF),* LNCS 9857. Frankfurt/Heidelberg/New York/Dordrecht/London: Springer: pages21-37.

Bigo, D. (2000): When two become one: Internal and external securitisations in Europe. In Kelstrup, M. Williams, M. (eds.), *International Relations Theory and the Politics of European Integration. Power, Security and Community*. London: Routledge: pages 171-204.

Bigo, D. (2008): Globalized (In)Security: The field and the Ban-Opticon. In Bigo, D., Tsoukala, A. (eds.), *Terror, Insecurity and Liberty. Illiberal practices of liberal regimes after 9/11*. Oxon/New York: Routledge: pages 10-48.

Blumberg, A.J., Eckersley, P. (2009): *On locational privacy, and how to avoid losing it forever*. Electronic Frontier Foundation. Available online at: https://www.eff.org/wp/locational-privacy

Boffey, D. (2017): Google fined record €2.4bn by EU over search engine results. *The Guardian*, June 27, https://www.theguardian.com/business/2017/jun/27/google-braces-for-record-breaking-1bn-fine-from-eu

Bonneau, J., Anderson, J., Anderson, R. Stajano, F. (2009): Eight friends are enough: Social graphs approximation via public listings. In *Proceedings of the Second ACM EuroSys Workshop on Social Network Systems (SNS)*, pages 13-18. Available online at: http://www.cl.cam.ac.uk/~rja14/Papers/8_friends_paper.pdf

Bonneau, J., Herley, C., van Oorschot, P. C., Stajano, F. (2012): The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. In *IEEE Symposium on Security and Privacy*. Available online at: https://www.cl.cam.ac.uk/~fms27/papers/2012-BonneauHerOorSta-password--oakland.pdf

Boulianne, S. (2015): Social media use and participation: a meta-analysis of current research. *Information, Communication & Society*, 18(5): 524-538.

Boyd, D. M., Ellison, N. B. (2007): Social Network Sites: Definition, History, and Scholarship, *Journal of Computer-Mediated Communication*, 13(1): 210-230.

Boyd, D., (2010). Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In Papacharissi, Z. (ed.), *Networked Self: Identity, Community, and Culture on Social Network Sites.* New York: Routledge: pages 39-58.

Brandom, R. (2016): Apple's new facial recognition feature could spur legal issues. *The Verge*, June 16, https://www.theverge.com/2016/6/16/11934456/apple-google-facial-recognition-photos-privacy-faceprint

Brandtdaeg, P. B., Heim, J., (2009): Why people use social networking sites. In Ozok, A. A. and Zaphiris, P. (eds.), *Online Communities*, LNCS 5621, Springer: pages 143-152.

Brin, D. (1998): *The Transparent Society. Will Technology Force Us to Choose Between Privacy and Freedom?* Reading, MA: Perseus Publishing.

Bromwich, J. E., Victor, D., Isaac, M. (2016): Police Use Surveillance Tool to Scan Social Media, A.C.L.U. Says. *New York Times*, October 11, https://www.nytimes.com/2016/10/12/technology/aclu-facebook-twitter-instagram-geofeedia.html

Brugger, J., Fraefel, M., Riedl, R. (2014): Raising Acceptance of Cross-Border eID Federation by Value Alignment. *Electronic Journal of e-Government*, 12(2): 179-189.

BSI – Bundesamt für Sicherheit in der Informationstechnik (2008): *BSI Standard 100-2: IT-Grundschutz Methodology.* Version 2.0. Bonn, Germany. Available online at: https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/BSIStandards/standard_100-2_e_pdf.pdf?__blob=publicationFile&v=1

Buettner, R. (2016): Predicting user behavior in electronic markets based on personality-mining in large online social networks - A personality-based product recommender framework. *Electronic Markets - The International Journal on Networked Business*: 1-19. doi:10.1007/s12525-016-0228-z

Buhl, M. (2015): Millennials Boast Huge Social Networking Growth and Engagement on Smartphones, But Older Users Surprisingly Outpace Them on Tablets. *Comscore*, August 12, http://www.comscore.com/Insights/Blog/Millennials-Boast-Huge-Social-Networking-Growth-and-Engagement-on-Smartphones

Burns, M. (2015): Leaked Palantir Doc Reveals Uses, Specific Functions And Key Clients. *Techcrunch*, November 01, https://techcrunch.com/2015/01/11/leaked-palantir-doc-reveals-uses-specific-functions-and-key-clients/

Butler, J. (2006): *Gender trouble: feminism and the subversion of identity.* New York: Routledge.

Buzan, B., Weaver, O., de Wilde, J., (1998): *Security: A New Framework for Analysis*. Boulder: Lynne Rienner.

Buzzetto-More, N. A. (2013): Social Media and Prosumerism. *Issues in Informing Science and Information Technology*, 10: 67-80. Available online at: http://iisit.org/Vol10/IISITv10p067-080Buzzetto0040.pdf

Cachia, R. (2008): *Social Computing: Study on the Use and Impact of Online Social Networking. IPTS Exploratory Research on the Socio-economic Impact of Social Computing:* JRC Scientific and Technical Reports - Institute for Prospective Technological Studies (IPTS) - European Commission. Luxembourg: office for official publications of the European Communities. Available online at: http://ftp.jrc.es/EURdoc/JRC48650.pdf

Calude, C. S., Longo, G. (2016): The Deluge of Spurious Correlations in Big Data. *Foundations of Science* (2016): 1-18. doi:10.1007/s10699-016-9489-4 Available online at: http://link.springer.com/article/10.1007/s10699-016-9489-4

Cameron, K. (2005): The laws of identity. *Identityblog*, November 5, http://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf

Cao, Y., Song, L., Wijmans, E. (2017): (Cross-)Browser Fingerprinting via OS and Hardware Level Features. In *Proceedings of the 24th NDSS Symposium, 26 February - 1 March 2017, San Diego, CA.* Internet Society. Available online at: http://dx.doi.org/10.14722/ndss.2017.23152

CASE Collective (2006): Critical Approaches to Security in Europe: A Networked Manifesto. *Security Dialogue 37* (4): 443-487.

Castells, M. (1996): *The Rise of the Network Society: The Information Age: Economy, Society, and Culture Volume I*. Oxford: Blackwell.

Castells, M. (2000): Materials for an exploratory theory of the network society. *British Journal of Sociology*, 51(1): 5-24.

Castells, M. (2003): *Das Informationszeitalter. Der Aufstieg der Netzwerkgesellschaft*. Leske+Budrich: Opladen.

Castro, D., McQuinn, A. (2015): Beyond the USA Freedom Act: How U.S. surveillance still subverts U.S. competitiveness. *Information Technology and Innovation Foundation*, June, http://www2.itif.org/2015-beyond-usa-freedom-act.pdf?_ga=1.114044933.369159037.1433787396

Cavoukian, A. (2012a): *Privacy by Design and the Emerging Personal Data Ecosystem*. Information and Privacy Commissioner, Ontario, Canada. Available online at: https://www.ipc.on.ca/wp-content/uploads/Resources/pbd-pde.pdf

Cavoukian, A. (2012b): *Operationalizing Privacy by Design: A Guide to Implementing Strong Privacy Practices*. Available online at: http://www.cil.cnrs.fr/CIL/IMG/pdf/operationalizing-pbd-guide.pdf

Cavoukian, A., Dix, A., Emam, K., E. (2014): *The unintended consequences of privacy paternalism*. Policy paper, Information and Privacy Commissioner, Ontario, Canada. Available online at: http://www.comm.utoronto.ca/~dimitris/JIE1001/levin4.pdf

CEN – Commité Européen Normalisation (2004): Towards an electronic ID for the European Citizen, a strategic vision. *CEN/ISSS Workshop eAuthentication, Brussels*. Available online at: http://www.umic.pt/images/stories/publicacoes/Towards%20eID.pdf

CFREU (2000): *Charter of Fundamental Rights of the European Union* http://www.europarl.europa.eu/charter/pdf/text_en.pdf

Chandler, J. (2009): Privacy versus national security – clarifying the trade-off. In Kerr, I., Steeves, V., Lucock, C. (eds.), *Lessons from the identity trail – anonymity, privacy and identity in a networked society*. Oxford University Press: pages 121-138.

Chaum, D. (1985): Security without identification: transaction systems to make big brother obsolete. *Communications of the ACM*, 28(10): 1030-1044. Available online at: https://gnunet.org/sites/default/files/10.1.1.48.4680.pdf

Chin, J. (2015): China Is Requiring People to Register Real Names for Some Internet Services. *Wall Street Journal*, February 4, http://www.wsj.com/articles/china-to-enforce-real-name-registration-for-internet-users-1423033973

Christl, W., Spiekermann, S. (2016): *Networks of Control. A Report on Corporate Surveillance, Digital Tracking, Big Data & Privacy*. Vienna: Facultas. Available online at: http://www.privacylab.at/wp-content/uploads/2016/09/Christl-Networks__K_o.pdf

Christman, J. (2015): Autonomy in Moral and Political Philosophy. In *Stanford Encyclopedia of Philosophy*. First published 2003, updated 2015. Available online at: http://plato.stanford.edu/entries/autonomy-moral/

Cisco (2016): The Zettabyte Era – Trends and Analysis, white paper, http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html

CJEU – Court of Justice of the European Union (2014): The Court of Justice declares the Data Retention Directive to be invalid. Press release no. 54/14. Luxembourg, 8 April 2014. Judgment in Joined Cases C-293/12 and C-594/12 Digital Rights Ireland and Seitlinger and Others. http://curia.europa.eu/jcms/upload/docs/application/pdf/2014-04/cp140054en.pdf

CJEU – Court of Justice of the European Union (2016). The Members States may not impose a general obligation to retain data on providers of electronic communications services. Press release no. 145/16, Luxembourg, 21 December 2016. Judgment in Joined Cases C-203/15 Tele2 Sverige AB v Post-och telestyrelsen and C-698/15 Secretary of State for the Home Department v Tom Watson and Others. http://curia.europa.eu/jcms/upload/docs/application/pdf/2016-12/cp160145en.pdf

Clarke, R. (1988): Information Technology and Dataveillance. *Communications of the ACM*, 31(5): 498-512. Available online at: http://www.rogerclarke.com/DV/CACM88.html

Clarke, R. (1994a): Human Identification in Information Systems: Management Challenges and Public Policy Issues. *Information Technology & People*, 7(4): 6-37. Available online at: www.rogerclarke.com/DV/HumanID.html

Clarke, R. (1994b): The Digital Persona and its Application to Data Surveillance. *The Information Society*, 10(2): 77-92. Available online at: http://www.rogerclarke.com/DV/DigPersona.html#DP

Clarke, R. (2001): *Biometrics and Privacy*. www.rogerclarke.com/DV/Biometrics.html

Clarke, R. (2006): *What's 'privacy'?* http://www.rogerclarke.com/DV/Privacy.html

Clarke, R. (2009): Privacy Impact Assessment: Its Origins and Development. *Computer Law & Security Review*, 25(2): 123-135. http://rogerclarke.com/DV/PIAHist-08.html

Clarke, R. (2012a): Location tracking of mobile devices: Ueberveillance stalks the streets. *Computer Law & Security Review* 29(3): 216-228. Available online at: http://www.rogerclarke.com/DV/LTMD.html

Clarke, R. (2012b): PIAs in Australia: A Work-In-Progress Report. In Wright, D., De Hert, P. (eds.), *Privacy Impact Assessment*, Springer:Dordrecht. pages 119-148.

CNBC (2016a): Alphabet earnings: $7.50 per share, vs expected EPS of $7.97. *CNBC*, April 21, http://www.cnbc.com/2016/04/21/alphabet-reports-first-quarter-results.html

CNBC (2016b): Facebook shatters Wall Street estimates, proposes new share structure. *CNBC*, April 27, http://www.cnbc.com/2016/04/27/facebook-reports-first-quarter-earnings.html

CNIL – Commission Nationale de l'Informatique et des Libertés (2015a): *Privacy Impact Assessment (PIA) – Methodology (how to carry out a PIA).* June 2015 edition. Available online at: https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-1-Methodology.pdf

CNIL – Commission Nationale de l'Informatique et des Libertés (2015b): *Privacy Impact Assessment (PIA) – Tools (templates and knowledge bases).* June 2015 Edition. https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-2-Tools.pdf

Cohen, J., E. (2012): *Configuring the networked self: law, code, and the play of everyday practice.* Yale: Yale University Press.

Colvile, R. (2013): Eric Schmidt interview: 'You have to fight for your privacy or you will lose it' *The Telegraph*, May 25, http://www.telegraph.co.uk/technology/eric-schmidt/10076175/Eric-Schmidt-interview-You-have-to-fight-for-your-privacy-or-you-will-lose-it.html

Comor, E. (2010): Digital prosumption and alienation. *emphemera – theory & politics in organization*, 10(3/4): 439-454. Available online at: http://www.ephemerajournal.org/contribution/digital-prosumption-and-alienation

Confessore, N., Hakim, D. (2017): Data Firm Says 'Secret Sauce' Aided Trump; Many Scoff. *The New York Times*, March 6, https://www.nytimes.com/2017/03/06/us/politics/cambridge-analytica.html

Connor, N. (2016): Female Chinese students 'asked to hand over nude photos to secure loans'. *The Telegraph*, June 15, http://www.telegraph.co.uk/news/2016/06/15/female-chinese-students-asked-to-hand-over-nude-photos-to-secure/

Constine, J. (2016): Facebook Messenger now allows payments in its 30,000 chat bots. *Techcrunch*, September 12, https://techcrunch.com/2016/09/12/messenger-bot-payments/

Cuthbertson, A. (2016): Facial Recognition Can Be Tricked With Facebook Photos. *Newsweek Europe*, August 22, http://europe.newsweek.com/facial-recognition-can-be-tricked-facebook-photos-492329?rm=eu

Dahlgren, P. (2013): Do social media enhance democratic participation? – The importance and difficulty of being realistic. Policy Paper No. 4/2013, edited by Rosa Luxemburg Stiftung Berlin. Available online at: http://www.rosalux.de/fileadmin/rls_uploads/pdfs/Standpunkte/policy_paper/Policy Paper_04-2013.pdf

Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J., Le Métayer, D., Tirtea, R., Schiffner, S. (2014): *Privacy and Data Protection by Design – from policy to*

engineering. European Union Agency for Network and Information Security (ENISA), December 2014.

De Andrade, N. N. G., Monteleone, S., Martin, A. (2013): *Electronic Identity in Europe: Legal Challenges and Future Perspectives (e-ID 2020).* JRC Scientific and Policy Reports. Joint Research Centre, European Commission.

De Cristofaro, E., Du, H., Freudiger, J., Norcie, G. (2014): A Comparative Usability Study of Two-Factor Authentication. *8th NDSS Workshop on Usable Security (USEC 2014)*, briefing paper, Internet Society, available online at: http://www.internetsociety.org/sites/default/files/01_5-paper.pdf

De Hert, P. (2008): Identity management of e-ID, privacy and security in Europe. A human rights view. *Information Security Technical Report*, 13: 71-75.

De Hert, P. (2012): A Human Rights Perspective on Privacy and Data Protetion Impact Assessments. In Wright, D., De Hert, P. (eds.), *Privacy Impact Assessment*, Springer:Dordrecht: pages 33-76.

Debatin, B., Lovejoy, J. P. (2009): Facebook and Online Privacy: Attitudes, Behaviors, and Unintended Consequences. *Journal of Computer-Mediated Communication*, 15(1): 83-108.

Denyer, S. (2016): China wants give all its citizens a score – and their rating could affect every area of their lives. *The Independent*, October 22, http://www.independent.co.uk/news/world/asia/china-surveillance-big-data-score-censorship-a7375221.html

Dewey, T., Kaden, J., Marks, M., Matsushima, S., and Zhu, B. (2012): *The Impact of Social Media on Social Unrest in the Arab Spring.* Final report prepared for: Defense Intelligence Agency. Stanford University. Available online at: https://stanford.box.com/shared/static/c5v3umqoc7oa2b9p6qjb.pdf

Dixon, P., Gellman, R. (2014): *The scoring of America: how secret consumer scores threaten your privacy and your future*. Research Report, World Privacy Forum. Available online at: https://www.ftc.gov/system/files/documents/public_comments/2014/04/00007-89171.pdf

Dmytrenko, O., Nardali, A. (2005): .Net Passport under the scrutiny of U.S. and EU privacy law: implications for the future of online authentication. *Journal of Law and Policy for the Information Society*, 1(2-3): 619-645. Available online at: https://kb.osu.edu/dspace/bitstream/handle/1811/72710/ISJLP_V1N2-3_619.pdf?sequence=1

Doctrow, C. (2014): If you read Boing Boing, the NSA considers you as target for deep surveillance. *Boing Boing*, July 3, http://boingboing.net/2014/07/03/if-you-read-boing-boing-the-n.html

Doctrow, C. (2016): Researchers trick facial recognition systems with facial features printed on big glasses. *BoingBoing*, November 2, http://boingboing.net/2016/11/02/researchers-trick-facial-recog.html

DPD – Data Protection Directive (1995): EU Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046&from=DE

Dreged, S. (2013): What is Tor? A beginner's guide to the privacy tool. *The Guardian*, November 5, https://www.theguardian.com/technology/2013/nov/05/tor-beginners-guide-nsa-browser

Dunn, J., E. (2017): 22 of the most infamous data breaches affecting the UK. *Techworld*, July 14, http://www.techworld.com/security/uks-most-infamous-data-breaches-3604586/

Dupont, B. (2008): Hacking the panopticon: distributed online surveillance and resistance. *Sociology of Crime Law and Deviance*, 10: 259-280.

Dwork, C., Roth, A. (2014): The Algorithmic Foundations of Differential Privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4): 211–407.

EB – Eurobarometer (2015): *Special Eurobarometer 431 – Data protection*. European Union. Available online at: http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_431_en.pdf

EB – Eurobarometer (2016): *Flash Eurobarometer 443 – Report e-privacy*. TNS political & social, European Union. Available online at: https://ec.europa.eu/digital-single-market/en/news/eurobarometer-eprivacy

ECHR (1953): *European Convention on Human Rights*. Available online at: http://www.echr.coe.int/Documents/Convention_ENG.pdf

Eckersley, P. (2010): How unique is your web browser? In Atallah, M., Hopper, N. (eds.), *Proceedings of the 10th international conference on Privacy enhancing technologies (PETS'10),* Heidelberg: Springer: pages1-18. Available online at: https://panopticlick.eff.org/static/browser-uniqueness.pdf

EDPS – European Data Protecetion Supvervisor (2016a): *Opinion 5/2016 - Preliminary opinion Preliminary EDPS Opinion on the review of the ePrivacy Directive (2002/58/EC).* Available online at: https://edps.europa.eu/sites/edp/files/publication/16-07-22_opinion_eprivacy_en.pdf

EDPS - European Data Protecetion Supvervisor (2016b): *Opinion 8/2016 - Opinion on coherent enforcement of fundamental rights in the age of big data*. Available online at: https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/EDPS/Events/16-09-23_BigData_opinion_EN.pdf

Edwards, L., McAuley, D. (2013): What's in a name? Real name policies and social networks. In *Proceedings of the 1st workshop on Internet Science and Web Science Synergies, Paris, May 2013*. Available online at: http://www.cs.nott.ac.uk/~pszdrm/papers/2013_NamesHavePower%20paris%20vn.pdf

EGE - European Group on Ethics in Science and New Technologies (2014): *Ethics of Security and Surveillance Technologies*. Opinion No. 28 of the European Groups on Ethics in Science and New Technologies. Brussels, May 20.

Egelman, S. (2013): My Profile is My Password, Verify Me! The Privacy/Convenience Tradeoff of Facebook Connect. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, April 27 – May 02*, pages 2369-2378. Available online at: http://www.guanotronic.com/~serge/papers/chi13a.pdf

Ellison, N. (2013): *Future Identities: Changing identities in the UK – the next 10 years. DR3: Social Media and Identity*. Government Office for Science Forsesight United Kingdom. Research report. Available online at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/275752/13-505-social-media-and-identity.pdf

Ellison, N., B., Boyd, D., M. (2013): Sociality through Social Network Sites. In Dutton, W. H. (ed.): *The Oxford Handbook of Internet Studies*, Oxford: Oxford University Press: pages 151-172.

Elmer, G. (2004): *Profiling Machines*. Cambridge/MA: MIT Press.

EPIC – Electronic Privacy Information Center (2003): Sign Out of Passport! https://epic.org/privacy/consumer/microsoft/

EPIC – Electronic Privacy Information Center (2016): *EPIC v. FBI - Next Generation Identification – Seeking documents about the FBI's expansive biometric identification database.* http://epic.org/foia/fbi/ngi/

EU-C – European Commission (2006): i2010 eGovernment Action Plan: Accelerating eGovernment in Europe for the Benefit of All. COM(2006) 173 final. Brussels, April 25, http://ec.europa.eu/smart-regulation/impact/ia_carried_out/docs/ia_2006/sec_2006_0511_en.pdf

EU-C – European Commission (2010a): *Digitizing Public Services in Europe: Putting ambition into action, 9th benchmark measurement*, December 2010, Directorate General for Information Society and Media, Unit C.4 Economic and Statistical Analysis. http://ec.europa.eu/newsroom/document.cfm?action=display&doc_id=747

EU-C – European Commission (2010b): *A comprehensive approach on personal data protection in the European Union*. Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions, COM (2010) 609 final, Brussels, Nov. 4. http://ec.europa.eu/justice/news/consulting_public/0006/com_2010_609_en.pdf

EU-C – European Commission (2016a): *EU eGovernment Action Plan 2016-2020 Accelerating the digital transformation of government*. COM(2016) 179 final. Brussels, April 19, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=15268

EU-C – European Commission (2016b): *Online Platforms and the Digital Single Market Opportunities and Challenges for Europe.* Communication from the Commission to the European Parliament, the Council, the European economic and social Committee and the Committee of the Regions, COM (2016) 288 final. Brussels, May 25, http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52016DC0288&from=EN

EU-CHR – European Court of Human Rights (2017): *Factsheet – data protection*, July 2017 http://www.echr.coe.int/Documents/FS_Data_ENG.pdf

EuroPriSe – European Privacy Seal (2017): *EuroPriSe Criteria for the certification of IT products and IT-based services.* "GDPR ready" version, January 2017. https://www.european-privacy-seal.eu/AppFile/GetFile/e5ed7122-74b1-4f75-a5af-fb0c317bd20b

Farivar, C. (2012, July 18). Washington State will enable voter registration via Facebook. *Ars Technica.* http://arstechnica.com/business/2012/07/washington-residents-to-be-able-to-register-to-vote-via-facebook

Finley, K. (2014): Encrypted web traffic more than doubles after NSA revelations. *Wired*, May 16, https://www.wired.com/2014/05/sandvine-report/

Finn, R. L., Wright, D., Friedewald, M. (2013): Seven Types of Privacy. In Gutwirth, S., Leenes, R., De Hert, P., Poullet, Y. (eds.), *European Data Protection: Coming of Age,* Dordrecht: Springer: pages 3-32.

Floridi, L. (2010): Ethics after the Information Revolution. In Floridi, L. (ed.): *The Cambridge Handbook of Information and Computer Ethics.* Cambridge/UK: Cambridge University Press. pages 3-19.

Floridi, L. (2013): *The Ethics of Information.* Oxford: Oxford University Press.

Foucault, M. (1977): *Discipline and Punish: the birth of the prison*. (translated from the French by A. Sheridan, 2nd edition 1995), New York: Vintage Books/Randomhouse.

FRA – European Union Agency for Fundamental Rights (2010): *Data protection in the European Union: the role of national Data Protection Authorities. Strengthening the fundamental rights architecture in the EU II*. Luxembourg: Publications Office of the European Union.

FRA – European Union Agency for Fundamental Rights (2014): *Handbook on European data protection law.* Luxembourg: Publications Office of the European Union.

FRA – European Union Agency for Fundamental Rights (2016): *Fundamental Rights Report 2016.* Luxembourg: Publications Office of the European Union. Available

online at: http://fra.europa.eu/sites/default/files/fra_uploads/fra-2016-fundamental-rights-report-2016-2_en.pdf

Frazer, N. (2007): Transnationalising the Public Sphere: On the Legitimacy and Efficacy of Public Opinion in a Post-Westphalian World. *Theory, Culture and Society*, 24: 7-30.

Friedewald, M., van Lieshout, M., Rung, S., Ooms, M., Ypma, J. (2015): Privacy and Security Perceptions of European Citizens: A Test of the Trade-Off Model. In Camenisch, J., Fischer-Hübner, S., Hansen, M. (eds.), *Privacy and Identity for the Future Internet in the Age of Globalization*, series IFIP AICT Vol. 457. Heidelberg: Springer: pages 39-53.

Friedwald, M., Burgess, P. J., Čas, J., Bellanova, R., Peissl, W. (eds.) (2017): *Surveillance, privacy, and security – citizens' perspectives*. London/New York: Roudledge.

Froomkin, D. (2015): FBI Director Claims Tor and the "Dark Web" Won't Let Criminals Hide From His Agents. *The Intercept*, October 9, https://theintercept.com/2015/09/10/comey-asserts-tors-dark-web-longer-dark-fbi/

Fuchs, C. (2010): How can surveillance be defined? Remarks on theoretical foundations of surveillance studies. *The Internet & Surveillance – Research Paper Series* (no. 1). Unified Theory of Information Research Group, Available online at: http://www.sns3.uti.at/wp-content/uploads/2010/10/The-Internet-Surveillance-Research-Paper-Series-1-Christian-Fuchs-How-Surveillance-Can-Be-Defined.pdf

Fuchs, C. (2014): Digital prosumption labour on social media in the context of the capitalist regime of time. *Time & Society*, 23 (1): 97-123. Available online at: http://fuchs.uti.at/wp-content/time.pdf

Fuchs, C. (2015): Social media surveillance. In Coleman, S., Freelon, D. (eds.), *Handbook of digital politics.* Cheltenham: Edward Elgar: pages 395-414

Fuchs, C., Hofkirchner, W. (2005): The Dialectic of Bottom-up and Top-down Emergence in Social Systems. *tripleC*, 3(2): 28-50. Available online at: http://www.triple-c.at/index.php/tripleC/article/view/21/20

Fuchs, C., Sandoval, M. (2015): 2015. The Political Economy of Capitalist and Alternative Social Media. In Atton, C (ed.): *The Routledge Companion to Alternative and Community Media.* London: Routledge: pages 165-175.

Gafni, R., Nissim, D. (2014): To social login or not login? - Exploring factors affecting the decision. *Issues in Informing Science and Information Technology*, 11: 57-72. Available online at: http://iisit.org/Vol11/IISITv11p057-072Gafni0462.pdf

Gallagher, R., Hager, N. (2016): The Raid - In Bungled Spying Operation, NSA Targeted Pro-Democracy Campaigner. *The Intercept,* August 15, https://theintercept.com/2016/08/14/nsa-gcsb-prism-surveillance-fullman-fiji/

Gandy, O. (1993): *The panoptic sort: A Political Economy of Personal Information*. Boulder: Westview.

GAO – US Government Accountability Organization (2016): *Face recognition technology - FBI Should Better Ensure Privacy and Accuracy.* Report to the Ranking Member, Subcommittee on Privacy, Technology and the Law, committee on the Judiciary, U.S. Senate. Available online at http://www.gao.gov/assets/680/677098.pdf

Garvie, C., Bedoya, A. M., Frankle, J. ((2016): *The perpetual line-up – unregulated police face recognition in America*. Georgetown Law Center for Privacy & Technology, Washington D.C. Available online at: https://www.perpetuallineup.org

GDPR – General Data Protection Regulation (2016): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&qid=1485427623759&from=en

Geambasu, R., Kohno, T., Levy, A., Levy, H.M. (2009): Vanish: Increasing Data Privacy with Self-Destructing Data. In *Proceedings of the 18th USENIX Security Symposium (SSYM'09), Montreal, Canada*, pages 299-316. Available online at: https://vanish.cs.washington.edu/pubs/usenixsec09-geambasu.pdf

Geels, F.W. (2004): Understanding system innovations: a critical literature review and a conceptual synthesis. In Elzen, B., Geels, F.W., Green, K., (eds.), *System innovation and the transition to sustainability: Theory, evidence and policy*. Cheltenham, UK/Northamption: Edward Elgar: pages 19-47.

Geels, F.W., Schot, J. (2007): Typology of sociotechnical transition pathways. *Research Policy*, (36): 399-417.

Gellman, B., Makron, J. (2013): Edward Snowden says motive behind leaks was to 'expose surveillance state', *Washington Post*, June 10, https://www.washingtonpost.com/politics/edward-snowden-says-motive-behind-leaks-was-to-expose-surveillance-state/2013/06/09/aa3f0804-d13b-11e2-a73e-826d299ff459_story.html?utm_term=.1cabd9f22bdf

Gemalto (2014): National Mobile ID schemes – Learning from today's best practices. White paper, http://www.securitydocumentworld.com/creo_files/upload/article-files/wp_mobileid_overview_en.pdf

Gibbs, S. (2016a): WhatsApp to give users' phone numbers to Facebook for targeted ads. *The Guardian*, August 25, https://www.theguardian.com/technology/2016/aug/25/whatsapp-to-give-users-phone-number-facebook-for-targeted-ads

Gibbs, S. (2016b): US border control could start asking for your social media accounts, *The Guardian*, June 28, https://www.theguardian.com/technology/2016/jun/28/us-customs-border-protection-social-media-accounts-facebook-twitter

Giddens, A. (1984): *The Constitution of Society. Outline of the Theory of Structuration*. Cambridge: Polity Press.

Giddens, A. (1991): *Modernity and Self Identity: self and society in the late modern age.* Cambridge: Polity Press.

Giddens, A. (1997): *Die Konstitution der Gesellschaft*. 3. Auflage, Campus Verlag: Frankfurt/NewYork.

Gillespie, T. (2010): The Politics of 'Platforms'. *New Media & Society*, 12(3): 347-364.

Gillings, M. R., Hilbert, M., Kemp, D. J. (2016): Information in the Biosphere: Biological and Digital Worlds. *Trends in Ecology and Evolution*, 31(3): 180-189.

Glässer, U., Vajihollahi M. (2010): Identity Management Architecture. In Yang, C.C., Chau, M.C., Wang, J.-H., Chen, H. (eds.), *Security Informatics, Annals of Information Systems Vol. 9*, Springer: pages 97–116.

Goel, V., Wyatt, E.(2013): Facebook Privacy Change Is Subject of F.T.C. Inquiry. *The New York Times*, September 11, http://www.nytimes.com/2013/09/12/technology/personaltech/ftc-looking-into-facebook-privacy-policy.html

Goldstein, I. S. (2011): Regulating privacy by design. *Berkeley Technology Law Journal*, 26(3): 1409-1456.

González-Fuster, G. (2014): *The emergence of personal data protection as a fundamental right of the EU. Law, Governance and Technology Series Vol. 16*, Cham/Heidelberg/New York etc.: Springer.

Gordon, D. (1987): The Electronic Panopticon: A Case Study of the Development of the National Crime Records System. *Politics and Society*, 15 (4): 483-511.

Granovetter, M. S., (1973): The Strength of Weak Ties. *American Journal of Sociology* 78(6): 1360-1380.

Grassi, P. A., Garcia, M. E., Fenton, J. L. (2017): *Digital identity guidelines. NIST special publication 800-63-3.* NIST - US National Institute of Standards and Technology. Available online at: https://www.nist.gov/itl/tig/special-publication-800-63-3

Greenberg, A. (2015): MIT's Bitcoin-Inspired 'Enigma' Lets Computers Mine Encrypted Data. *Wired*, June 30, https://www.wired.com/2015/06/mits-bitcoin-inspired-enigma-lets-computers-mine-encrypted-data/

Greenberg, A. (2016a): Hacker Lexicon: What Is Perfect Forward Secrecy? *Wired*, November 28, https://www.wired.com/2016/11/what-is-perfect-forward-secrecy/

Greenberg, A. (2016b): Signal, the Cypherpunk App of Choice, Adds Disappearing Messages. *Wired*, October 11, https://www.wired.com/2016/10/signal-cypherpunk-app-choice-adds-disappearing-messages/

Greenleaf, G. (2017): Renewing Convention 108: The CoE's 'GDPR Lite' Initiatives. 142 *Privacy Laws & Business International Report,* 14-17; UNSW Law Research Paper

No. 17-3. Available online at: https://ssrn.com/abstract=2892947 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2892947

Greenwald, G. (2013): XKeyscore: NSA tool collects 'nearly everything a user does on the internet'. *The Guardian*, July 31, https://www.theguardian.com/world/2013/jul/31/nsa-top-secret-program-online-data

Greenwald, G. (2014): *No place to hide – Edward Snowden, the NSA and the Surveillance State.* London: Hamish Hamilton/Penguin Books.

Guild, E., Carrera, S., Balzacq, T. (2008): *The changing dynamic of security in an enlarged European Union*. Research paper No. 12, CEPS Programme Series. Available online at: www.ceps.eu http://aei.pitt.edu/11457/1/1746.pdf

Gulyás, G. G., Simon, B. Imre, S. (2016): An Efficient and Robust Social Network De-anonymization Attack. In *Proceedings of the ACM Workshop on Privacy in the Electronic Society, October 2016, Vienna, Austria*. pages 1-11. doi: 10.1145/2994620.2994632

Gürses, S., Troncoso, C., Diaz, C. (2011): Engineering Privacy by Design. Paper presented at the *Conference on Computers, Privacy & Data Protection (CPDP), 25-28 January 2016, Brussels, Belgium*. Available online at: https://www.esat.kuleuven.be/cosic/publications/article-1542.pdf

Habermas, J. (1989): *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society* (original work: Strukturwandel der Öffentlichkeit 1962, Hermann Luchterhand Verlag), Cambridge MA: The MIT Press.

Haggerty, K. D. (2006): Haggerty, Tear down the walls: On demolishing the panopticon. In Lyon, D. (ed.), *Theorizing surveillance: The panopticon and beyond*. Cullompton: Willan Publishing. pages 23–45.

Haggerty, K. D., Ericson, R. V. (2000): The surveillant assemblage. *British Journal of Sociology*, 51(4): 605-622.

Haggerty, K. D., Samatas, M. (eds.) (2010): *Surveillance and democracy*. Oxon: Routledge-Cavendish.

Halperin, R., Backhouse, J. (2008): A roadmap for research on identity in the information society, *Identity in the information society*, 1(1): 71-87.

Hamacher, K., Katzenbeisser, S. (2011): Public Security: Simulations need to replace conventional wisdom. In *Proceedings of the New Security Paradigms ACM Workshop (NSPW11), September 2-15, Marin County, USA*. pages 115-124.

Hansen, M., Berlich, P., Camenisch, J., Clauß, S., Pfitzmann, A., Waidner, A. (2004): Privacy-enhancing identity management. *Information Security Technical Report*, 9(1): 35-44.

Hansen, M., Jensen, M., Rost, M. (2015): Protection Goals for Privacy Engineering. In *Proceedings of the 2015 IEEE Security and Privacy Workshop*. pages 159-166.

Hazari, S., Brown, C. (2013): An Empirical Investigation of Privacy Awareness and Concerns on Social Networking Sites. *Journal of Information Privacy and Security*, 9(4): 31-51.

Heins, M. (2014): The brave new world of social media censorship. Harvard Law Review, 127(8): 325-330.

Heller, C. (2011): *Prima leben ohne Privatsphäre*. C. H. Beck Verlag: München,.

Henriksen-Bulmer, J., Jeary, S. (2016): Re-identification attacks – a systematic literature review. *International Journal of Information Management,* 36(6): 1184-1192.

Hern, A. (2016): Facebook launches facial recognition app in Europe (without facial recognition). *The Guardian*, May 11, https://www.theguardian.com/technology/2016/may/11/facebook-moments-facial-recognition-app-europe

Heylighen, F. and Campell, D. (1995): Selection of Organization at the Social Level: obstacles and facilitators of metasystem transitions. *World Futures: the Journal of General Evolution*, 45(1-4): 181-212.

Hildebrandt, M. (2006): Privacy and Identity. In Claes, E., Duff, A., Gutwirth, S. (eds.), *Privacy and the criminal law*, Antwerpen/Oxford: Intersentia. pages 43-57.

Hildebrandt, M. (2008): Profiling and the rule of law. *Identity in the Information Society (IDIS),* 1(1): 55-70. Available online at: http://link.springer.com/article/10.1007/s12394-008-0003-1#Fn1

Hildebrandt, M. (2011): Introduction: a multifocal view of human agency in the era of autonomic computing. In Hildebrandt/Rouvroy (2011): *Law, Human agency and autonomic computing*. London/New York: Routledge: pages 1-11.

Hildebrandt, M., Gutwirth, S. (eds.) (2008): Profiling the European citizen – cross-disciplinary perspectives. Amsterdam: Springer Netherlands.

Hildebrandt, M., Rouvroy A. (2011): *Law, Human agency and autonomic computing*. London/New York: Routledge.

HNS - HelpNetSecurity (2011): Interpol chief calls for global electronic identity card system, *Help Net Security*, April 6, https://www.helpnetsecurity.com/2011/04/06/interpol-chief-calls-for-global-electronic-identity-card-system/

Hofkirchner W. (2013): *Emergent Information: A unified theory of information framework.* World Scientifc Series in Information Studies: Vol 3.

Hofkirchner, W. (2005): Ludwig von Bertalanffy - Forerunner of Evolutionary Systems Theory. In Gu J., Chroust G. (eds.), *The New Role of Systems Sciences For a Knowledge-based Society. Proceedings of the First World Congress of the International Federation for Systems Research, Kobe, Japan*. Available online at: http://www.bcsss.org/wp-content/uploads/2011/09/pdf41.pdf

Hofkirchner, W. (2010): How to Design the Infosphere: the Fourth Revolution, the Management of the Life Cycle of Information, and Information Ethics as a Macroethics. Knowledge. *Technology and Policy*, 23(1-2): 177-192.

Hofkirchner, W. Schafranek, M. (2011): General System Theory. In Hooker, C. (ed.): *The Philosophy of Complex Systems.* Series: Handbook of the Philosophy of Science Vol 10. North Holland/Elsevier: Oxford: pages 177-194.

Holpuch, A. (2013): Eric Schmidt says government spying is "the nature of our society". *The Guardian*, September 13, https://www.theguardian.com/world/2013/sep/13/eric-schmidt-google-nsa-surveillance

Hood, C.C., Margetts, H.Z. (2007): *The Tools of Government in the Digital Age*, 2nd edition. Public Policy and Politics. Hampshire: Palgrave Mcmillan.

Hornung, G., Schnabel, C. (2009): Data protection in Germany I: The population census decision and the right to informational self-determination. *Computer Law & Security Report,* 25(1): 84-88.

Horvát, EÁ, Hanselmann M., Hamprecht F.A., Zweig K.A. (2013:) You Are Who Knows You: Predicting Links Between Non-members of Facebook. In Gilbert T., Kirkilionis M., Nicolis G. (eds.), *Proceedings of the European Conference on Complex Systems 2012*. Springer Proceedings in Complexity. Dordrecht/Heidelberg/New York/London:Springer. pages 309-315.

Hosanagar, K. (2016): Blame the echo chamber on Facebook. But blame yourself, too. *Wired*, https://www.wired.com/2016/11/facebook-echo-chamber/

http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2010/wp169_en.pdf

IBM (2013): Data Monetization: Telco's Bit-Pipe Dystopia, Cloud and the Redistribution of Value. *IBM Telecom, Media and Entertainment Blog*, May 2, https://www.ibm.com/blogs/insights-on-business/telecom-media-entertainment/data-monetization-telcos-bit-pipe-dystopia-cloud-and-the-redistribution-of-value/

ICO – UK Information Commissioner's Office (2014): *Conducting privacy impact assessments – code of practice.* Version 1.0, Available online at: https://ico.org.uk/media/for-organisations/documents/1595/pia-code-of-practice.pdf

International Telecommunication Union – ITU (2010): *Baseline identity management terms and definitions. Series X: Data Networks, Open System Communications and Security. Cyberspace security – Identity management. Recommendation ITU-T X.1252*. Available online at: http://www.itu.int/SG-CP/example_docs/ITU-T-REC/ITU-T-REC_E.pdf

Isaac, M. (2010): Facebook to Add Facial Recognition Software to Photo Tagging. *Forbes Magazine*, December 15,

https://www.forbes.com/sites/mikeisaac/2010/12/15/facebook-to-add-facial-recognition-software-to-photo-tagging

ISO – International Standards Organisation (2010): *Information technology – Identification cards – On-card biometric comparison*. ISO/IEC 24787:2010(E). First edition 2010-12-15. Available online at: http://bcc.portal.gov.bd/sites/default/files/files/bcc.portal.gov.bd/page/adeaf3e5_cc55_4222_8767_f26bcaec3f70/ISO_IEC_24787.pdf

ISO – International Standards Organisation (2011): *Information technology – Security techniques – Privacy framework*. ISO/IEC 29100:2011(E). First edition 2011-12-15.

ITU – International Telecommunication Union (2005): *Privacy and ubiquitous network societies*. Background paper, ITU workshop on ubiquitous network societies, ITU new initiatives programme April 6-8. UNS/05. Available online at: https://www.itu.int/osg/spu/ni/ubiquitous/Papers/Privacy%20background%20paper.pdf

ITU – International Telecommunication Union (2016): *ICT facts and figures 2016*. http://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx

Jardine, E. (2015): *The Dark Web Dilemma: Tor, Anonymity and Online Policing. Global Commission on Internet Governance*, paper series: no. 21. Centre for International Governance Innovation and Chatham House – the Royal Institute of International Affairs. Available online at: https://www.cigionline.org/sites/default/files/no.21.pdf

Jardine, E. (2016): Tor, what is it good for? Political repression and the use of online anonymity-granting technologies. *New Media & Society* (March 31, Online first) 1-18. Available online at: http://journals.sagepub.com/doi/10.1177/1461444816639976

Jin, L., Chen, Y., Wang, T., Hui, P., Vasilakos, A. V. (2013): Understanding user behavior in online social networks: a survey, *IEEE Communications Magazine*, 50(9): 144-150.

Johnson, B. (2010): Privacy no longer a social norm, says Facebook founder. The *Guardian*, Jan 11, https://www.theguardian.com/technology/2010/jan/11/facebook-privacy

Jolly, R., Ray, D. B. (2006): *The Human Security Framework and National Human Development Reports: A Review of Experiences and Current Debates*. NHDR occasional paper no. 5, United Nations Development Programme, National Human Development Report Unit.New York: UNDP. Available online at: http://www.pogar.org/publications/other/undp/governance/Human-Security-Guidance-Note.pdf

Jøsang, A., AlZomai, M., and Suriadi, S. (2007): Usability and Privacy in Identity Management Architectures. In *Proceedings of the 5th Australasian symposium on ACSW frontiers – Vol. 68*, Australian Computer Society: pages 143-152.

Kahn, J. D. (2003): Privacy as a legal principle of identity maintenance. *Seton Hall Law Review*, 33 (2): 371-410.

Kannenberg, A. (2016): Telefónica Deutschland will Bewegungsdaten von Mobilfunkkunden vermarkten. *Heise Online*, September 22, https://www.heise.de/newsticker/meldung/Telefonica-Deutschland-will-Bewegungsdaten-von-Mobilfunkkunden-vermarkten-3329545.html

Kant, I. (1785/1997): *Grundlegung zur Metaphysik der Sitten.* Werkausgabe Band VIII (herausgegeben von Weischedel, W.) Erstmals erschienen 1785. Suhrkamp Taschenbuch Wissenschaft: Berlin.

Kaplan, A. M., Haenlein, M. (2011). Two hearts in three-quarter time: How to waltz the social media/viral marketing dance. *Busines Horizons*, 54: 253-263.

Kaplan, A. M., Haenlein, M. (2012): The Britney Spears universe: Social media and viral marketing at its best. *Business Horizons*, 55: 27-31.

Katz, M. L., Shapiro, C. (1994): Systems competition and network effects. *Journal of Economic Perspectives*, 8(2): 93-115. Available online at: http://socrates.berkeley.edu/~scotch/katz_shapiro.pdf

Kemp, R., Rip, A., Schot, J. (2001): Constructing transition paths through the management of niches. In Garud, R., Karnoe, P. (eds.), *Path Dependence and Creation*. Mahwa/London: Lawrence Erlbaum. pages 269-299.

Kerr, O. (2015): Edward Snowden's impact. *The Washington Post*, April 9, https://www.washingtonpost.com/news/volokh-conspiracy/wp/2015/04/09/edward-snowdens-impact/

Khurana, T. (2013): Paradoxes of Autonomy: On the Dialectics of Freedom and Normativity. Symposium. *Canadian Journal of Continental Philosophy*, 17: 50-74.

Kim, Y., Sohn, D., Choi, S. M. (2011): Cultural difference in motivations for using social network sites: A comparative study of American and Korean college students. *Computers in Human Behavior*, (27): 365-372.

Kleinfeld, J. S. (2002): The small world problem. *Society*, 39(2): 61-66. Available online at: http://www.stat.cmu.edu/~fienberg/Stat36-835/Kleinfeld_SWP.pdf

Kleinz, T. (2016): Datenschutzbedenken: Mozilla entfernt Akku-Fingerprinting aus Firefox. *Heise Online*, Oktober 31, https://www.heise.de/newsticker/meldung/Datenschutzbedenken-Mozilla-entfernt-Akku-Fingerprinting-aus-Firefox-3405099.html

Klíma, V. (2005): Finding MD5 Collisions – a Toy for a Notebook. *Cryptology ePrint Archive*, Report 2005/075, Available online at: https://eprint.iacr.org/2005/075.pdf

Klitou, D. (2014): Privacy, Liberty and Security. In *Privacy-Invading Technologies and Privacy by Design – Safeguarding Privacy, Liberty and Security in the 21st Century*. Information Technology and Law Series. The Hague: Springer/TMC Asser Press. pages 13-25.

Klopfer, P. H., Rubenstein, D. I. (1977): Privacy and its biological basis. *Journal of Social Issues*, 33(3): 53-65.

Kneer, G., Nassehi, A. (2000): *Niklas Luhmanns Theorie sozialer Systeme*. 4. Auflage, (1. Auflage 1993). W. Fink Verlag UTB: Paderborn.

Kolb, D.G. (2008): Exploring the metaphor of connectivity: attributes, dimensions and duality. *Organization Studies*, 29 (1): 127-144.

Korfmacher C. (2006): Personal Identity. In *Internet Encyclopaedia of Philosophy*. available online at: http://www.iep.utm.edu/person-i/

Kozierok, C., M: (2005): The TCP/IP Guide, Version 3, http://www.tcpipguide.com/free/t_TCPPortsConnectionsandConnectionIdentification.htm

Krawczyk, H. (2005): Perfect Forward Secrecy. In van Tilborg, H.C.A. (ed.), *Encyclopedia of Cryptography and Security*. New York:Springer: pages 457-458.

Krempl, S. (2015): 31C3: CCC-Tüftler hackt Merkels Iris und von der Leyens Fingerabdruck. *Heise Online*, December 28, https://www.heise.de/security/meldung/31C3-CCC-Tueftler-hackt-Merkels-Iris-und-von-der-Leyens-Fingerabdruck-2506929.html

Krempl, S. (2016): 'Terror Score': Ex-Bundesdatenschützer greift Innenminister an. *Heise Online*, October 7, https://www.heise.de/newsticker/meldung/Terror-Score-Ex-Bundesdatenschuetzer-greift-Innenminister-an-3343177.html

Krieg, G. (2015): No-fly nightmares: The program's most embarrassing mistakes DNA database. *CNN*, July 12, http://edition.cnn.com/2015/12/07/politics/no-fly-mistakes-cat-stevens-ted-kennedy-john-lewis/index.html

Kubicek, H.; Noack, T. (2010a): The path dependency of national electronic identities. A comparison of innovation processes in four European countries. *Identity in the Information Society (IDIS),* 3(1): 111-153.

Kubicek, H.; Noack, T. (2010b): Different countries-different extended comparison of the introduction of eIDs in eight European countries. *Identity in the Information Society (IDIS),* 3(1): 235-245.

Kuchler, H. (2014): Tech companies step up encryption in wake of Snowden. *Financial Times,* November 4, https://www.ft.com/content/3c1553a6-6429-11e4-bac8-00144feabdc0

Kupka, A. (2012): Facebook Timeline now mandatory for everyone. *Forbes Magazine,* January 24 2012, http://www.forbes.com/sites/annakupka/2012/01/24/facebook-timeline-now-mandatory-for-everyone/

Lachance, N. (2016): Facebook's Facial Recognition Software Is Different From The FBI's. Here's Why. *National Public Radio (NPR)*, May 18, http://www.npr.org/sections/alltechconsidered/2016/05/18/477819617/facebooks-facial-recognition-software-is-different-from-the-fbis-heres-why

Laszlo, A., Krippner, S. (1998): Chapter 3 – systems theories: their origins, foundations, and development. In Jordan, J. S. (ed.), *Systems theories and a priori aspects of perception. Advances in Psychology*, 126, Amsterdam: Elsevier Science. pages 47-74. Doi: doi.org/10.1016/S0166-4115(98)80017-4

Le Blond, S., Manils, P., Abdelberi, C., Kaafar, M. A., Claude Castelluccia, C., Legout, A., Dabbous, W. (2011): One Bad Apple Spoils the Bunch: Exploiting P2P Applications to Trace and Profile Tor Users. In *4th USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET '11), Mar 2011, Boston, US*. Available online at: https://hal.inria.fr/inria-00574178/en/

Le Grand, G., Barrau, E. (2012): Prior Checking, a Forerunner to Privacy Impact Assessments. In Wright, D., De Hert, P. (eds.), *Privacy Impact Assessment*, Springer: Dordrecht: pages 97-116.

Leber, J. (2013): How Wireless Carriers Are Monetizing Your Movements. *Technology Review*, April 12, https://www.technologyreview.com/s/513016/how-wireless-carriers-are-monetizing-your-movements/

LeCun, Y., Benigo, Y., Hinton, G. (2015). Deep Learning. *Nature,* 521: 436-444.

Lee, D. (2012): Facebook surpasses one billion users as it tempts new markets. *BBC News*, October 5, http://www.bbc.co.uk/news/technology-19816709

Leenes, R. (2010): Context is everything – Sociality and Privacy in Online Social Network Sites. In Bezzi, M., Duquenoy, P., Fischer-Hübner, S., Hansen, M., Zhang, G. (eds.), *Privacy and identity management for life*. IFIP AICT Vol. 320, Heidelberg/Berlin/New York: Springer: pages 48-65.

Leenes, R., Schallaböck, J. and Hansen, M. (2008): Prime (Privacy and Identity Management for Europe) White Paper. Third and final version, EU: The PRIME consortium.

Leimbach, T., Hallinan, D., Bachlechner, D., Weber, A., Jaglo, M., Hennen, L., Nielsen, R., Nentwich, M., Strauß, S., Lynn, T., Hunt, G. (2014): *Potential and Impacts of Cloud Computing Services and Social Network Websites – Study*. Report no. IP/A/STOA/FWC/2008-096/Lot4/C1/SC8; Science and Technology Options Assessment – European Parliamentary Research Service: Brussels

Leiner, B. M., Cerf, V. G., David D. Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., Postel, J., Roberts, L. G., Wolff, S. (2003): *Brief History of the Internet.* Internet Society. Available online at: http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet

Lemos, R. (2006): Private identities become a corporate focus. *SecurityFocus*, Feburary 20, http://www.securityfocus.com/news/11377

Leskovec, J., Horvitz, E. (2008): Worldwide buzz: Planetary-scale views on a large instant-messaging network. In *Proceedings of the 17th international conference on World Wide Web, April 21-25, Beijing:* pages 915-924.

Leskovec, J., Rajaraman, A., Ullman, J. D. (2014): *Mining of massive datasets.* Second edition, Cambridge: Cambridge University Press.

Lessig, L. (2006): *Code Version 2.0*. New York: Basic Books.

Lever, A. (2006): Privacy rights and democracy – a contradiction in terms? *Contemporary Political Theory*, 5(2): 142-162.

Leyden, J. (2016): Mastercard rolls out pay-by-selfie across Europe. *The Register*, October 5 2016, http://www.theregister.co.uk/2016/10/05/mastercard_selfie_pay/

Li, N., Li, T., Venkatasubramanian, S. (2007): t-Closeness: Privacy Beyond k-Anonymity and l-Diversity. In *Proceedings of the 23rd IEEE International Conference on Data Engineering (ICDE):* pages106-115. DOI: 10.1109/ICDE.2007.367856

Lips, A. Miriam B., Taylor, John A., Organ, J. (2009): Managing Citizen Identity Information in E-Government Service Relationships in the UK. *Public Management Review*, 11(6): 833-856.

Long, D. (2010): Eugene Kaspersky: introduce ID system for all Internet users. *PC authority*, May 17, https://www.pcauthority.com.au/News/174767,eugene-kaspersky-introduce-id-system-for-all-internet-users.aspx

Luepnitz, D. A. (2003): *Schopenhauer's porcupines. Intimacy and its dilemmas: five stories of psychotherapy.* New York: Basic Books.

Luhmann, N. (1991): *Soziale Systeme. Grundriß einer allgemeinen Theorie*. Suhrkamp Taschenbuch Wissenschaft, Frankfurt. 4. Auflage (Erste Auflage 1984).

Lyon, D. (1994): *The Electronic Eye: The Rise of Surveillance Society.* Minneapolis: University of Minnesota Press.

Lyon, D. (2001): *Surveillance Society: Monitoring Everyday Life.* Oxford: University Press.

Lyon, D. (2003): *Surveillance as Social Sorting. Privacy, Risk and Automated Discrimination*. London: Routlegde.

Lyon, D. (2006): *Theorizing surveillance. The panopticon and beyond*. Cullompton: Willan Publishing.

Lyon, D. (2009): *Identifying Citizens – ID Cards as Surveillance*. Cambridge: Polity Press.

Lyon, D. (2014): Surveillance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data & Society,* July-December: 1-13. Available online at: http://journals.sagepub.com/doi/abs/10.1177/2053951714541861

MacAskill, E., Borger, J., Hopkins, N., Davies, N., Ball, J. (2013): Mastering the internet: how GCHQ set out to spy on the world wide web. *The Guardian*, June 21, https://www.theguardian.com/uk/2013/jun/21/gchq-mastering-the-internet

Machanavajjhala, A., Gehrke, J., Kifer, D. (2007): ℓ-Diversity: Privacy Beyond k-Anonymity. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1). DOI: 10.1145/1217299.1217302 Available online at: http://arbor.ee.ntu.edu.tw/archive/ppdm/Anonymity/MachanavajjhalaLP06.pdf

Madden, M. (2014): Public Perceptions of Privacy and Security in the Post-Snowden Era. Report. *Pew Research Center Internet & Technology*. Available online at: http://www.pewinternet.org/2014/11/12/public-privacy-perceptions/

Madden, M., Lenhart, A., Cortesi, S., Gasser, U., Duggan, M., Smith, A., Beaton, M. (2013): *Teens, social media and privacy.* Research Report, Pew Research Center Internet & Technology. Available online at: http://www.pewinternet.org/2013/05/21/teens-social-media-and-privacy/

Martin, A. K., De Adrane, N. N. G. (2013): Friending the taxman: On the use of social networkingservices for government eID in Europe. *Telecommunications Policy*, 37(9): 715-724.

Marx, G. T. (2001): Identity and Anonymity : Some Conceptual Distinctions and Issues for Research. In Caplan, J., Torpey, J. (eds.) *Documenting Individual Identity.* Princeton: Princeton University Press. Available online at: http://web.mit.edu/gtmarx/www/identity.html

Marx, G. T. (2015): Surveillance studies. In Wright, J. D. (ed.), *International Encylopedia of the Social & Behavioral Sciences*, Second Edition, 23: pages 7133-741. DOI: 10.1016/B978-0-08-097086-8.64025-4

Marx, G.T. (2002): What's New About the "New Surveillance?" Classifying for Change and Continuity. *Surveillance & Society*, 1(1): 9-29.

Matikainen, J. (2015): Motivations for content generation in social media. Participations – *Journal of Audience & Reception Studies*, 12 (1): 41-58.

Mayer-Schönberger, V. (2009): Delete: *The Virtue of Forgetting in the Digital Age*. Princeton: Princeton University Press.

Mayer-Schönberger, V., Cukier, K. (2013): *Big Data: A Revolution that Will Transform how We Live, Work and Think*. New York: Houghton Mifflin Harcourt.

McCallister, E., Grance, T., Scarfone, K. (2010): *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII). Recommendations of the National Institute of Standards and Technology (NIST), Special publication 800-122*. U.S. Department of Commerce. Available online at: https://www.nist.gov/publications/guide-protecting-confidentiality-personally-identifiable-information-pii

McCarthy, K. (2016): Privacy advocates rail against US Homeland Security's Twitter, Facebook snooping. *The Register,* August 23, http://www.theregister.co.uk/2016/08/23/homeland_security_social_media_snooping/

McCarthy, K. (2017a): Trump signs 'no privacy for non-Americans' order – what does this mean for the rest of us? *The Register*, January 26, https://www.theregister.co.uk/2017/01/26/trump_blows_up_transatlantic_privacy_shield/

McCarthy, K. (2017b): Russia, China vow to kill off VPNs, Tor browser. *The Register*, July 11, https://www.theregister.co.uk/2017/07/11/russia_china_vpns_tor_browser/

McGoogan, C. (2016): Dark web browser Tor is overwhelmingly used for crime, says study. *The Telegraph*, February 2, http://www.telegraph.co.uk/technology/2016/02/02/dark-web-browser-tor-is-overwhelmingly-used-for-crime-says-study/

McLuhan, M. (1964): *Understanding Media. The Extensions of Man.* London: Routledge & Kegan Paul.

McQuail, D. (2010): *Mass communication theory: an introduction*. Sixth edition 2005 (first edition published 1983). London: Sage Publications

Mearian, L. (2012): By 2020, there will be 5,200 GB of data for every person on Earth. *Computerworld*, December 11, http://www.computerworld.com/article/2493701/data-center/by-2020--there-will-be-5-200-gb-of-data-for-every-person-on-earth.html

Menezes, A., van Oorschot, P. C., Vanstone, S. A. (1996): *Handbook of Applied Cryptography*. First edition 1996. London: CRC Press.

Metz, R. (2012): More passwords, more problems. *Technology Review*, September 21, http://www.dailymail.co.uk/sciencetech/article-2174274/No-wonder-hackers-easy-Most-26-different-online-accounts--passwords.html

Micciancio, D. (2010): Technical Perspective: A First Glimpse of Cryptography's Holy Grail. *Communications of the ACM*, 53(3): 96. Available online at: http://www.yildiz.edu.tr/~aktas/courses/CE-0112822/06-05-2-2.pdf

Milan, S. (2015): When Algorithms Shape Collective Action: Social Media and the Dynamics of Cloud Protesting. *Social Media + Society (SM+S),* 1(2) (Online first December 30): 1-10. Available online at: http://journals.sagepub.com/doi/pdf/10.1177/2056305115622481

Milgram (1967): The small world problem. *Psychology Today*, 2(1): 60-67.

Monckton, P. (2017): Is Google About To Start Sharing Your Facial Recognition Data? *Forbes Magazine*, April 28, https://www.forbes.com/sites/paulmonckton/2017/04/28/google-photos-ramps-up-facial-recognition/

Moor, J. H. (1998): reason, relativity and responsibility in computer ethics. *Computers and Society*, 28(1): 14-21.

Moore Jr., B. (1984): *Privacy: Studies in social and cultural history*. New York: Pantheon Books.

Motwani, R., Y. Xu. (2007): Efficient algorithms for masking and finding quasi-identifiers. In *Proceedings of the Conference on Very Large Data Bases (VLDB), September 23-28, Vienna, Austria*. pages 83-93.

Nanz, P. (2007): Multiple Voices: An Interdiscursive Concept of the European Public Sphere. In Fossum, J. E., Schlesinger, P. and Kvaerk, G. (eds.), *Public Sphere and Civil Society? Transformations of the European Union*. ARENA Report Series No 2/07. Centre for European Studies, Oslo, Norway. pages 11-28.

Narayanan, A., Shmatikov, V. (2009): De-anonymizing social networks. In *Proceedings of the 30th IEEE Symposium on Security and Privacy*. pages 173-187. DOI: 10.1109/SP.2009.22

Naumann, L., Hobgen, G. (2009): *Privacy features of European eID Card Specifications*. Position paper, European Network and Information Security Agency ENISA. Available online at: https://www.enisa.europa.eu/publications/eid-cards-en/at_download/fullReport

Nentwich, M., König, R. (2012): *Cyberscience 2.0. Research in the Age of Digital Social Networks*. Series Interactiva, Vol. 11, Frankfurt/New York: Campus.

Newman, L., H. (2016): Hack Brief: Hackers Breach a Billion Yahoo Accounts. A Billion. *Wired*, December 14, https://www.wired.com/2016/12/yahoo-hack-billion-users/

Nilizadeh, S., Kapadia, A., Ahn, Y. (2014): community-enhanced de-anonymization of online social networks. In Ahn, G., Yung, M., Li, N. (eds.), *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security.* ACM: New York. pages 537-548.

Nissenbaum, H. (2010): *Privacy in context – technology, policy, and the integrity of social Life*. Stanford: Stanford University Press.

O'Connell, A. (2014): Facebook Buys WhatsApp: what does it mean for WhatsApp privacy? *The Online Privacy Blog*, March 5, https://www.abine.com/blog/2014/whatsapp-privacy-and-facebook-acquisition/

O'Donnell, A. (2017): How to Disable Facebook's Facial Recognition Feature. *Lifewire*, June 26, https://www.lifewire.com/how-to-disable-facebooks-facial-recognition-feature-2487265

O'Mahony, J. (2012): Text messaging at 20: how SMS changed the world. *The Telegraph*, December 3, http://www.telegraph.co.uk/technology/mobile-phones/9718336/Text-messaging-at-20-how-SMS-changed-the-world.html

O'Reilly, T. (2005): What is Web 2.0? Design patterns and business models for the next generation of software. *O'Reilly Media*, September 30, http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html

OECD – Organization for Economic Co-Operation and Development (2013b): *The OECD Privacy Framework*. OECD Publishing. Available online at: https://www.oecd.org/sti/ieconomy/privacy-guidelines.htm

OECD – Organization for Economic Co-Operation and Development (2013a): Building Blocks for Smart Networks. *OECD Digital Economy Papers*, No. 215, DSTI/ICCP/CISP(2012)3/FINAL, Paris: OECD Publishing. Available online at: http://dx.doi.org/10.1787/5k4dkhvnzv35-en

OECD – Organization for Economic Co-Operation and Development (2011): *National Strategies and Policies for Digital Identity Management in OECD Countries*. OECD Digital Economy Papers, No. 177, Paris: OECD Publishing. Available online at: http://dx.doi.org/10.1787/5kgdzvn5rfs2-en

OECD – Organization for Economic Co-Operation and Development (2004): *The Security Economy*. Organisation for Economic Co-operation and Development, Paris: OECD Publishing. Available online at: http://www.oecd.org/futures/16692437.pdf

Oetzel, M. C., Spiekermann, S., Grüning, I., Kelter, H., Mull, S. (2011): *Privacy impact assessment guidline*. Bundesamt für Sicherheit in der Informationstechnik – BSI.

Oracle (2015): Oracle Buys Datalogix Creates the World's Most Valuable Data Cloud to Maximize the Power of Digital Marketing. Company presentation, January 23, http://www.oracle.com/us/corporate/acquisitions/datalogix/general-presentation-2395307.pdf

Oracle (2016): Oracle Data Cloud. T*he data source magazine*, Issue No. 3, Fall 2016, https://cloud.oracle.com/opc/saas/resources/the-data-source-magazine-fall-2016.pdf

Osborne, C. (2016): How to access Tor, even when your country says you can't. *ZDNet*, August 4, http://www.zdnet.com/article/how-to-dance-around-censorship-and-access-tor-even-when-blocked/

Owen, T. (2004): Challenges and opportunities for defining and measuring human security, Human Rights, Human Security and Disarmament, *disarmanet forum 2004*, 3: 15-24.

Panzarino, M. (2011): Facebook introduces radical new profile design called Timeline: The story of your life. *The Next Web*, September 22, http://thenextweb.com/facebook/2011/09/22/facebook-introduces-timeline-the-story-of-your-life/

Papacharissi, Z. Gibson, P., L. (2011): Fifteen Minutes of Privacy: Privacy, Sociality, and Publicity on Social Network Sites. In Trepte, S. and Reinecke, L. (eds.), *Privacy Online – Perspectives on Self-Disclosure in the Social Web*. Berlin/Heidelberg: Springer: pages 75-89.

Pariser, E. (2011): *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin Press.

Parsons, T. (1991): *The social system*. New edition (2005), first published 1991, (first edition 1951), Sociology Classics, London: Routledge.

Patrizio, A. (2016): ICQ, the original instant messenger, turns 20. *NetworkWorld*, November 18, http://www.networkworld.com/article/3142451/software/icq-the-original-instant-messenger-turns-20.html

Pavone, V., Degli Esposti, S. (2012): Public assessment of new surveillance-orientated security technologies: Beyond the trade-off between privacy and security. *Public Understanding of Science*, 21 (5): 556-572.

Perrin, A. (2015): Social media usage 2005-2015. Report, Pew Research Center Internet & Technology, October 8, http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/

Perry, W.L., McInnis, B., Price, C.C.; Smith, S., Hollywood, J.S. (2013): *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. RAND Cooperation: Santa Monica, CA, USA. Available online at: http://www.rand.org/pubs/research_reports/RR233

Pew (2014): World Wide Web timeline. Pew Research Center Internet & Technology, March 11, http://www.pewinternet.org/2014/03/11/world-wide-web-timeline/

Pfitzmann, A. and Borcea-Pfitzmann, K. (2010): Lifelong Privacy: Privacy and Identity Management for Life. In Bezzi, M., Duquenoy, P., Fischer-Hübner, S., Hansen, M., Zhang, G. (eds.), *Privacy and Identity Management for Life, 5th IFIP/PrimeLife, International Summer School,* IFIP AICT 320, Heidelberg: Springer: pages 1-17.

Pfitzmann, A., Hansen, M. (2010): *A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity*. Version 0.34. Available online at: http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf

Poster, M. (1990): *The Mode of Information*. Chicago: University of Chicago Press.

Prabhakar, S., Pankanti, S., Jain, A. K. (2003): Biometric recognition: security and privacy concerns. *IEEE Security & Privacy*, 99 (2): 33-42.

Preneel, B. (2005): Collision attack. In van Tilborg, H.C.A. (ed.), *Encyclopedia of Cryptography and Security*. New York: Springer: pages 220-221.

Prevelakis, V., Spinellis, D. (2001): Sandboxing Applications. In *USENIX 2001 Technical Conference Proceedings: FreeNIX Track*, pages 119-126.

Raab, C. D. (2006): Social and Political Dimensions of Identity. In Fischer-Hübner, S., Duquenoy, P., Zuccato, A., Martucci, L. (eds.), *Proceedings of the 3rd IFIP/FIDIS International Summer School on The Future of Identity in the Information Society, Karlstad University, Sweden, August 4–10*. Heidelberg: Springer: pages 3-19.

Raab, C., D. (2009): Identity: difference and categorisation. In Kerr, I., Steeves, V., Lucock, C. (eds.), *Lessons from the identity trail: anonymity, privacy and identity in a networked society.* Oxford: Oxford University Press: pages 227-244.

Raab, C., D., Wright, D. (2012): Surveillance: extending the limits of privacy impact assessment. In Wright, D., De Hert, P. (eds.), *Privacy Impact Assessment*, Dordrecht: Springer: pages 363-383.

Raeburn, S. (2013): Fury at Facebook as login requests "Government ID" from users. *The Drum News*, October 29, http://www.thedrum.com/news/2013/10/29/fury-facebook-login-requests-government-id-users#0SAYzdXek51viRHm.99

Rannenberg, K., Royer, D., Deuker, A. (eds.) (2009): *The future of identity in the information society – Challenges and opportunities*. Berlin: Springer.

Reardon, S. (2012): FBI launches $1 billion face recognition project. *New Scientist*, September 7, https://www.newscientist.com/article/mg21528804-200-fbi-launches-1-billion-face-recognition-project/

Redecker, C., Ala-Mutka, K., and Punie, Y. (2010): *Learning 2.0 - The Impact of Social Media on Learning in Europe*. Policy Brief, Joint Research Centre of the European Commission, JRC56958, Luxembourg: European Communities Publications. Available online at: http://www.ict-21.ch/com-ict/IMG/pdf/learning-2.0-EU-17pages-JRC56958.pdf

Reiman, J. H. (1976): Privacy, Intimacy, and Personhood. *Philosophy & Public Affairs*, 6(1): 26-44. Available online at: https://www.jstor.org/stable/2265060?seq=1#page_scan_tab_contents

Reitman, R. (2012): What Actually Changed in Google's Privacy Policy. *EPIC*, Feb. 1, https://www.eff.org/deeplinks/2012/02/what-actually-changed-google%27s-privacy-policy

Ricoeur, P. (1992): *Oneself as Another*. (Translated by Kathleen Blamey). Chicago: University of Chicago Press.

Riley, S. (2006): Password Security: What Users Know and What They Actually Do. *Usability News*, February 14, Software Usability Research Laboratory, Wichita State University, http://usabilitynews.org/password-security-what-users-know-and-what-they-actually-do/

Robinson, N., Bonneau, J. (2014): Cognitive disconnect: understanding Facebook Connect login permissions. In Sala, A., Goel, A., Gummadi, K., P. (eds.), *Proceedings of the 2nd ACM conference on Online social networks (COSN), Dublin, Ireland, October 1-2,* pages 247-258. DOI: 10.1145/2660460.2660471

Rogers, R. (2009): The Googlization Question, and the Inculpable Engine. In Stalder, F. and Becker, K. (eds.). *Deep Search: The Politics of Search Engines*. Edison, NJ:. Innsbruck: StudienVerlag/Transaction Publishers: pages 173-184.

Rötzer, F. (2014): Wer seine Privatsphäre schützt, ist für die NSA ein Extremist. *Telepolis*, July 3, http://www.heise.de/tp/artikel/42/42165/1.html

Rouvroy, A., Poullet, Y. (2009): The right to informational self-determination and the value of self-development: reassessing the importance of privacy for Democracy.

In Gutwirth, S. et al. (eds.), *Reinventing Data Protection?* Dordrecht: Springer: pages 45-76.

Rubinstein, I. S. (2011): Regulating Privacy by Design. *Berkeley Technology Law Journal*, 26(3): 1409-1456. Available online at: http://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1917&context=btlj

Ruddick, G. (2016): Admiral to price car insurance based on Facebook posts. *The Guardian*, November 2, https://www.theguardian.com/technology/2016/nov/02/admiral-to-price-car-insurance-based-on-facebook-posts

Rundle, M., Blakley, B., Broberg, J., Nadalin, A., Olds, D., Ruddy, M., Guimarares, M. T. M., Trevithick, P. (2008): At a crossroads: 'Personhood' and digital identity in the information society. *STI Working paper* 2007/7, no. JT03241547 29-Feb-2008, Directorate for Science, Technology and Industry, OECD Publishing. Available online at: http://www.oecd.org/dataoecd/31/6/40204773.doc

Ruoti, S., Andersen, J., Zappala, D., Seamons, K. (2016): Why Johnny still, still can't encrypt: Evaluating the Usability of a Modern PGP Client. Version 2, January 13. *ArXiv ePrint*, Cornell University, Available online at: https://arxiv.org/abs/1510.08555

Russell, J. (2016): Facebook is testing social commerce payments in Southeast Asia. *Techcrunch*, June 9, https://techcrunch.com/2016/06/09/facebook-is-testing-social-commerce-payments-in-southeast-asia/

Salimkhan, G., Manago, A., & Greenfield, P. (2010). The Construction of the Virtual Self on MySpace. Cyberpsychology: *Journal of Psychosocial Research on Cyberspace*, 4(1), article 1. Available online at: http://www.cyberpsychology.eu/view.php?cisloclanku=2010050203

Sapelova,S., Jerman-Blažič, B. (2014): Privacy Issues in Cross-Border Identity Management Systems: Pan-European Case. In Hansen, M., Hoepman, J., Leenes, R., Whitehouse, D. (eds.), *Privacy and Identity Management for Emerging Services and Technologies: 8th IFIP/Primelife International Summer School, Nijmegen, The Netherlands, June 17-21 2013*, Revised Selected Papers, IFIP AICT 421, Springer: pages 214-223.

Sartre, J. P. (2014): *Das Sein und das Nichts. Versuch einer phänomenologischen Ontologie.* 18. Auflage Januar 2014 (Original 1943). Hamburg: Rowohlt.

Sayer, P. (2013): Fujitsu names UniCredit as first European customer for palm-scan authentication. *NetworkWorld*, March 4, http://www.networkworld.com/article/2164100/byod/fujitsu-names-unicredit-as-first-european-customer-for-palm-scan-authentication.html

Schneier B., (2006b): No-Fly list. *Schneier on Security*, October 6, https://www.schneier.com/blog/archives/2006/10/nofly_list.html

Schneier, B. (2003): *Beyond Fear: Thinking Sensibly about Security in an Uncertain World.* New York: Copernicus Books.

Schneier, B. (2006a): The eternal value of privacy. Published in *Wired*. May 18, Available online at: *Schneier on Security*, https://www.schneier.com/essays/archives/2006/05/the_eternal_value_of.html

Schneier, B. (2014): Metata = Surveillance. *Schneier on Security*, March 13, https://www.schneier.com/blog/archives/2014/03/metadata_survei.html

Schrittwieser , S., Kieseberg, P., Echizen, I., Wohlgemuth, S., Sonehara, N., Weippl, E. (2011): An Algorithm for k-Anonymity-Based Fingerprinting. In Y. Q. Shi, H. J. Kim, F. Perez-Gonzalez (eds.), *Proceedings of the 10th international conference on Digital-Forensics and Watermarking – IWDW, LCNS 7128*, Heidelberg: Springer: pages 439-452. Available online at: https://www.sba-research.org/wp-content/uploads/publications/k_anonymity_algorithm_2011.pdf

Schwartz, P., M., Solove, D., J. (2011): The PII problem: privacy and a new concept of personally identifiable information. *New York University Law Review* (86): 1814-1894.

SGTF – Smart Grid Task Force (2014): *Data Protection Impact Assessment Template for Smart Grid and Smart Metering systems*. Expert Group 2: regulatory recommendations for privacy, data protection and cyber - security in the smart grid environment. Brussels, European Commission. Available online at: https://ec.europa.eu/energy/sites/ener/files/documents/2014_dpia_smart_grids_forces.pdf

Shannon, C., E., (1948): A mathematical theory of communication. *The Bell System Technical Journal*, 27(4): 623–656. Available online at: http://math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf

Shapiro, S. (2005): Agency theory. *Annual Review of Sociology*, 31: 263-284. http://www.annualreviews.org/doi/abs/10.1146/annurev.soc.31.041304.122159

Sharif, M., Bhagavatula, S., Bauer, L., Reiter, M. K. (2016): Accessorize to a Crime: Real and Stealthy Attacks on State-of-the-Art Face Recognition. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, October 24-28, Vienna, Austria*, pages 1528-1540. Available online at: https://www.cs.cmu.edu/~sbhagava/papers/face-rec-ccs16.pdf

Shirky, C. (2009): A Speculative Post on the Idea of Algorithmic Authority. *Blogpost*, November 15, http://www.shirky.com/weblog/2009/11/a-speculative-post-on-the-idea-of-algorithmic-authority/

Singh, N., Lehnert, K. and Bostick, N. (2012): global social media usage – insights into reaching consumers worldwide. *Wiley Online Library*, DOI: 10.1002/tie.21493 Available online at: http://beople.es/wp-content/uploads/2013/04/GLOBAL-SOCIAL-MEDIA-USAGE-INSIGHTS-INTO-REACHING-CONSUMERS-WORLDWIDE.pdf

Skinner, J. (2012): Social Media and Revolution: The Arab Spring and the Occupy Movement as Seen Through Three Information Studies Paradigms. *Sprouts Working Papers on Information Systems* 483. Available online at: http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1482&context=sprouts_all

Smith, A. (2011): Why American use social media. Report, *Pew Research Center Internet & Technology*, November 15, http://www.pewinternet.org/2011/11/15/why-americans-use-social-media/

Smith, A., Sterling, A., Berkhout, F. (2005): The governance of sustainable socio-technical transitions. *Research policy*, (34): 1491-1510.

Solove, D. J. (2004): *The digital person – technology and privacy in the information age*. New York and London: New York University Press.

Solove, D. J. (2006): a taxonomy of privacy. *University of Pennsylvania Law Review*, 154 (3): 477-560.

Solove, D. J. (2011): *Nothing to hide: the false tradeoff between privacy and security*. New Haven/London: Yale University Press.

Spiekermann, S., Acquisti, A., Böhme, R., Hui, K.L. (2015): the challenges of personal data markets and privacy . *Electronic Markets* 25(2): 161-167.

Spivack, N. (2013): The post-privacy world. *Wired*, July, https://www.wired.com/insights/2013/07/the-post-privacy-world/

Stanton, J. M. (2008): ICAO and the biometric RFID passport: History and analysis. In Bennett, C. J., Lyon, D. (eds.), *Playing the identity card - surveillance, security and identification in global perspective*, London /New York: Routledge: pages 253-267.

StatB – Statistic Brain (2016): Google Plus demographics & statistics, *Statistic Brain*, September 4, http://www.statisticbrain.com/google-plus-demographics-statistics/

Statista (2017a): Facebook's advertising revenue worldwide from 2009 to 2016 (in million U.S. dollars). *Statista*, http://www.statista.com/statistics/271258/facebooks-advertising-revenue-worldwide/

Statista (2017b): Advertising revenue of Google websites from 2001 to 2016 (in billion U.S. dollars). *Statista*, http://www.statista.com/statistics/266242/advertising-revenue-of-google-sites

Statista (2017c): Number of monthly active Facebook users worldwide as of 1st quarter 2017 (in millions), *Statista*, https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

Stefanone, M. A., Lackaff, D., Rosen, D. (2010): The Relationship between Traditional Mass Media and 'Social Media': Reality Television as a Model for Social Network Site Behavior. *Journal of Broadcasting & Electronic Media*, 54(3): 508–525. DOI: 10.1080/08838151.2010.498851

Steinfield, C., Ellison, N. B., Lampe, C. (2008): Social capital, self-esteem, and use of online socialnetworksites: A longitudinal analysis, *Journal of Applied Developmental Psychology,* 29(6): 434-445.

Sterling, G. (2016): Nearly 80 percent of social media time now spent on mobile devices. *Marketing Land*, April 4, http://marketingland.com/facebook-usage-accounts-1-5-minutes-spent-mobile-171561

Stiglitz, J. E. (2002): Information and the Change in the Paradigm in Economics. The *American Economic Review*, 92 (3): 460-501.

Storm, D. (2015): ACLU: Orwellian Citizen Score, China's credit score system, is a warning for Americans. *Computerworld*, October 7, http://www.computerworld.com/article/2990203/security/aclu-orwellian-citizen-score-chinas-credit-score-system-is-a-warning-for-americans.html

Strauß, S. (2011): The Limits of Control - (Governmental) Identity Management from a Privacy Perspective. In Fischer-Hübner, S. Duquenoy, P., Hansen, M., Leenes, R., Zhang, G. (eds.), *Privacy and Identity Management for Life, 6th IFIP/PrimeLife International Summer School, Helsingborg, Sweden, August 2-6 2010, Revised Selected Papers*. Dordrecht: Springer: pages 206-218.

Strauß, S. (2015a): Datafication and the Seductive Power of Uncertainty – A Critical Exploration of Big Data Enthusiasm. *Information*, 6: 836-847. Available online at: http://www.mdpi.com/2078-2489/6/4/836/pdf

Strauß, S. (2015b): *Citizen summits on privacy, security and surveillance: synthesis report. Deliverable 6.10 of the SurPRISE project*. Available online at: http://surprise-project.eu/wp-content/uploads/2015/02/SurPRISE-D6.10-Synthesis-report.pdf

Strauß, S. (2017a): A game of hide and seek? Unscrambling the trade-off between privacy and security. In Friedewald et al., (eds.), *Surveillance, privacy, and security – citizens' perspectives.* London/New York: Roudledge: pages 255-272.

Strauß, S. (2017b): Privacy Analysis – Privacy Impact Assessment. In Hansson, S. O. (ed.): *The Ethics of Technology – Methods and Approaches.* London/New York: Rowman&Littlefield International: pages 143-156.

Strauß, S., Aichholzer, G. (2010): National Electronic Identity Management: The Challenge of a citizen-centric Approach beyond Technical Design, *In International Journal on Advances in Intelligent Systems,* 3(1/2): 12-23.

Suhl, S. O. (2012): Google führt Dienste trotz Datenschutzbedenken zusammen. *Heise Online*, 01. März, https://www.heise.de/newsticker/meldung/Google-fuehrt-Dienste-trotz-Datenschutzbedenken-zusammen-1446292.html

Sun, Y.,Edmundson, A., Vanbever, L., Li, O. (2015): RAPTOR: Routing attacks on privacy in Tor. In *Proceedings of the 24th USENIX Security Symposium, Washington D.C.,* pages 271-286. Available online at: https://www.usenix.org/system/files/conference/usenixsecurity15/sec15-paper-sun.pdf

Sweeney, L. (2002): k-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, 10(5): 557-570. Available online at: https://epic.org/privacy/reidentification/Sweeney_Article.pdf

Taigman, Y., Yang, M., Ranzato, M. A., Wolf, L. (2014): DeepFace: Closing the Gap to Human-Level Performance in Face Verification. *Facebook Research*, June 24, available online at: https://research.facebook.com/publications/480567225376225/deepface-closing-the-gap-to-human-level-performance-in-face-verification/

Toor, A. (2016): Facebook begins tracking non-users around the internet. *The Verge*, May 27, http://www.theverge.com/2016/5/27/11795248/facebook-ad-network-non-users-cookies-plug-ins

Travis, A. (2009): Innocent suspects' profiles still reaching DNA database, *The Guardian*, October 28, http://www.guardian.co.uk/politics/2009/oct/28/dna-database-innocent-profiles

Treiger-Bar-Am, K. (2008): In defense of autonomy: An ethic of care. *New York University Journal of Law & Liberty*, 3(2): 548-590.

Trenz, H. J. (2008): In search of the European Public Sphere. Between Normative Overstretch and Empirical Disenchantment. *ARENA working paper* 12/08. Available online at: http://www.sv.uio.no/arena/english/research/publications/arena-working-papers/2001-2010/2008/wp08_12.pdf

Turchin, V. F. (1977): *The Phenomenon of Science: a cybernetic approach to human evolution*. New York: Columbia University Press.

Tynan, D. (2016): Facebook v Adblock: the anti-ad empire strikes back. *The Guardian*, August 11, https://www.theguardian.com/technology/2016/aug/11/facebook-advertising-changes-adblockers-strike-back

UDHR (1948): *The Universal Declaration of Human Rights*. Available online at: http://www.un.org/en/universal-declaration-human-rights/index.html

UN – United Nations (1994): New dimensions of Human Security. Human development report 1994, United Nations Development Programme, New York: Oxford University Press.

Ungerleider, N. (2012): Department of Homeland Security Tells Congress Why It's Monitoring Facebook, Twitter, Blogs. *FastCompany*, February 2, https://www.fastcompany.com/1816814/department-homeland-security-tells-congress-why-its-monitoring-facebook-twitter-blogs

Vaas, L. (2013): Nordstrom tracking customer movement via smartphones' WiFi sniffing. *Naked Security*, May 9, https://nakedsecurity.sophos.com/2013/05/09/nordstrom-tracking-customer-smartphones-wifi-sniffing/

Valkenburg, G. (2015): Privacy Versus Security: Problems and Possibilities for the Trade-Off Model. In Gutwirth, S., Leenes, R., De Hert, P. (eds.), *Reforming European Data Protection Law*, 20 of the series Law, Governance and Technology. Dordrecht: Springer: pages 253-269.

Van Dijk, J. (2014): Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, 12(2): 197-208.

Varela, F., Maturana, H., Uribe, R. (1974): Autopoiesis: The Organization of Living Systems, Its Characterization, and a Model. *Biosystems*, 5(4): 187-196.

Varela, F.J. (1997): Patterns of Life: Intertwining Identity and Cognition. *Brain and Cognition*, 34(1): 72-87.

Verbeek, P. (2011): Subject to technology – on autonomic computing and human autonomy, In Hildebrandt, M., Rouvroy A. (2011): *Law, Human agency and autonomic computing*. London/New York: Routledge: pages 27-45.

Warren, S. D., Brandeis, L. D. (1890): The Right to Privacy. *Harvard Law Review* 193 (1890), IV Dec. 15 1890, No. 5. Available online at: http://faculty.uml.edu/sgallagher/Brandeisprivacy.htm

Waters, J. K. (2004): The ABCs of identity management. *CSOOnline*, January 1, http://www.csoonline.com/article/2120384/identity-management/the-abcs-of-identity-management.html

Watson, S. (2011): The 'human' as referent object? Humanitarianism as securitization. *Security Dialogue*, 42(1): 3-20.

WEF – World Economic Forum (2015): Addressing barriers to digital trade. E15 Expert Group on the Digital Economy – Strengthening the global trade and investment system for sustainable development, Think piece, December 2015. Colongy/Geneva: World Economic Forum. Available online at: http://e15initiative.org/wp-content/uploads/2015/09/E15-Digital-Ahmed-and-Aldonas-Final.pdf

Wehage, J. C. (2013): *Das Grundrecht auf Gewährleistung der Vertraulichkeit und Integrität informationstechnischer Systeme und seine Auswirkungen auf das Bürgerliche Recht*. Dissertation, Universitätsverlag Göttingen. Available online at: http://www.univerlag.uni-goettingen.de/bitstream/handle/3/isbn-978-3-86395-123-8/Wehage_Diss.pdf

Wei, M., Grupp, L., M., Spada, F., E., Swanson, S. (2011): Reliably erasing data from flash-based solid state drives. In *9th USENIX Conference on File and Storage Technologies, San Jose, February 15-17*. Available online at: https://www.usenix.org/legacy/events/fast11/tech/full_papers/Wei.pdf

Weinger, M. (2016): Snowden's Impact Fades After Three Years. *The Cypher Brief*, June 5, https://www.thecipherbrief.com/article/exclusive/north-america/snowden%E2%80%99s-impact-fades-after-three-years-1089

Weisbuch, M., Ivcevic, Z., and Ambady, N., (2009): On being liked on the web and in the "real world": Consistency in first impressions across personal webpages and spontaneous behavior. *Journal of Experimental Social Psychology*, 45(3): 573-576.

Weiser, M. (1991): The computer of the 21st century. *Scientific American*, 265(3): 94-104. Available online at: https://www.lri.fr/~mbl/Stanford/CS477/papers/Weiser-SciAm.pdf

Westin, A. (1967): *Privacy and freedom*. New York: Atheneum.

WH – The White House (2011): *National Strategy for Trusted Identities in Cyberspace – Enhancing Online Choice, Efficiency, Security, and Privacy*. April 2011, The White House, Washington D.C. Available online at: https://www.hsdl.org/?view&did=7010

White, H. C. (2008): *Identity and Control – how social formations emerge.* Second edition, Princeton/Oxford: Princeton University Press.

Whitley, E. A., Gal, U., Kjaergaard, A. (2014): Who do you think you are? A review of the complex interplay between information systems, identification and identity. *European Journal of Information Systems*, 23(1): 17-35.

Whitley, E., Hosein, G. (2010): *Global Challenges for Identity Policies*. London: Palgrave Macmillan.

Wiener, N. (1954): *The Human Use of Human Beings – Cybernetics and Society*. (first published 1950, reprint of revised edition of 1954). Boston: Da Capo Press.

Wilson, C. (2010): *Biometric Modalities. Vein Pattern Recognition: A Privacy-Enhancing Biometric.* C. Wilson, New York: CRC Press.

Wimmer, J. (2009): The Publics behind Political Web Campaigning. The Digital Transformation of 'Classic' Counter-Public Spheres. In Baringhorst, S., Kneip, V., Niesyto, J. (eds.), *Political Campaigning on the Web*, Bielefeld: transcript: pages 31-51.

Wolchok, S., Hofmann, O.S., Heninger, N., Felten, E.W., Halderman, J.A., Rossbach, C.J., Waters, B., Witchel, E. (2010): Defeating Vanish with Low-Cost Sybil Attacks Against Large DHTs. In *Proceedings of the 17th Network and Distributed System Security Symposium (NDSS), San Diego, Feburary 28-March 3*. Available online at: https://www.internetsociety.org/sites/default/files/wol.pdf

Wondracek, G., Holz, T., Kirda, E., Kruegel, C. (2010): A practical attack to de-anonymize social network users. In *IEEE Symposium on Security and Privacy (SP), Washington D.C.*. pages 223-238. DOI: 10.1109/SP.2010.21 Available online at: https://www.syssec.rub.de/media/emma/veroeffentlichungen/2011/06/07/deanonymizeSN-Oakland10.pdf

Woodman, S. (2017): Palantir provides the engine for Donald Trump's deportation machine. *The Intercept*, March 2, https://theintercept.com/2017/03/02/palantir-provides-the-engine-for-donald-trumps-deportation-machine/

WP29 – Article 29 Data Protection Working party (2010): Opinion 1/2010 on the concepts of "controller" and "processor". 00264/10/EN WP 169. Available online at:

WP29 – Article 29 Data Protection Working Party (2016): Opinion 01/2016 on the EU–U.S. Privacy Shield draft adequacy decision. 16/EN WP 238. Available online at: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2016/wp238_en.pdf

Wright and De Hert (2012c): Findings and Recommendations. In Wright/De Hert (eds.), *Privacy Impact Assessment*. Law, Governance and Technology Series 6, Dordrecht: Springer: pages 445-481.

Wright, D., De Hert (2012b): Introduction to Privacy Impact Assessment. In Wright/De Hert (eds.), *Privacy Impact Assessment*. Law, Governance and Technology Series 6, Dordrecht: Springer: pages 3-32.

Wright, D., De Hert, P. (eds.) (2012a): Privacy Impact Assessment. Law, Governance and Technology Series 6, Dordrecht: Springer.

Wright, D., Kreissl, R. (eds.) (2015): *Surveillance in Europe*. London/New York: Routledge.

Wright, D., Rodrigues, R., Raab, C., Jones, R., Székely, I., Ball, K., Bellanova, R., Bergersen, S. (2015): Questioning surveillance. *Computer Law & Security Review*, 31(2): 280-292.

York, J. C. (2016): Censorship on Social Media: 2016 in Review. *EFF – Electronic Frontier Foundation,* December 24, https://www.eff.org/de/deeplinks/2016/12/censorship-social-media-2016-review

ZDNet (2013): PRISM: Here's how the NSA wiretapped the internet. *ZDNET*, June 8 2013, http://www.zdnet.com/article/prism-heres-how-the-nsa-wiretapped-the-internet/

Zhang Y., Zhou Y. (2006): Transparent Computing: A New Paradigm for Pervasive Computing. In Ma J., Jin H., Yang L.T., Tsai J.JP. (eds.), *Ubiquitous Intelligence and Computing. UIC 2006*. LNCS Vol. 4159. Berlin/Heidelberg: Springer: pages 1-10.

Zimmer, M. (2014): Mark Zuckerberg's theory of privacy. *The Washington Post*, February 3, https://www.washingtonpost.com/lifestyle/style/mark-zuckerbergs-theory-of-privacy/2014/02/03/2c1d780a-8cea-11e3-95dd-36ff657a4dae_story.html

Zuiderveen-Borgeswius, F.J. (2016): Informed Consent. We Can Do Better to Defend Privacy. *IEEE Security & Privacy*, 13(2): 103-107.

Zyskind, G., Nathan, O., Pentland, A. (2015): Enigma: Decentralized Computation Platform with Guaranteed Privacy. White paper. *ArXiv ePrint*, Cornell University, Available online at: https://arxiv.org/abs/1506.03471

Appendix

List of Abbreviations

CCTV – closed circuit television
CRM – Customer Relationship Management
DPA – Data protection authority
DPD – Data Protection Directive
DPI – Deep packet inspection
eID – electronic identity
eIDMS – electronic identity management system
GCHQ – Government Communications Headquarters
GDPR – General Data Protection Regulation
IDM – identity management
IDMS – identity management system
IMEI – International Mobile Equipment Identity
IMSI – International Mobile Equipment Identity
ISD – informational self-determination
ISD – informational self-determination
LBS – location-based service
MAC – media access control
MST – Metasystem Transition
NSA – National Security Agency
PbD – Privacy by Design
PET – Privacy Enhancing Technology
PIA – Privacy impact assessment
PII – Personally identifiable information
sCCTV – smart CCTV
SIM – Subscriber Identity Module
SNS – Social Networking Sites
TII – Technically identifiable information

Curriculum Vitae

Personal: Stefan Strauß, born 1978 in Upper Austria, living in Vienna.
Contact: sstrauss(a)oeaw.ac.at
Research interests: Technology assessment, science and technology studies, societal impacts of ICTs, privacy, security and surveillance studies, identity management, e-democracy & governance in the information society, information and computer ethics, computing and philosophy

Professional experience

- Since 2008: Researcher at the Institute of Technology Assessment (ITA), Austrian Academy of Sciences
- 2008: Software and systems engineer at a Viennese Software company
- 2006-2007: Freelancer at a Cross-Media company in Berlin

Education

- 2007: Diploma in information systems/business informatics (Mag.rer.soc.oec) at the Johannes Kepler University of Linz, emphasis on information engineering and -management. Diploma thesis about E-Government, public sector modernisation and perspectives of electronic democracy.
- 2000-2007: Studies of information systems/business informatics at the Johannes Kepler University of Linz/Upper Austria
- 1993-1998: Commercial School (Handelsakademie), Gmunden
- 1989-1993: Grammar School (Bundesgymnasium), Gmunden

Research projects (overview)

- 02/2016-on -going: MATCH – Markets, Actors, Technologies. A comparative study of smart grid solutions in Austria, Norway and Denmark. Funded within EU ERA-net programme.
- 01/2015-10/2015: Digitaler Stillstand (digital breakdown) – examining the vulnerability of critical infrastructures and implications of large-scale failure of IT-systems. Funded by Austrian Academy of Sciences.
- 01/2012-01/2015: SurPRISE – Surveillance, Privacy and Security – a large scale participatory assessment of criteria and factors determining acceptability and acceptance of security technologies in Europe. Funded within the 7th EU framework programme.
- 04/2012-12/2013: SNS-Cloud – European Potential and Impacts of Cloud Computing Services and Social Network Websites. Funded by Science and Technology Options Assessment (STOA) Panel of the European Parliament.
- 08/2009-04/2014: E2Democracy: Comparative analysis of e-participation in local initiatives for sustainable development and reduction of CO_2-emissions. Funded by the Austrian Science Fund (FWF).
- 03/2010-09/2011: eDemEU – e-Democracy in Europe – Prospects of Internet-based political participation. Funded by Science and Technology Options Assessment (STOA) Panel of the European Parliament.

- 09/2008-12/2008: E-identity – systemic change of citizen identification by the government. Comparative country study of electronic identity management systems in four European countries. Funded by the Deutsche Volkswagen-Stiftung.

Publications (selection)

Strauß, S. (2017): A game of hide and seek? Unscrambling the trade-off between privacy and security. In Friedewald et al., (eds.), *Surveillance, privacy, and security – citizens' perspectives,* London/New York: Roudledge: pages 255-272.

Strauß, S. (2017): Privacy Analysis – Privacy Impact Assessment. In Hansson, S. O. (ed.): *The Ethics of Technology – Methods and Approaches.* London/New York: Rowman&Littlefield International: pages 143-156.

Aichholzer, G., Strauß, S. (2016): Electronic Participation in Europe. In R. Lindner et al. (eds.), *Electronic Democracy in Europe. Prospects and Challenges of E-Publics, E-Participation and E-Voting*. Springer: pages 55-132.

Aichholzer, G., Allhutter, D., Strauß, S. (2016): Comparing Output and Outcome of Citizen–Government Collaboration on Local Climate Targets. In Aichholzer, G. et al. (eds.), *Evaluating e-Participation. Frameworks, Practice, Evidence.* Cham/Heidelberg/New York/Dordrecht/London: Springer: pages 167-193.

Strauß, S. (2015): Datafication and the Seductive Power of Uncertainty – A Critical Exploration of Big Data Enthusiasm. *Information*, 6: 836-847.

Strauß, S. (2015): Towards a taxonomy of social and economic costs. In Wright, D., Kreissl, R. (eds.): *Surveillance in Europe*. London/New York: Routledge: pages 212-218.

Strauß, S., Nentwich, M. (2013): Social network sites, privacy and the blurring boundary between public and private spaces. *Science and Public Policy*, 40 (6): 724-732.

Aichholzer, G.; Allhutter, D.; Strauß, S. (2012): Using online carbon calculators for participation in local climate initiatives. In Tambouris et al. (eds.), LNCS 7444 (4th Intl. ePart Conference, Kristiansand, Norway, 3-5 September); Berlin/Heidelberg/New York: Springer: pages 85-96.

Strauß, S. (2011): The Limits of Control – (Governmental) Identity Management from a Privacy Perspective, In Fischer-Hübner, S. et al (eds.), *Privacy and Identity Management for Life: IFIP Advances in Information and Communication Technology Vol. 352*, Berlin/Heidelberg: Springer: pages 206-218.

Strauß, S., Aichholzer, G. (2010): National Electronic Identity Management: The Challenge of a citizen-centric Approach beyond Technical Design, *International Journal on Advances in Intelligent Systems*, 3 (1&2): 12-23.

Aichholzer, G., Strauß, S., (2010): The Austrian case: multi-card concept and the relationship between citizen ID and social security cards, *Identity in the Information Society (IDIS)*, 3(1): 65-85.

Aichholzer, G., Strauß, S., (2010): Electronic Identity Management in e Government 2.0: Exploring a System Innovation exemplified by Austria. In *Information Polity – An International Journal of Government and Democracy in the Information Age*, 15(1-2): 139-152.