TU WIEN Informatics

# Innovation and Ethics in Software Engineering and Internet Computing

DIPLOMARBEIT

zur Erlangung des akademischen Grades

**Diplom-Ingenieur**

im Rahmen des Studiums

**Software Engineering & Internet Computing**

eingereicht von

**Klaus Nigsch, BSc**
Matrikelnummer 01025991

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Ao.Univ.Prof.i.R. Dr.phil. Wolfgang Hofkirchner

Wien, 19. August 2025

| Klaus Nigsch | Wolfgang Hofkirchner |
|---|---|

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

TU WIEN Informatics

# Innovation and Ethics in Software Engineering and Internet Computing

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

**Diplom-Ingenieur**

in

**Software Engineering & Internet Computing**

by

**Klaus Nigsch, BSc**
Registration Number 01025991

to the Faculty of Informatics

at the TU Wien

Advisor: Ao.Univ.Prof.i.R. Dr.phil. Wolfgang Hofkirchner

Vienna, August 19, 2025

| Klaus Nigsch | Wolfgang Hofkirchner |
|---|---|

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

# Erklärung zur Verfassung der Arbeit

Klaus Nigsch, BSc

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.
Ich erkläre weiters, dass ich mich generativer KI-Tools lediglich als Hilfsmittel bedient habe und in der vorliegenden Arbeit mein gestalterischer Einfluss überwiegt. Im Anhang „Übersicht verwendeter Hilfsmittel“ habe ich alle generativen KI-Tools gelistet, die verwendet wurden, und angegeben, wo und wie sie verwendet wurden. Für Textpassagen, die ohne substantielle Änderungen übernommen wurden, habe ich jeweils die von mir formulierten Eingaben (Prompts) und die verwendete IT-Anwendung mit ihrem Produktnamen und Versionsnummer/Datum angegeben.

Wien, 19. August 2025

Klaus Nigsch

# Acknowledgements

I would like to express my greatest appreciation to Ao.Univ.-Prof.i.R. Dr.phil. Wolfgang Hofkirchner for providing his constructive input during all phases of this work. Not only were his efforts in the help of planning, structuring, and correction important, but his deep understanding of Information Science, especially in the topic of the Global Sustainable Information Society, was of great help in gaining a strong fundamental understanding required for this thesis.

Furthermore, I wish to express my deepest thanks to my family for their encouragement and support throughout the work and my study.

# Kurzfassung

Das Internet und Computer im Allgemeinen sind weltweit von zentraler Bedeutung. Innovationen und Fortschritte in der Informationstechnologie beeinflussen weltweit Sektoren wie Industrie, Bildung, Regierung und Gesellschaft. Konzepte wie Ubiquitous Computing, Verteilte Systeme, Mobile Computing, Soziale Netzwerke, Internet of Things, Big Data und Künstliche Intelligenz gestalten unser Leben und spielen eine zentrale Rolle für die menschliche Zukunft.

Softwareentwickler sollten Verantwortung für die Softwaresysteme, die sie planen und entwickeln, übernehmen. Zwar sind die Möglichkeiten für die Erstellung von Software, die die Menschheit positiv beeinflusst, gegeben, jedoch darf das Gefahr für Software, die der Menschheit schadet nicht vernachlässigt werden.

In den letzten Jahren haben globale Herausforderungen wie geopolitische Konflikte (z.B. Kriege), ökonomische Instabilität, soziale Ungerechtigkeiten und digitale Transformation immer wieder neue, raffiniertere Methoden verlangt, um festzustellen, wie Software ethisch im Bezug auf diese globalen Herausforderungen entwickelt werden kann.

Weil Softwareentwickler sektorübergreifend eine zentrale Position im Vormarsch der Innovation einnehmen, ist es wichtig, Innovationen kritisch zu hinterfragen. Eine große Bandbreite an ethischen Grundkonzepten hilft dabei, zu entscheiden, ob diese Innovationen dem Fortschritt der Menschheit dienen oder Schaden verursachen.

Um einen groben Überblick über die Thematik zu bieten betrachten wir zunächst bestendende Ethikliteratur im Allgemeinen. Aufbauend darauf behandeln wir spezifischere ethische Ideen, insbesondere die state-of-the-art ethischen Codes in der Informatik. Mit diesem vorhandenen Wissen zeigen wir dann, welche blinden Flecken und offenen Herausforderungen in diesen ethischen Codes bestehen. Weiters zeigen wir auf, wie die existierenden ethischen Paradigmen mit Hilfe von zentralen Ideen der Global Sustainable Information Society auf die nächste Ebene gehoben werden können. Um einen strukturierten Weg vorzugeben, um Ethik in der Software Entwicklung zu bewerten, präsentieren wir eine Liste an Kriterien, an die sich Experten im Feld der Informationstechnologie halten sollten. Diese Kriterien dienen als zusätzliche ethische Ressourcen, welche die existierende ethische Basis der Profession bereichern. Zuletzt wenden wir diese Kriterien an einigen konkreten Beispielen an, um Innovationen im breiten Feld der Informationstechnologie der letzten Jahre zu bewerten. Damit zeigen wir sowohl, wie man mit Hilfe dieser Kriterien Fehlentwicklungen identifizieren, als auch bewerten kann.

# Abstract

The internet and computers in general play a major role on a global scale. Worldwide, sectors like industry, education, government and society are clearly influenced by innovations and advances in information technology. Concepts like ubiquitous computing, distributed systems, mobile computing, social networks, internet of things, big data, and artificial intelligence shape the way we live our lives and play a central role in what the future of humanity looks like.

Software Engineers should take responsibility for the software systems they plan and develop. While there are opportunities to create software that positively influences humanity, there is also a lot of potential for causing harm to people.

In recent years, huge global challenges, such as climate change, geopolitical conflicts (such as wars), economic instability, social inequalities and digital transformation require increasingly sophisticated methods to determine ways of developing software that are ethically aligned with solving these global challenges.

While software engineers are playing a major role in driving innovation across multiple sectors, it is important to critically question new innovations and consider a multitude of ethical ideas that aid in evaluating whether these innovations serve the progress of humanity as a whole or are of detriment to our world.

In order to provide a broader overview of this topic, in this thesis we first consider existing ethics literature in general. We next consider more specific ethical ideas, including state-of-the-art ethical codes in computer science. Furthermore, we then line out blind spots and open challenges in these ethical systems. Afterwards, we provide a way to overcome existing ethical paradigms by introducing key ideas of the Global Sustainable Information Society. In an effort to provide a structured way to evaluate ethics in software development, we then provide a list of criteria to adhere to by anyone in the information technology profession. These criteria serve as additional ethical resources, enhancing the existing ethical basis for the profession. Finally, we apply these criteria to judge a few concrete innovations in the broad field of information technlogy in recent years. This is being done in order to point out how to identify and evaluate misguided developments.

# Contents

CHAPTER 1

# Introduction

Ethical studies aim to answer the general question of what the right way to live is. This was traditionally understood to be a question of societal nature, focussing on human relationships. However, technology, and in particular information technology, change key aspects of our daily life. Software changes the ways humans communicate, provide abilities to collect and store data to a vast degree, and offer constant connectedness. Therefore, aligning technology with ethical principles must be treated as a serious priority. Hofkirchner claims that the global challenges we face today are reaching a critical state: "Earth civilisation as a whole is in a crisis." [33, p. 120]. In order to tackle these challenges, he proposes the vision of the *Global Sustainable Information Society* (*GSIS*) [33, p. 136ff]. This ontological model proposes a techno-eco-social transformation to resolve the global crisis in a positive outcome. In this way, humanity can avoid a breakdown of the world system and proceed to a new suprasystem which utilizes Globality, Sustainability and Informationality as key properties.

This thesis proposes a novelty in the field of software engineering ethics. The concepts of GSIS are used as a basis to put new ethical guidelines into place. These guidelines can be used to assess technological artifacts on an ethical level. The approach of this thesis does not in any way invalidate existing state-of-the-art ethical principles or codes, such as ACM Code of Ethics and Professional Conduct [1] or The Software Engineering Code of Ethics and Professional Practice [25]. Ethical guidelines proposed in this thesis enrich the existing ethical principles and serve as a new way of ethically evaluating technology and innovation. Having the vision of GSIS in mind, and therefore promoting fruitful human-human, human-nature, and human-technology relationships, the guidelines offer additional ethical criteria for stakeholders in the field of information technology. The guidelines not only address ethical means, but also ethical ends: Because of the transformative approach of GSIS, they provide ways to identify desirable goals. Furthermore, the guidelines promote interdisciplinary thought, as the focus is not strictly on ethics regarding professional competence, but also comprise the eco-social space.

This thesis is structured as follows: In Chapter 2 we progressively introduce ethical perspectives. We start by introducing broad classical ethical perspectives in order to build a foundational understanding about ethics. From there, we gradually consider more specialized ethical concepts. These include Digital Humanism and finally concrete codes that represent state-of-the-art ethical principles in computer science and software engineering. In Chapter 3 we then aim to identify blind spots, i.e. issues and open challenges, in existing ethical codes. As a means of addressing open challenges, in Chapter 4 we then introduce concepts required to overcome existing ethical paradigms, most notably the concept of the Global Sustainable Information Society (GSIS). Chapter 5 presents the reader with novel ethical guidlines derived from GSIS ideas. Finally, Chapter 6 examines concrete, practical examples of questionable developments in information technology with the help of the new guidelines from this thesis. This shows how to apply findings from this thesis to practical cases, outlining their capability of identifying misguided innovations and developments.

CHAPTER 2

# Ethics Literature

In order to provide a concept that addresses the ethical challenges in the field of software engineering and internet computing, we first establish ethical foundations. The findings of this chapter are essential in order to understand that ethics are evolving: As humanity faces new challenges, our ethical understandings and considerations require adaptation. This chapter provides a broad overview of the ethical foundations that are relevant for the field of Information Technology.

This chapter is structured as follows: In a first step we introduce main classical ethical perspectives found in literature [cf. Section 2.1]. Having these different perspectives in mind, we discuss a central philosophical standpoint called *Humanism* [cf. Section 2.2]. Building on Humanism, we then introduce *Digital Humanism* as a means to promote a humanistic digitalization. [cf. Section 2.3]. Concluding the ethics literature research, we explore ethical codes in the field of Computer Science and Software Engineering [cf. Section 2.4].

## 2.1 Main Ethical Perspectives

Chowning et al. [12, p. 18-23] list five main ethical perspectives, also referred to as ethical frameworks. Each of these perspectives in a way represents a possible guideline to moral life. These perspectives all have their strenghts and limitations that deserve consideration. Furthermore, these perspectives serve as a reference point to compare one's own position to. They might strengthen the structure or organization of one's way of determining what is morally wrong or right. However, it is crucial to make an effort to integrate these perspectives. In the end, if we cannot find a clear ethical path forwards, we are left uncertain about how to proceed. This is why a binding ethic that provides a moral framework is appealing. In order to work towards this, unification is used with the idea of *unity through diversity*: Each ethical perspective can have reasonable aspects, while still being conceptually different. Differences and possible objections can even enrich the

ethical principles: Reasonable and helpful parts of the concepts, even if they are just appropriate in one part or area, can be assigned a dedicated place in a new, binding ethic. This way, the "dispute" between the concepts is settled, allowing each concept to contribute to a unifying ethical concept. The concept of unity through diversity will be revisited and incorporated later in this work [cf. Sections 4.5, 5.1, 5.3.2].

The five ethical frameworks are briefly described and discussed below.

### 2.1.1 Duty-Based Ethics

In *duty-based* or *deontological* ethics, famously developed and proposed by philosopher Immanuel Kant, the ethical focus is on one's act itself, not on the consequences of an action. Acting in accordance with moral duties and respecting people are emphasized in duty-based ethics.

Kant's *categorical imperative* [37] is a basic principle on deciding whether an action is morally good or not and represents a cornerstone of duty-based ethics. The categorical imperative is stated in different forms in Kant's work, but the most cited *Gesetzesformel* is stated as follows.

"One must act only in such a way that one could will one's act to become a universal law or rule (maxim)."

Kant not only makes the distinction between duties that we have to ourselves and duties we have to others - he also introduces the concept of *perfect* and *imperfect duties*.

In the context of this thesis, perfect duties play a major role in understanding duty-based ethics. The idea is not to consider the consequences of an action, but instead consider to universalize the proposed maxims in question. If the universalization of a maxim causes a logical contradiction, we have a perfect duty not to act by this maxim in question.

The proposition: "Stealing is permissible" causes such a logical contradiction when one attempts to universalize it. Stealing requires the existence of personal property. However, if the proposition was to be universalized, personal property could not exist. Therefore, according to the categorical imperative, we have a perfect duty not to steal.

Perfect duties are considered to be morally reprehensible if violated - they are considered required, basic duties for every human. Furthermore, perfect duties do not allow for exceptions. This is being criticized, especially by consequentialists, who propose that an action is considered moral if it produces some nonmoral good. As one can think of circumstances where telling a lie produces generally more good than behaving honestly, consequentialists conclude that it can be moral to lie. In contrast, strict adherence to truth-telling as a perfect duty might require one to reveal harming information to others.

Imperfect duties, in contrast to perfect duties, depend on personal preference. In contrast to perfect duties, one is not supposed to receive blame for not completing an imperfect duty, instead they shall receive praise for completing it. Going beyond the perfect duties to take on additional duties is what makes one deserving of praise. One basic difference

to perfect duties is that imperfect duties are subject of circumstance and dynamic change. One could argue that they can never be entirely completed. Kant provides the example of cultivating one's own talents as an imperfect duty.

**Positives about duty-based ethics**

- offers consistent principles and rules that can be logically examized by checking for contradictions using the process of universalization
- values human dignity and does not consider a human as a means to an end.
- recognizes autonomy and individual rights
- encourages accountability by requiring to uphold to moral obligations

**Objections to duty-based ethics**

- unclear how to resolve potentially conflicting duties
- lack of flexibility, as perfect duties do not allow for exceptions and situational considerations
- prioritization of rules over consequences, leading to potentially questionable outcomes
- requires actively considering maxims - does not guide one into forming a set of moral rules
- no consideration for emotions

### 2.1.2 Virtue-based Ethics

In *virtue-based* ethics, the focus is on the *character traits* of the individual. Virtues are *positive*, or *admirable*, traits that an individual possesses which reflect a moral character. Virtues are to be distinguished from single actions: In virtue-based ethics the moral value of an action is based on a virtue. Therefore, in order to act morally, virtues need to be cultivated by the individuals.

Possessing a virtue means to be a certain type of human with a certain mindset. In virtue-based ethics, it is insufficient to consider a person honest just because they act honestly. The individual could be motivated by other goals, and only act honestly because of the fear of getting caught. Contrary, a reason to consider a person honest would be if the person recognizes and values the virtue of honesty.

What virtues should an individual according to virtue-based ethics cultivate? MacIntyre states that philosophers, such as Aristotle, Aquinas, or Franklin, all came up with their

own list of essential moral virtues [45]. However, there is no consensus on essential virtues, let alone an extensive list that is universally agreed on.

Furthermore, it remains an open question whether virtues can be *taught*.

**Positives about virtue-based ethics**

- emphasizes character traits of an individual, compared to rules or outcomes
- promotes the cultivation of character traits and working on human excellence
- cultivation acknowledges ethics as an ongoing process
- allows for more circumstancial flexibility than strict rules

**Objections to virtue-based ethics**

- lack of clear guidelines and decision-making framework
- subjectivity; no clear agreement about core virtues of a human
- unclear whether virtues can be taught or are inborn to humans
- people can be of good character traits and arguably still commit morally wrong acts - the reverse applies as well

### 2.1.3 Consequentialist Ethics

In *consequentialist* or *utilitarian* ethics, co-founded and promoted by English philosophers John Stuart Mill and Jeremy Bentham [36, 9], not the actions themselves are of focus, but the consequences of them. A moral action is one that maximizes the amount of positive outcomes that result from performing that action.

In this maximization strategy, the consequentialist aims to perform the actions that cause the most good to the largest number of people. There are different measures or ideas about what is considered a good action. Traditonally though, the idea is that pleasure is generally preferrable to pain. Therefore, actions that cause some kind of pleasure are considered moral and are to be maximized. Conversely, actions that cause pain are considered immoral and are to be minimized.

Considering this traditional consequentialist approach, typical values that are considered good are life (opposed to death), pleasure, health. Beyond this traditional form, literature also identifies other good values like knowledge or friendship.

**Positives about consequentialist ethics**

- stimulates thought about outcomes; focus on best possible consequences of actions
- flexibility based on context; avoids dogmatism
- equal treatment of all people
- encourages practical solutions, even for many dilemmas

**Objections to consequentialist ethics**

- interest of the greater mass of people can overrule interests of minorities
- neglects the nature of virtues themselves as the consequences of actions are of focus
- subjectivity in defining good consequences
- potentially difficult to predict consequences, since consequences are not assured beforehand

### 2.1.4 Principle-Based Ethics

*Principle-based* or *deontological* ethics, in contrast to consequentialist or virtue-based ethics, revolve around following fundamental moral principles. They provide clear guidelines that rely on recognized principles. This approach ensures consistency in reasoning while respecting situational factors. However, principle-based ethics can appear rigid in complex situations: Principles can cause conflicts in complex situations, but there is no clear guideline on resolving these conflicts. Despite their limitations, principle-based ethics are respected in many professions.

The philosophers Beauchamp and Childress provide moral guidelines for the field of health profession [7]. They advocate four main moral pillars of health care, which are also commonly adapted as general goals of principle-based ethics. The four pillars are stated as follows.

| Principle | Explanation |
|---|---|
| Respect for Autonomy | Acknowledgement of people's right to make chocies, to hold views, and to take actions based on personal values and beliefs |
| Justice | Treat others equitably, distribute benefits / burdens fairly |
| Nonmaleficence - do not harm | Obligation to not inflict harm intentionally |
| Beneficence - do good | Provide benefits to people and contribute to their welfare (an action done for the benefit of others) |

Table 2.1: Four Pillars of Principle-Based Ethics

**Positives about principle-based ethics**

- clear moral framework with specific, easy to understand guidelines
- in line with duty-based and consequentialist ethics (duties are respect for autonomy, justice, while nonmaleficience and beneficence are outcome-based)
- takes into account specific situations, having to balance the pillars in question

**Objections to principle-based ethics**

- rigidity, as the fixed principles sometimes fall short in judging complex, real-world situations
- lacks a main theory that unifies the pillars; no common grounds that provide a unified morality
- specific situations might lead to principles conflicting with each others, without clear guidelines on how to resolve the conflicts
- prioritization of rules over emotions
- fixed principles offer potential for dogmatism

### 2.1.5 Care-based Ethics

In *care-based* ethics (or *ethics of care*), the focus lies on relationships and the underlying power structures in a given situation. While duty-based or consequentialist ethics aim to generalize and establish standards, the care-based approach considers individuals. The idea is that not only individual situations differ, but individuals themselves have different dependencies on each others. Carol Gilligan, considered the originator of the ethics of care, criticized the application of generalized standards as "morally problematic, since it breeds moral blindness or indifference" [21].

Therefore, care-based ethics stand in opposition to ethics that try to generalize and universalize moral standards. While generalized ethics try to answer the question of "What is just?", ethics of care address the question of "How to respond?". However, since care-based ethics are not attempting to generalize standards, it is tougher to obtain clear guidelines. According to care-based ethics, the specific, individual case needs to be considered in order to act morally.

**Positives about care-based ethics**

- recognizes context and considers individual situations, people, and their relationships
- counteracts principle-based ethics and generalized standards

- attempts to not become morally blind or indifferent by always considering the individual case
- advocates respect for minorities, emphasizing importance of addressing inequalities

**Objections to care-based ethics**

- relationships or individual structures might not always be perceivable accurately
- does not provide clear principles that are straight-forward to apply
- potential inconsistency because of subjectivity and emotional influence
- effective for personal relationships, but difficult to apply on larger-scale or broad ethical issues

## 2.2 Humanism

Humanism is a progressive philosophical standpoint centered around human well-being. Instead of ethics grounded in religious or supernatrual ideas, It considers the human as the main focus in ethical considerations. Humanism is not a dogma: Instead, humanism based ethics change and develop as humankind's understanding of what is best for human autonomy, freedom, and general well-being, progresses.
There is not one clear definition of Humanism, but a commonly referred manifest is Humanism and Its Aspirations: Humanist Manifesto III, a Successor to the Humanist Manifesto of 1933 [5]. In the following, we will refer to this manifest as *Humanist Manifesto III*. This manifest published by the American Humanism Association has changed and adapted since its first Manifesto from 1933.

We will now briefly look at the history of the Humanist Manifesto.
The first Manifesto from 1933 [3] critiques religion and takes a strong position against theism which is demonstrated especially in its sixth affirmation:

- We are convinced that the time has passed for theism, deism, modernism, and the several varieties of "new thought."

Breeden discusses objections to the manifesto in detail. He lists one addressing this sixth affirmation as follows: "Many Humanists at the time thought that this statement precluded even the discussion of theism." [10]. This was clearly a problem, since encouraging discussion and debate between humans has always been an important part of humanism. Other objections were documented by the Humanist Manifesto itself in the second manifesto from 1973, such as presenting an over-optimistic outlook. [4].

The affirmations from the original manifesto have since been reworked completely, and Humanist Manifesto III focuses more about how we as humans acquire knowledge, how

we can derive ethical values, where life's fulfillment comes from, and what society means to humankind. Humanist Manifesto III therefore has evolved significantly: The original Humanist Manifesto I included the universe and religions as an important topic. Humanist Manifesto III, however, does not even include the words "universe" and "religion".

**Affirmations of Humanist Manifesto III** In the following we will briefly cover the six affirmations that Humanist Manifesto III makes.

- Knowledge of the world is derived by observation, experimentation, and rational analysis.
- Humans are an integral part of nature, the result of unguided evolutionary change.
- Ethical values are derived from human need and interest as tested by experience.
- Life's fulfillment emerges from individual participation in the service of humane ideals.
- Humans are social by nature and find meaning in relationships.
- Working to benefit society maximizes individual happiness.

**Objections to Humanism** A concise list of objections to humanism was presented in a talk by Andrew Copson in Conway Hall, United Kingdom on the May 22, 2015 [13]. There, Copson listed the six most common objections to Humanism and provided rebuttals to them.

In the following, we will cover these six objections against humanism and include a brief rebuttal to them. This part gives a loose summary of Copson's talk, but integrates some own thoughts.

- Humanism diminishes the dignity of humanity. This objection claims that if humans are not a special creation of a higher being, then humans are not special, and therefore it is not important what we (ethically) do as humans.
  - Response: The evidence, especially evolution, just does not show that we are the apex of creation or that the universe is created for us. Humans are not special in that sense.
- Humanists can not have morals. If there is no objective morality (guided from a higher being) one can never know what is right or wrong.
  - Response: There is no need to hypothesize an extrahuman original point of morality. Our foundation of morality might not be perfect, but we accept and learn from our mistakes and strive for improvement.

- Humanism is too dry, too rational. Humanists can not explain love with science.
  - Response: Science is the demonstrable way to understand the world. It can also be an inspiration and spark interest. Just because there are feelings or things we can not yet explain, we can not conclude a supernatural origin behind them.
- Humanism is a secular religion.
  - Response: Humanism is a *descriptive* word applied *retrospectively* to a certain set of *beliefs and values*. Someone that has never heard of christianity could not be a christian, but someone that has never heard of humanism could be a humanist. They might just not use the label, but live by the beliefs and values of humanism.
- Humanism is a myth of progress. Humanists are too optimistic, ignore historical tragedies. It is unwarranted utopianism.
  - Response: Humanists take a realistic view. The idea of making an improvement itself can ultimately lead to an improvement. Quality of life has genereally improved a lot as humankind progresses.
- Pointlessness of it all. Humanism is nihilism. If there is no god, no future life and the universe is a totally natural phenomenon, then the universe has no purpose.
  - Response: Identify your fears. In this case, death. Life is meaningful *now*. Using your computer today is not useless when it stops working in a few years. In life, the meaning comes with living. Embrace the here and now, the goals we have, our experiences we make. Life is its own reward. Do not focus solely on a better future neglecting your life today.

## 2.3 Digital Humanism

Digitalization and digital innovation offers numerous possibilites and opportunities for human life, but it also impacts and changes our society notably. It is a challenge to keep the human well-being in the center of interest in the time of digital change. Instead of technology defining, commanding, or guiding our lives, in Digital Humanism, technology is seen as a means to assist and enrich human life. This means prioritizing the human interest as a fundamental social good and designing technology to best serve human society.

In his book, Nida-Rümelin discusses Digital Humanism and gives a brief idea about Digital Humanism in German; we provide our own loose translation of this passage here: “Digital Humanism is not defensive. It does not intend to slow down, but instead encourages technical or technological progress in the age of artificial intelligence. Digital

Humanism pleads for acceleration of human progress using digital possibilites in order to colour our lives richer, more efficent, and more sustainable" [48, p. 207, translation ours].

Primarily, Digital Humanism represents a type of digitalisation that values the human. Even though it mainly aims for a humane digitalisation, it is important to point out that also a digital humanisation is a goal. More specifically, it is not the goal to digitalize the humane, but do so in a humane way to begin with. In other words: A humane digitalization of the humane.

Digital Humanism builds on Humanism in order to provide a digitalistic humanisation.

**Core Principles of Digital Humanism** In The Vienna Manifesto on Digital Humanism [61], Werther proclaims the following core principles.

- Digital technologies should be designed to promote democracy and inclusion.
- Privacy and freedom of speech are essential values for democracy and should be at the center of our activities.
- Effective regulations, rules and laws, based on public discourse, must be established.
- Regulators need to intervene with tech monopolies.
- Decisions with consequences that have the potential to affect individual or collective human rights must continue to be made by humans.
- Scientific approaches crossing different disciplines are a prerequisite for tackling the challenges ahead.
- Universities are the place where new knowledge is produced and critical thought is cultivated. Hence, they have a special responsibility and have to be aware of that.
- Academic and industrial researchers must engage openly with wider society and reflect upon their approaches.
- Practitioners everywhere ought to acknowledge their shared responsibility for the impact of information technologies.
- A vision is needed for new educational curricula, combining knowledge from the humanities, the social sciences, and engineering studies.
- Education on computer science / informatics and its societal impact must start as early as possible.

**Objections to Digital Humanism** There are objections to Digital Humanism. Unlike for Humanism, where we briefly discussed objections and rebuttals in this chapter, we provide more detail for Digital Humanism. This discussion leads to first findings in regards to ethical blind spots in its related chapter [cf. Chapter 3]. Therefore, we dedicate a separate section to these issues [cf. Section 3.1].

## 2.4 Ethics in Computer Science

There have been efforts to establish an ethical standard to the profession of Computing. In the following we cover the two most well-established ethical codes, namely ACM Code of Ethics and Professional Conduct [cf. Section 2.4.1] and The Software Engineering Code of Ethics and Professional Practice [cf. Section 2.4.2]. These codes reflect the state-of-the-art in computing or software engineering ethics. In the following examination of these codes, we strive to make sure not to take the principles and ideas out of context. We took extra care in order not to view the principles in isolation to justify possible errors. For the full details, always refer to the originals [1, 25].

### 2.4.1 ACM Code of Ethics and Professional Conduct

The "ACM Code of Ethics and Professional Conduct" [1], as adopted by the Association for Computing Machinery Council on June 22nd, 2018, is a code that is directed towards all people professing in computing, including those who use computers or software in an impactful way. Within this subsection, we will refer to the ACM Code of Ethics and Professional Conduct as *Code*.

The focus of the Code is to guide computing professionals in ethical and responsible behaviour. It emphasizes the importance of the public good as the main consideration. Principles of the Code are structured in the form of "A computing professional should ...". These principles also contain explanations and guidelines on understanding and applying them. There are four broad sections in the original document: General ethical principles, professional responsibilities, professional leadership responsibilities, and compliance with the code.

The principles of the ACM Code of Ethics and Professional Conduct are listed below. The explanations of the principles in the Code are formulated in own words, except when a direct quote is used, denoted by quotation marks (""). Quotes and the original Code are taken from the single-page website of the Code [1].

#### 1. General Ethical Principles

The following principles comprise fundamentals that encompass a basis for the rest of the Code.

**Principle 1.1** *A computing professional should contribute to society and to human well-being, acknowledging that all people are stakeholders in computing.*

This principle outlines the obligation to focus on the benefit of human society. In the detailed explanation to this principle, it is stated that this includes the benefit of the surrounding environment (especially environmental sustainability). Another goal is to minimize negative consequnces like health, safety, security, and privacy.

We emphasize that this principle holds a special position within the Code: By emphasizing human well-being and contribution to society, it addresses ethical ends. With the descriptions of the other principles, it will become clearer that the rest of the Code predominantly addresses ethical means. These means generally revolve around ideas of Principle-Based ethics [cf. Table 2.1], like minimization and avoidance of harm, acknowledgement of various rights, upholding respect, and establishing fairness. Note that this is not an exhaustive list of ideas that the concrete Code comprises.

**Principle 1.2** *A computing professional should avoid harm.*

First, this principles explains the concept of *harm*. Harm, in the context of the ACM's code of ethics, is considered as negative consequences, in particular ones that are significant and unjust. At first glance, this principle seems obvious, but there is more to it in the detailed explanations that leaves room for discussion.

It is mentioned that even well-intended actions may lead to (unintended) harm, viewing the topic from a somewhat consequentialist perspective. Note that consequentialist ideas suggest that an action is considered moral only if the consequences of that action are good. The Code reflects this by stating that avoiding harm begins with "begins with careful consideration of potential impacts on all those affected by decisions".

The principle then discusses the case "When harm is an intentional part of the system, ...", mentioning that "those responsible are obligated to ensure that the harm is ethically justified. In either case, ensure that all harm is minimized." This, however, raises questions about systems that include *intentional harm* in the first place. We will revisit this subject later in this work [cf. Sections 3.3, 2].

Furthermore, the principle discusses how to minimize the possibility of indirectly or unintentionally harming others, referring to best practices except when ethical considerations suggest otherwise. It also discusses careful consideration of consequences of data aggregation.

Finally, the principle mentions responsibility "to report any signs of system risks that might result in harm". It even mentions importance of carefully analyzing the situation, as unjustified reporting might again result in harm.

**Principle 1.3** *A computing professional should be honest and trustworthy.*

This principle acknowledges that transparency and full disclosure about the system as well as human capabilities, limitations and potential problems is part of the code. Honesty

and trustworthyness are acknowledged as a form of virtues, while "deliberately false or misleading claims, fabricating or falsifying data, offering or accepting bribes, and other dishonest conduct" are seen as violations of the Code.

**Principle 1.4** *A computing professional should be fair and take action not to discriminate.*

Promoting fairness and avoiding discrimination, as well as providing means to redress grievances are emphasized in this principle. Ideas of this principle comprise two areas:

Firstly, it is emphasized that computer professionals should strive for fair treatment and inclusion of all people, and that discrimination on any inappropriate basis violates the Code.

Secondly, it is addressed that information technology is capable of introducing new or altering existing forms of inequities. The principle pushes for technologies and systems that are inclusive and accessible. Systems that oppress people should be taken action against.

**Principle 1.5** *A computing professional should respect the work required to produce new ideas, inventions, creative works, and computing artifacts*

Respect for the work, the ideas themselves, and the artifacts that arise from that are emphasized in this principle. Furthermore, custom and law exceptions that limit a creator's control are recognized as a public good. The principle mentions that creators should not oppose reasonable uses of their work in favor for the public good. Furthermore, open source and free software, where people help each others and contribute to the works of others, are given as a positive aspect of this principle.

**Principle 1.6** *A computing professional should respect privacy.*

In Principle 1.1, minimizing negative consequences for people's privacy was mentioned as one of many points to consider. Here, privacy is given a separate principle. It is mentioned that in computing, privacy is of particular concern. Collecting and processing data is an embedded part of information systems. Therefore, it makes sense to give this topic its own principle. The details of this principle emphasize that system design and policies should always enable individuals to have control over their personal data, enabling them to obtain, alter, and delete their own data. Furthermore, it is mentioned that only the minimum amount of data required should be collected, and that the system has to disclose details about the process transparently.

**Principle 1.7** *A computing professional should honor confidentiality.*

This principle acknowledges that there is a mass of confidential information that computer professionals have to work with. As examples, the principle lists "trade secrets, client data, nonpublic business strategies, financial information, research data, pre-publication scholarly articles, and patent applications". The description of the principle mentions that confidential data should be protected, unless there is violation of law, of organizational regulations, or of the Code. For these exceptions, the computer professional should consider thoroughly what authorities to appropriately disclose this information with and if such a disclosure is consistent with the Code.

### 2. Professional Responsibilities

The following principles address additional ethical means, concerns, and specific professional responsibilities.

**Principle 2.1** *A computing professional should strive to achieve high quality in both the processes and products of professional work.*

In this principle, it is mentioned that a lack of quality of work can negatively impact stakeholders of the project. Computing professionals should therefore promote high quality work from themselves and colleagues and insist on this high quality.

**Principle 2.2** *A computing professional should maintain high standards of professional competence, conduct, and ethical practice.*

Outlining the importance of high quality standards, this principle mentions the importance of professional competence. This professional competence encompasses technical knowledge, awareness of the social context in which their work may be deployed, skill in communication, skill in reflective analysis and skill in recognizing and navigating ethical challenges. It is also mentions that upgrading these skills should be an ongoing process and that "Professional organizations and employers should encourage and facilitate these activities".

**Principle 2.3** *A computing professional should know and respect existing rules pertaining to professional work.*

In the context of this principle, a rule is referred to as "local, regional, national, and international laws and regulations, as well as any policies and procedures of the organizations to which the professional belongs". It is encouraged that a computer professional abides and respects these rules, but challenges them if they judge a rule as unethical. The standard by which to judge the rules is phrased as "A rule may be unethical when it has an inadequate moral basis or causes recognizable harm".

Apart from that, the explanations to the principle advise to not just violate rules that are judged unethical, but take appropriate measures through existing channels first. It is also stated that the computing professional must consider the potential consequences of challenging or violating an unethical rule and take responsibility for that action.

**Principle 2.4** *A computing professional should accept and provide appropriate professional review.*

This principle emphasizes the importance of professional review at all stages that required for high quality professional work. This should be established via peer and stakeholder review. Furthermore, computing professionals should themselves thoroughly review work of others.

**Principle 2.5** *A computing professional should give comprehensive and thorough evaluations of computer systems and their impacts, including analysis of possible risks.*

This principle is designed to prevent negative consequences or unforseen impacts and risks of computing systems: Systems should be frequently evaluated as thoroughly and objectively as possible. A special focus should be put on identifying and mitigating potential risks regarding machine learning systems. When future risks cannoot be reliably predicted, frequent reassessment of the potential risks is required, and not deploying the system should be an option to consider. Furthermore, when major risks are identified, these should be reported to appropriate stakeholders.

**Principle 2.6** *A computing professional should perform work only in areas of competence.*

In this principle, the importance of honesty and openness in regard to one's own competence is outlined: A computing professional has to make a judgement about their own work competence in regards to a possible work assignment. Should a lack of expertise for a work assignment be identified, this has to be disclosed to appropriate parties. Furthermore, it is mentioned that a computing professional's ethical judgement should be the deciding factor on whether to work on the assignment or not.

**Principle 2.7** *A computing professional should foster public awareness and understanding of computing, related technologies, and their consequences.*

This principle outlines that impacts, limitations, vulnerabilities, and opportunities of computer systems should be disclosed with the public. This communication is to be done in in a "clear, respectful, and welcoming" way. Furthermore, inaccurate or misleading information regarding computing should be respectfully addressed by a computing professional.

**Principle 2.8** *A computing professional should access computing and communication resources only when authorized or when compelled by the public good.*

According to this principle, access to systems and data should be restricted and respected by a computing professional. Therefore, a system should not be accessed unless authorized or there is a compelling belief in place that accessing the system is in the interest of the public good. It is mentioned that under exceptional circumstances, unauthorized access is warranted to disrupt or inhibit the functioning of malicious systems. Extraordinary precautions are to be taken in order to prevent harm to others. Also, a publicly accessible system alone is not sufficient grounds on its own to imply authorization.

**Principle 2.9** *A computing professional should design and implement systems that are robustly and usably secure.*

This principle discusses robustness and security of systems. System security should be a primary consideration in design and implementation of a system, as security breaches cause harm. Security is an ongoing process as computer systems get deployed or change over time, and should be monitored, patched, and reported appropriately. In cases where harm or misuse of a system is predictable or even unavoidable, it is mentioned that not implementing the system at all may be the best option.

### 3. Professional Leadership Principles

The following principles are aimed at computing professionals, especially those acting in a leadership role.

**Principle 3.1** *A computing professional, especially one acting as a leader, should ensure that the public good is the central concern during all professional computing work.*

In this principle, the public good is emphasized as the central focus of computing in this principle. No matter the process, system type, or stage in development: The public good should always be of concern.

**Principle 3.2** *A computing professional, especially one acting as a leader, should articulate, encourage acceptance of, and evaluate fulfillment of social responsibilities by members of the organization or group.*

As this principle outlines, technical groups and organizations carry an effect to broader society, and leaders of such should accept associated responsibilites. Leaders should encourage full participation of computing professionals in meeting relevant social responsibilites and discourage tendencies to do otherwise. The welfare of society is emphasized in the form of reducing “harm to the public” and raising “awareness of the influence of technology in our lives”.

**Principle 3.3** *A computing professional, especially one acting as a leader, should manage personnel and resources to enhance the quality of working life.*

According to this principle, leaders should strive to enhance the quality of working life. This includes personal and professional development, accessibility requirements, physical safety, psychological well-being, and human dignity.

**Principle 3.4** *A computing professional, especially one acting as a leader, should articulate, apply, and support policies and processes that reflect the principles of the Code.*

This principle discusses support for policies that reflect ethics of this Code: Clearly defined organizational policies that are consistent with the Code should be defined by leaders. Furthermore, compliance with the policies should be encouraged and rewarded. In cases of violation of policies appropriate action should be taken.

**Principle 3.5** *A computing professional, especially one acting as a leader, should create opportunities for members of the organization or group to grow as professionals.*

This principle mentions growth opportunities for professionals: All of organization or group members should experience opportunities for education and growth as professionals. This growth should include knowledge and skills in professionalism, in the practice of ethics, and in their technical specialties.

**Principle 3.6** *A computing professional, especially one acting as a leader, should use care when modifying or retiring systems organization or group.*

Changing, modifying, and retiring computing systems are emphasized as impactful on productivity of users and quality of work. Therefore, change, modification, and discontinuation of systems should be thoroughly investigated. Users should also be notified about the risks associated with using an outdated or unsupported system, even before the support of such a system ends. Computing professionals should also assist users of a system in understanding when and why a replacement of an inappropriate or outdated system may be needed.

**Principle 3.7** *A computing professional, especially one acting as a leader, should recognize and take special care of systems that become integrated into the infrastructure of society.*

In this principle, impacts of integration systems into society are discussed: Computing systems have impact on all aspects of society. Leaders should be good stewards of systems that integrate into social infrastructure. This encompasses establishing policies

for fair system access and continual monitoring of the level of integration of their systems into society. As level of integration changes, impact and ethical responsibilites of the associated organizations are likely to change. Therefore, standards of care should be developed and respected.

**4. Compliance with the Code**

The following principles address compliance with the Code. Compliance with the code is required of every ACM member, ACM *Special Interest Group* (SIG) member [18], ACM award recipient, and ACM SIG award recipient.

**Principle 4.1** *A computing professional should uphold, promote, and respect the principles of the Code.*

Here, technical and ethical excellance are mentioned to set the standard for the future of computing. Therefore, the Code should be adhered and contributed to in order to improve it. As breaches of the Code are identified, appropriate actions are to be taken.

**Principle 4.2** *A computing professional should treat violations of the Code as inconsistent with membership in the ACM.*

This principle is aimed at ACM membership. ACM members should encourage and support adherance to the Code by all computing professionals (regardless of ACM membership). ACM members should report violations of the Code to the ACM that "may result in remedial action as specified in the ACM's Code of Ethics and Professional Conduct Enforcement Policy".

### 2.4.2 The Software Engineering Code of Ethics and Professional Practice

In a joint ACM/IEEE-CS initiative, The Software Engineering Code of Ethics and Professional Practice was adopted and approved in 1997 [24]. This Code of Ethics comprises eight principles that apply to all positions in the field of professional software engineers. There is a short version of the code that considers ethical aspirations at a high level of abstraction [25, p. 232]. It comprises a preamble and eight principles.

The full version provides details in the form of clauses that explain how the aspirations influence our actions in all phases of the software engineering profession [25, p. 232-238]. It comprises a preamble and a full version of these eight principles, each containing several clauses.

Within this subsection, we will refer to the The Software Engineering Code of Ethics and Professional Practice as *Code*.

We will present the Code as follows: First, we will summarize the preambles of both the short, and the full version. Then, for each of the eight principles given, we will state the principle, followed by an assessment of the detailed version in the form of a summary. The assessments of the preamble and principles in the Code are formulated in own words, except when a direct quote is used, denoted by "". In these cases, we refer to the original Code [25]. Again, we will make sure not to take the principles out of context, as well as not to view the principles in isolation to justify possible errors.

We now give some thoughts about the preambles [25, p. 232-233].

### Short version of the Preamble

The short version of the preamble states that it represents aspirations on a high level of abstraction. It mentions how the details from the full version change the aspirations mentioned here, and outlines the importance of both the short, and the full version: Without the full version, the short version might appear "high-sounding but empty", and without the abstractions of the short version, the details of the full version "the details can become legalistic and tedious". It is mentioned that only together, these elements create a comprehensive and meaningful code.

Also, the short version mentions a list of commitments for Software engineers, such as striving to make all phases of software development a respected profession. Finally, a call for committing to "health, safety, and welfare of the public" is made before going into the short version of the eight principles.

### Full version of the Preamble

The full version of the preamble first addresses the central role and increasing significance of computer science in different fields, and society itself. Since software engineers develop software systems for others, they influence outcomes for good or for harm. This influence is recognized and a it is stated that ethical practices therefore are of highest importance. It is stated that software engineers should adhere to this Code.

Then, the preable outlines the eight principles the Code comprises and states that it applies to all stakeholders in the profession. Then it mentions obligations within relationships with individuals, groups, and organizations. It is mentioned that "These obligations are founded in the software engineer's humanity, in special care owed to people affected by the work of software engineers, and in the unique elements of the practice of software engineering".

Furthermore, instructions for handling and understanding the code are given. It is mentioned that individual parts are not to be considered in isolation, as this may cause "errors of omission or commission". The list of principles is not to be considered exhaustive and the clauses should not be understood as ethical decision makers in all practical situations in the profession. Furthermore, the code is not to be seen as an algorithm

generating ethical decisions. Also, ethical tensions are to be expected and addressed in a thoughtful way using fundamental principles.

It is then stated that the code is suitable for individuals and whole teams alike.

Finally, this full version of the preamble mentions its function as an educational tool. By providing a consensus within the profession on ethical matters, this Code serves as a tool to educate the public, and future professionals in the field about ethical responsibilities of software engineers.

### Eight Principles of the Code

The eight principles of The Software Engineering Code of Ethics and Professional Practice are listed below.

**Principle I. Public** *Software engineers shall act consistently with the public interest.*

This principle states that acting in favor of public interest is a central role for software developers. In earlier versions of the Code, the principle *Product* was listed before *Public*. This got adjusted in order to reflect the importance of the *Public* principle.

This principle puts the public interest into central focus, emphasizing that "The ultimate effect of the work should be to the public good.". Software should only be approved if safety standards and specifications are met, quality of life and privacy are respected, and the environment is not harmed. It is mentioned that danger should always be disclosed to appropriate parties. Then, a call for fairness and avoiding deception in all statements regarding all of the artifacts of the profession is given. Disadvantages in any form should always be considered, whether they are in the form of economics, resources, physical impairments, or any other form that can hinder accessibility of software. Finally, volunteering of professional skills to good causes such as public education is stated.

**Principle II. Client and Employer** *Software engineers shall act in a manner that is in the best interests of their client and employer, consistent with the public interest.*

In this principle, honesty and transparency about qualifications, limitations, and issues is emphasized. Confidentiality ensures proper authorization and consensual use of resources. Also, collecting evidence about any issues - including project failure - and properly reporting it to client or employer is emphasized. Finally, conflicts of interests are to be avoided and reported to appropriate authorities, always prioritizing higher ethical concerns.

**Principle III. Product** *Software engineers shall ensure their products and related modifications meet the highest professional standards possible.*

The fifteen clauses that this principle comprises emphasize highest standards, ethical awareness, and professionalism in general. Engineers have to consider cost-effectiveness, achievable goals and objectives, while being aware of ethical, cultural, legal, and environmental factors. Thorough documentation, testing, debugging, and reviews are to be adequately applied to ensure project success. Other important practices mention an adequate documentation of respecting privacy, as well as always adhering to ethics and laws. Accurate data use and maintaining data integrity are further responsibilities mentioned along with maintenance of these artifacts. Finally, it is emphasized that all forms of software maintenance shall be treated with the same amount of care as development itself.

**Principle IV. Judgement** *Software engineers shall maintain integrity and independence in their professional judgment.*

This principle puts special focus on the ethical integrity, objectivity and transparency, while grounding technical decisions in support of maintaining human values. It is mentioned that professional objectivity in regards to evaluating software and related documents should be maintained. Furthermore, an deceptive and improper financial practices are to be avoided. Finally, conflicts of interest should be disclosed to all stakeholders and an encouragement to refuse to participate in professional practices where undisclosed potential conflicts are present is given.

**Principle V. Management** *Software engineering managers and leaders shall subscribe to and promote an ethical approach to the management of software development and maintenance.*

This principle is aimed towards managers and leaders specifically and emphasizes on responsible management, fairness, and ethical practices. Quality-focus, proper training, clear policies regarding confidentiality and transparency in management are a few key guidelines given. Fair conditions for software engineers should be upheld. Adequate assignments matching the individual's skills and opportunities for growth should be provided, while maintaining transparent and fair employment conditions. Finally, it is mentioned that ethical concerns of software engineers should be respected and that nobody should be asked to do any task that breaches this very Code.

**Principle VI. Profession** *Software engineers shall advance the integrity and reputation of the profession consistent with the public interest.*

This principle focuses on building an ethical organizational environment and upholding integrity. Compliance with this code is to be promoted. Engineers should advance public knowledge of their field and respect and promote the profession. Compliance with laws - as

long as there is no conflict with the public interest - is crucial. Furthermore, it is mentioned that responsibility for errors should be taken. Accountability should be demonstrated by addressing conflicts, or violations of this Code. Finally, it is mentioned that significant violations of this Code should be reported to appropriate authorities in cases where direct consulatation with the envolved people seems impossible, counterproductive, or dangerous.

**Principle VII. Colleagues** *Software engineers shall be fair to and supportive of their colleagues.*

This principle focuses on collaboration, fairness, and professional growth among colleagues. It is mentioned that colleagues are to be encouraged to adhere to this Code. Colleagues are to be assisted in their professional development and credit is to be given for others work. When reviewing work of others, it should be done in a objective, candid, and properly-documented way. Finally, it is mentioned that assistance of colleagues should also contain raising awareness of standard work practices, including policies, procedures, and security measures.

**Principle VIII. Self** *Software engineers shall participate in lifelong learning regarding the practice of their profession and shall promote an ethical approach to the practice of the profession.*

*This principle emphasizes lifelong learning, ethical conduct, and professionalism. It is mentioned that engineers should strive to improve their technical knowledge, ability to create high-quality products and documentation, and knowledge about relevant standards and laws. A bias-free, fair treatment of others is encouraged. Avoiding practices that breach this very Code are mentioned. Finally, it is expressed that a violation of this Code is inconsistent with being a professional in this field.*

CHAPTER 3

# Blind Spots

In this chapter, we consider critique from literature regards the discussed ethical concepts, focussing on Digital Humanism [cf. Section 2.3] and concrete software ethics [cf. Section 2.4]. We do this in order to identify *Blind Spots* in ethics of computer science. We classify a Blind Spot as an issue or open challenge that is not sufficiently addressed by existing software development ethical codes.

Discrediting the ethical codes or concepts is not the idea behind the process: The identified Blind Spots serve as anchor points where this thesis contributes and builds upon in order to overcome existing limitations of ethical paradigms.

This chapter is structured as follows: First, we discuss general issues and challenges related to Digital Humanism in order to identify problems and areas of interest that need to be addressed in order to improve the human well-being in an age of digitalization [cf. Section 3.1]. We then apply software ethic codes [cf. Section 2.4] in order to address these challenges [cf. Section 3.2]. This illustrates how software ethics address challenges of Digital Humanism. Furthermore, we briefly state open challenges in Digital Humanism. Then, we look at critique aimed towards ethics in computer science [cf. Section 3.3]. Once again, we identify open challenges. To conclude this chapter, we summarize the distinct identified Blind Spots in state-of-the-art software ethics [cf. Section 3.4].

## 3.1 Issues of Digital Humanism

Neidhardt et al. (2022) specify a non-exhaustive list of so-called “simple-complicated” issues that Digital Humanism has to address [47]. These issues are called simple-complicated because at a first glance, the solution sounds simple, but it in fact is often multi-facetted, faces conflicts of interest, or is just way more complicated than it initially seems. In the following, these issues are listed and summarized in our own words.

- *Interdisciplinarity.* We face negative phenomena of the digitalization, such as monopolization of the Web, issues related to automatisation of work, problems with respect to AI and decision making, the emergence of filter bubbles, the spread of fake news, the loss of privacy, and the prevalence of digital surveillance. Furthermore, the discipline of Informatics is often viewed from a problem solving side, while side- and long-term effects of its work are not always seen. The proposed solution is interdisciplinarity. This however, comes with its own challenges like common language and terminology or funding (touching different communities).

- *Humans decide.* As it is written in The Vienna Manifesto For Digital Humanism, "Decisions with consequences that have the potential to affect individual or collective human rights must continue to be made by humans" [61]. However, humans tend do simplify the world by building models in order to understand it. We even do so in complicated cases, and aid from machines and statistical models is often required. Therefore, the issue is when and how strongly we incorporate machines into our decision making: Is it just us humans deciding, even when we base our decisions on results produced by machines or algorithms? In fact, even if we decide to let machines do the thinking for us, we as humans take responsibility for deciding to let that happen.

- *IT for the good.* As discussed, Digital Humanism not only tries to minimize the downsides of thechnological advance, but also encourages technological advance in the interest of the societal good. This appears simple, but is complicated. When does society improve? Income equality and a growing market gap between companies are examples of problems that are ongoing during the digitalization. The causes of these societal developments are not obvious and hard to identify, and so is finding the correct technological (and political) answers to these problems.

- *Ethical technology.* We need to be aware of the potential impact that digital systems have. Therefore there is a need for guidelines for ethical designs of our systems. Such guidelines already are in place to a degree, as discussed in 2.4. However, it is not as simple as just blindly following these guidelines. We are facing big IT platform companies with data, financial resources, and outstanding expertise, that play a major role in AI research. Is this research really independent? Independence of research is critical, and so are reproducibility of results and transparency of funding. Furthermore, there is a problem of identifying if we are overall acting ethically at all. Ethical standards, paired with technological evolution, do not evolve on their own. In the work it is claimed that ethics alone may not be enough.

- *It is about the economy.* The economy of our planet has massively adapted with the rise of digital technologies. Due to automization, manual human work has become less important and machine work and data plays an important role for our economy. It is not obvious what role humans play in the economy. Is it the personalized data we leave in the web? Do we still provide value to companies in the traditional sense of labour, or are we mainly providing value with our personal data?

- *And about politics.* If humans, and human good, are the center of interest, we should let people partake in the future by voting or other participatory initiatives. However, in practical terms, it is difficult. Even though social media clearly provides a basis for a participative process, the effect can go in the opposite direction. Critical topics like who needs (and who is allowed to) to regulate platforms, should it be done on a regional or national scale, and should new infrastructures and services be built need to be addressed and definitely are not simple to solve issues.

- *We as academics.* This issue states that at first glance, academics provide solution to open questions, build models, and follow a scientific process, which in some cases leading to innovation or just making the world a better place - hopefully. However, the impact of our innovations and academic work is not always obvious. New technology also leads to automatization which means some jobs will disappear, so there is a lot of room for criticizing practices in technology and science. Specifically, the issue of "how and when to combine humans and machines" is metioned as complicated.

## 3.2 Addressing Issues

We now apply ideas, Principles, and Clauses of software ethics [cf. Section 2.4] in order to address issues in Digital Humanism. In this chapter, we regularly refer to The ACM Code of Ethics and Professional Conduct as *Code*, or *The Code*.

### 3.2.1 Interdisciplinarity

The ACM Code of Ethics and Professional Conduct recognizes: "The entire computing profession benefits when the ethical decision-making process is accountable to and transparent to all stakeholders. Open discussions about ethical issues promote this accountability and transparency." Furthermore, general ethical Principles of the Code, especially 1.1 [cf. Section 2.4.1], make sure to put human well-being first and essentially declare the public good as the most important consideration.

In doing so, the Code here addresses the issue that information technology is mainly seen from a problem solving point of view. The Code recognizes the impact of software and systems on society and human well-being.

However, open issues remain, such as common terminology across disciplines. Research landscape for interdisciplinary research is another issue, as researchers are often not specialized enough in one field to receive funding. In a world where software plays a central role in a wide array of professions, interdisciplinarity deserves special focus. While the issue of interdisciplinarity is not listed as a distinct Blind Spot in Section 3.4, it is implicitly addressed by Hofkirchner's concept of the *Global Sustainable Information Society*, in particular *Sustainability* [cf. Section 4.2].

### 3.2.2 Humans decide

The Code puts the public good and human well-being first. The issue stated is: How do we decide what is best for us? The problem seems to be that we already use technology to make such decisions: We rely on statistical algorithms and models that outperform humans. Multiple regression models are mentioned as an example, but artificial intelligence models come to mind as well. Furthermore, when attempting to solve problems or make decision, we often simplify in order to understand the world, thereby building a model that does not accuately depict the reality.

So, while we recognize that technology should be built in the best interest of human well-being, the issue of how does one decide what is in the best interest of human well-being remains an open challenge.

### 3.2.3 IT for the good

The Code recognizes the importance of the public good, stating "the public good is the paramount consideration." The issue is that what the societal good comprises is not a trivial question.

To some degree, this can be seen in the Code in Principle 1.2, Avoid harm: "When harm is an intentional part of the system, those responsible are obligated to ensure that the harm is ethically justified. In either case, ensure that all harm is minimized." When harm is an intentional part of a system, clearly the good has to be larger than the harm in order for the system to still be seen in favor of the public good. Furthermore, where do we draw the line? In medicine, an MRI scan can cause claustrophobic reactions [14], which could certainly be classified as harm. However, it seems apparent that the developers - including software engineers - of MRI scanners did not act unethically by developing said scanners because harm can be caused to patients. It can be argued that the insights gained from such a scan clearly outweigh the downsides, and as technology improves, the downsides decrease. As MRI scanners improve, the rate of claustrophobic reactions decreases [14].

Yet, it can still be discussed if the harm caused by an MRI scanner is "intentional" or rather an undesirable side-effect. Without going into details of the specific case, the next question to consider is if from an ethical perspective harm should ever be an "intentional" part of a system. How can it be ensured that this intentional harm is ethically justified. It is apparent that this remains an open challenge.

### 3.2.4 Ethical technology

The Code itself is built in a way that addresses the ethical conduct of all computing professionals, so everyone should in theory be responsible for one another. Rules 4.1 and 4.2 point out that appropriate action needs to be taken when other computing professionals violate the Code. However, as this issue states: Even when we develop

technology along the best ethical standards we have, is it enough? To what degree can we hold important stakeholders in big IT companies, or politics, accountable?

Also, the issue mentions that even when acting according to Kant's categorical imperative, can we guarantee a good ethical behaviour or a good outcome? As discussed in 2.1.1, there are objections to Kant's categorical imperative, like perfect duties not allowing exceptions and conflicting duties. As described in 2.1, we however consider diverse existing ethical perspectives, unifying them through diversity. The open challenge revolves around the bigger picture: Even if we act according to a unified ethical perspective that leads us to act ethically, is it enough?

One might be hard pressed to prove that we can guarantee a good outcome even with perfect ethics. Existing power relationships in society and economics influence the outcome. Are ethics alone good enough? If no, what is missing? This remains an open challenge.

### 3.2.5 It is about the economy

The Code recognizes that digitalization fundamentally changes the way we work, and live. Because of this massive impact, the Code points out the importance of ethical guidelines in the computing profession. Rule 3.7 points out that even simple systems can have large impact on society. Therefore, ethical guidelines need to be evaluated as systems integrate. Behaviour needs to be monitored, and if ethical standards do not exist yet for a new technology, they need to be developed.

Similarly as in the issues in regards to *IT for the good* [cf. Section 3.2.3], it remains unclear how to resolve the question of the role that humans play in the economy. Addressing the ethical ends of software engineering is crucial.

### 3.2.6 And about politics

The Code puts special focus on the responsibilities of computing professionals. While it recognizes the massive impact of the profession to society, how can interdisciplinarity be addressed by the Code? Politics, and people participating in shaping their own future, is increasingly influenced by digitalization, automization, and social media. Here, the Code can help in the design process of such social media technologies and global platforms, but the question remains of how one can evaluate that we are truly developing such artifacts in an ethical manner? Rule 3.7 of the Code does point out that integration of new systems has to be done with care, and behavioural observations of sytems need to be made in order to enrich our ethical understandings.

However, it might be hard to detect bad developments. Possible, the impact of a new technology might prove useful, but come with negative repercussions later down the life-cycle. Are new ways of communicating globally dangerous or do they just benefit? There are no trivial answers to questions like these. Furthermore, even if we detect bad developments: Once a system is in place on a global scale, discussions may end up

ineffective, yet costly. Addressing these ethical ends in regards to politics remains an open challenge.

### 3.2.7 We as academics

The Code addresses that all people are stakeholders computing in Principle 2.4.1. It is mentioned that they should focus on the public good and the societal good. We consider academics as important stakeholders in innovation and progress of professions. However, existing ethical software ethics fall short in addressing the effects of academic contributions. We believe that the ethical ends of academics and innovation remain an open challenge.

## 3.3 Issues of Ethics in Computer Science

Seth Aaditeshwar considers the ACM Code of Ethics and Professional Conduct and identifies matters unidentified by the code [53]. His main critique is that the code addresses ethical means, but largely ignores the topic of defining ends. He summarizes his insights as follows [53, p. 44]:

- The ACM Code of Ethics and Professional Conduct largely focuses on an ethics of the means to avoid harm, but does not clearly define ethical ends that computing systems should aim to achieve.
- The code is ineffective in flagging unjust and undesirable goals for which technologies are built or used.
- The code should embrace goals such as achieving equality and overturning unjust social and economic structures through technological inventions.

As seen in the discussion of the Code [cf. Section 2.4.1], the very first principle [cf. Principle 1.1] already outlines the the importance of contributing to the public good and human society. Principle 1.4 encourages inclusive and accessible technologies and practices. Then there is Principle 3.7, pointing to special care required for systems that become integrated into society. These principles address ethical ends, but the rest of the principles predominantly focuses on means. The means go into great detail, such as avoiding harm, promoting honesty, fairness, respecting privacy, foster work quality, taking responsibility, etc. The ethical ends however fall relatively short.

We recognize the that you cannot view the Principles of the Code in separation. The Preamble mentions that the Principles in Section 1 of the Code [cf. Section 2.4.1] outline "fundamental ethical principles that form the basis for the remainder of the Code". Yet, Aaditeshwar identifies that the Code stays very general in addressing ethical ends, making it hard to judge undesirable goals or even embracing positive goals.

First, we identify the Code's recognition for the public good and human society. However, the Code does not address in detail what is meant by that. Secondly, highly integrated, often cloud-based software solutions shape the communication and society. In such systems, ethical judgement is hard: Is the software system itself aimed towards a usage for the public good? Might it be misused? How can we judge the impact a highly interconnected software has into society? Therefore, as Aaditeshwar points out, we need a mechanism to actively "embrace goals such as achieving equality and overturning unjust social and economic structures through technological inventions".

We recognize this shortcoming of addressing ethical ends as an open challenge.

Now, we consider the The Software Engineering Code of Ethics and Professional Practice regarding the same issue. Again, we make sure not to view the Code in separation in order not to make errors of omission or commission. The full version of its Preamble recognized the significant opportunities to cause good or harm that software systems contain. It claims that "To ensure, as much as possible, that their efforts will be used for good, software engineers must commit themselves to making software engineering a beneficial and respected profession.". Furthermore, it claims that in order to conform to this commitment, software engineers shall follow this Code. Also, the full version of the Preamble re-emphasizes that the *Public Interest* is crucial to this Code, stating that "health, safety and welfare of the public is primary". Similarly to The ACM Code of Ethics and Professional Conduct, we recognize an early emphasis on the public good with statements like "The ultimate effect of the work should be to the public good.", or "Cooperate in efforts to address matters of grave public concern caused by software, its installation, maintenance, support or documentation." [cf. Principle I. Public]. Compared to The ACM Code of Ethics and Professional Conduct, we recognize this as a slightly stronger effort on addressing the ends.

Even though the interest of the public good is mentioned as the crucial concern over all of the Code, the other Principles and their Clauses are directed at addressing the means. In this Code, we find a vast spectrum of clauses that deal with various types of avoiding harm, respecting peers, taking responsibilities, ensuring a high quality product, judging fairly, conforming to standards, promote the profession, etc. However, we ultimately end up with a situation where addressing the ends falls short. Other than the all-encompassing public health, safety and welfare, it remains unclear what ethical goals computer systems should achieve. We conclude that the Code can not sufficiently identify software systems that are geared towards undesirable goals. Yes, ethical concerns towards the software may always be raised, but there are cases where unjust goals were credited to be beneficial to society [38, 56]. In cases like these, concerns about surveillance capitalistic concepts and data collection were allegedly justified by claiming that these software systems benefit the welfare of the users. Often, the intentions of a software technology are unclear to the public, but also stakeholders and developers. This complicates issues further, re-inforcing the requirement for defining clear ethical ends that software should strive for. Finally, the Code falls short in the ability to actively promote the requirement for goals that challenge existing unjust social and economic systems. We recognize goals like a united humanity,

recognizing environmental health and developing software for humans as important ends that a Code should address.

Finally, we now focus on another issue in Software Ethics. Principle 1.2 from the ACM Code of Ethics and Professional Conduct which states: "When harm is an intentional part of the system, those responsible are obligated to ensure that the harm is ethically justified. In either case, ensure that all harm is minimized."

This statement raises questions, as one could ask what what *ethically justified harm* even means. Is it harm that is outweighted by a larger good in total? A surgical procedure causes immediate damage to the human body, but is in many necessary in order to improve that human's well-being or even save their life. One can argue similarly for the MRI case [14] discussed earlier [cf. Section 3.2]. These examples would probably represent an ethically justified harm, but it is not trivial to come up with such an example in the software engineering context. Just providing a small amount of personal data in order to be able to use a software might have privacy implications and usage of personal data that was not intended by the user of the software, violating at least some aspect of his well-being. From a standpoint of care-based ethics, this could potentially be a large problem, as the individual case and context will differ.

Furthermore, not only ethically justified harm is potentially unclear, but also *harm as an intentional part of the system.* Again, in the case of a surgical procedure this could probably be argued to be the case, but for what artifact in the software engineering context would it be appropriate to be designed with harm as an intentional part of the system? We cannot clearly identify if the data-privacy example would even qualify here. In the mentioned case, following ethical guidelines of the Code, the data would only be collected as an act of necessity. This minimizes privacy issues to the least amount necessary, and not cause harm with intention. Shifting our focus to another aspect, software for military purposes comes to mind as an example where harm is an intentional part of the system. This would respresent intentionality, but then again, the note that *In either case, ensure that all harm is minimized.* does not really match the example. Such a software would intentionally be designed to cause harm, and not to be minimized.

To summarize, we recognize this Principle as challenging, necessitating further study beyond this thesis.

## 3.4 Identified Blind Spots or Open Challenges

In this Section we formulate Blind Spots on the issues discussed in the last Sections [cf. Sections 3.2, 3.3].

We recognize the first Blind Spot as the main open challenge in regards to addressing ethical means and ends in Digital Humanism and software ethics. The Blind Spot is summarized as follows:

**Blind Spot 1** *Software Ethics, while thoroughly addressing the ethical means of software development fall short in the ability to address ethical ends. Full-fledged Software Ethics must be capable of addressing both the ethical means and the ends. This results in a responsible software process that also embraces desirable, ethical goals. Ideas like the public good, healthy human-human relationships, an alliance with nature and technology made for humans should be key aspects that Software Ethics address. Addressing these desirable goals should be done to an extent that makes identifying and flagging Software with unjust goals possible using these Software Ethics.*

We consider this Blind Spot the key issue in existing state-of-the-art Software Ethics. It is not a detail that is being criticized here. Software Ethics can clearly identify malicious development practices, unjust treatments of stakeholders, and raise awareness for several issues like transparency, data privacy, etc. These aspects add to making Software development a respected profession. However, in the big picture, the encouragement to actively embrace ethical goals that benefit society falls somewhat short: Goals can remain unclear or incomplete, or even disguised as beneficial for society. Enforcing clear, detailed ethical descriptions of the ethical ends of Software systems has to be of key focus.

We now consider our second, more nuanced Blind Spots. This Blind Spot reflects open questions about harm as intentional part of software systems. We formulate this Blind Spot as follows:

**Blind Spot 2** *Software ethics leave open questions about software systems where harm is an intentional part of the system. What is a computer professional's responsibility when harm is an intentional part of the system? How can harm which is an intentional part of the system be ethically justified? Is it even moral to justify (even though minimizing) harm that is an intentional part of the system, or should an ethical framework prevent further advancements on such systems?*

We believe that this Blind Spot is insufficiently addressed by existing Software Ethics. Consequences of systems where intentional harm is an integral part might be fatal. There should be no room for faulty usage of ethical systems regards such a sensitive topic. We believe that this Blind Spot warrants an investigation beyond this thesis.

CHAPTER 4

# Overcoming paradigms

In this chapter, given humanity takes its ethical responsibilities seriously, the need for a transformation of our techno-eco-social system is discussed as well as how such a transformation can be accomplished. This transformational, visionary approach is discussed in this chapter. Within this chapter, relevance to the topic of thesis is established: Overcoming existing ethical paradigms and limitations.

This chapter is structured as follows. First, the circumstances (a global crisis and *The Great Bifurcation*) that lead to the necessity that humanity needs an evolutionary step are discussed [cf. Section 4.1]. Then, the concept of the *Global Sustainable Information Society (GSIS)* [33, p. 136ff] by Hofkirchner is provided as a vision that defines the conditions to make the evolutionary step [cf. Section 4.2]. Furthermore, additional important concepts, starting with *Informatisation*, *Informationalisation* and their variations are introduced [cf. Section 4.3]. Then, *Global Consciousness* [cf. Section 4.4] and *Reorganisation* are explored [cf. Section 4.5]. Finally, relevance of GSIS for ethical studies in computer science is given [cf. Section 4.6]

## 4.1 Crisis And The Great Bifurcation

Scientific literature identifies that human impact causes massive threats for the whole future of the planet earth and our civilisation: Randle and Eckersley claim that there is "growing scientific evidence that humanity faces serious risks of a catastrophic coalescence of environmental, social, economic and technological threats within the next 100 years" [51]. Ehrlich and Ehrlich identify the impact and risks civilisation causes on the planet as the human predicament [16]. They list several points on the predicament, including (but not limited to) population growth, warming athmosphere and disrupted climate, and economic growth among the already rich. Furthermore, Ehrlich and Ehrlich state that "determining how to prevent [the predicament] from generating a global collapse is perhaps the foremost challenge confronting humanity" [15]. Gleick claims that "it now

seems unavoidable that human misery, caused by our own actions, will be more extensive than it otherwise would have been and worse than it should be" [23].

In this thesis, the state where global threats to humanity cause serious risks to the societal and economic future is referred to as a *crisis*. Note that the term crisis is also used interchangably in a more abstract, systemic point of view, where a crisis causes amplified fluctuations to a system. The crisis of our society calls for a solution, calling for responsibility of humanity to act on it. This means that the concepts discussed as follows apply to systems theory in general, but should be applied in the real-world context to our techno-eco-social systems.

Instead of trying to exhaustively list crises humanity is facing, Wolfgang Hofkirchner [33, p. 150ff] merges different kinds of crises on the epistemic layer, and he does so along the social relations. He identifies prevailing logics of our social systems in three main categories: Human-to-human, human-nature, and human-technology. The emerging risks to these logics require a new step in the sociogenesis, which is an evolution of a social system. In the following, the three main categories, their prevailing logics and the possible next step in the sociogenesis [33, p. 150ff] are briefly discussed.

| | | **sociogenesis** | |
|---|---|---|---|
| | | **logics of the second step** | **logics for a third step** |
| **social relations** | **among humans** | self-centricity: exceptionalism | pan-humanism: united humanity |
| | **with nature** | anthropocentricity: colonisation | anthropo-relational humanism: alliance with nature |
| | **to techno-logy** | power-centricity: super empowerment | digital humanism: ethically aligned technology |

Figure 4.1: Table of social relations. Current second step (current prevailing logics) and their proposed third step. [33, p. 150]

The first social relation is the human-to-human relation, that is societal relations. This relation is determined by a logic of *self-centricity*, that is exceptionalism. The proposed solution is replacing self-centricity with *pan-humanism*, that is the idea of a united humanity.

The second social relation is the human-to-nature relation, which relates humans to nature. This relation is determined by a logic of *anthropocentricity*, that is colonisation of nature. The proposed solution is replacing anthropocentricity with *anthropo-relational humanism*, that is alliance with nature.

The third social relation is the human-to-technology relation, linking humans and technology. This relation is determined by a logic of *power-centricity*, that is super empowerment. The proposed solution is replacing power-centricity with *Digital Humanism* [cf. Section 2.3], that is ethically aligned technology.

As Hofkirchner describes in detail [33, p. 150], the prevailing social relations are rendered dysfunctional by the new requirements of *Globality*, *Sustainability*, and *Informationality* [cf. Section 4.2]. Earlier in his work [33, p. 117ff], he describes how the conditions of a crisis and consequently a state of bifurcation (*The Great Bifurcation*) demand an evolution from an ontological point of view: "The trajectory of social systems evolution is on the point of reaching a bifurcation between existence and extermination – the Great Bifurcation." A bifurcation means that there is a space of possible trajectories, or paths, that a system can follow. The system however, at the tipping point of this space of trajectories, has to follow one of the paths.

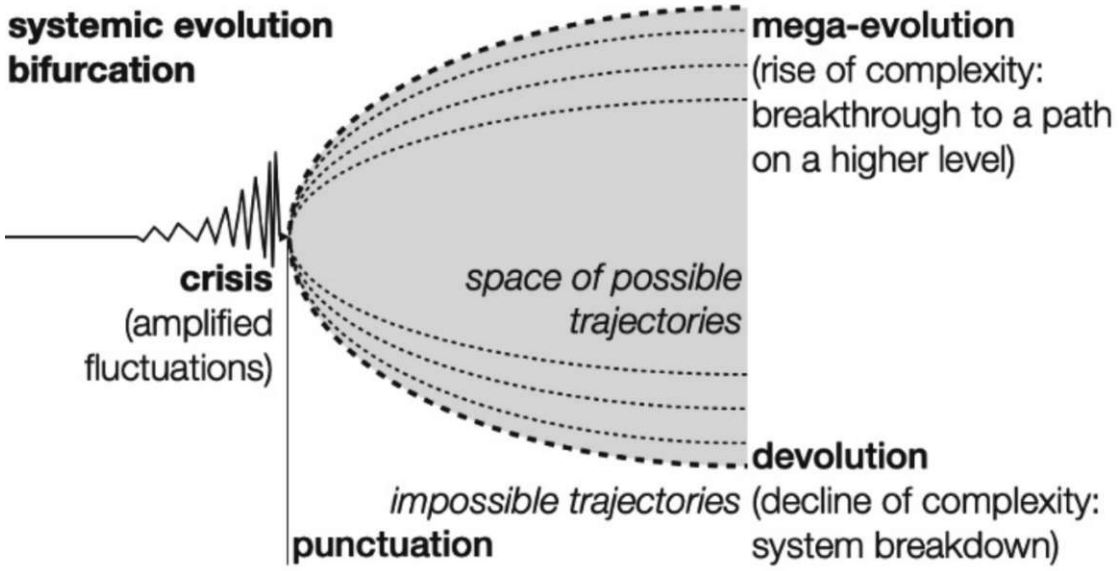


Figure 4.2: Bifurcations. Mega-evolution. [33, p. 119]

In Figure 4.2, the crisis leading up to such a tipping point is modelled as fluctuations. These fluctuations represent the consequences of the global challenges humanity is facing. Furthermore, the trajectories can either lead to a so-called *mega-evolution* [29, 50], or a *devolution*. In the figure, a mega-evolution is represented by the uppermost path and results in a systemic evolution, rise of complexity, and ultimately reaching a higher level from a systemic point of view. In this scenario, the global challenges are mastered. The resolution to these global challenges is the *Global Sustainable Information Society (GSIS)*.

A devolution, depicted by the lowermost path, represents a trajectory of systemic failure, a decline of complexity, and ultimately leads to the system breaking down.

Considering the possible scenarios of the Great Bifurcation, it is in humanity's best interest to aim towards a positive outcome rather than a systematic breakdown. Using the main ethical perspectives, [cf. Chapter 2] it can be argued that humans have an ethical responsibility to resolve the Great Bifurcation towards the best possible outcome. Whether one subscribes to duty-based, virtue-based, consequentialist, principle-based, or care-based ethics, or identifies as a humanist: Aiming towards mastering the global challenges is not just a goal, it is a necessity. Opting out of contributing to mastering the global challenges stands in direct conflict with moral responsibilities.

## 4.2 Global Sustainable Information Society

Hofkirchner [33, p. 136ff] defines a *Global Sustainable Information Society (GSIS)* as a society that demands the following three properties.

1. Globality. "Regarding the spatio-temporal dimension, globality would be, for the first time in history, an integration at the level of all humanity."
2. Sustainability. "Regarding the level of organisation, sustainability would be a reorganisation between the interdependent social systems in which humanity is currently scattered. It would also involve reorganisation within the social systems such that sociogenic dysfunctions of the cultural, political, economic, ecological and technological subsystems can be kept below a threshold that would endanger the continuation of sociogenesis. That task requires fully integrating all of humanity."
3. Informationality. "Regarding the state of collective intelligence, informationality would be a capacity to create information that is required for the reorganisation [34]".

The GSIS is a vision and describes an ontological model aiming towards mastering the global challenges and leading the Great Bifurcation to its uppermost path, therefore resolving crisis and reaching mega-evolution. That being said, the GSIS does not in itself provide a blueprint for the new society. It merely states that without fulfilling the properties, i.e. Globality, Sustainability, and Informationality, the global challenges can not be mastered. This in turn would mean that we possibly face a breakdown of the system and can not proceed to mega-evolution.

This mega-evolution ultimately represents a new *Meta-/Suprasystem*, which is the GSIS, requiring a transformation and emergence of our current techno-eco-social systems.

## 4.3 Informatisation, Informationalisation, and Informationality

In order to discuss ethical relevance of Information and Communications Technology in the context of the GSIS, a few terms are introduced.

Nora and Mirc introduce the term *Informatisation* or *Informatization* as the extent to which a society becomes increasingly responsive to information [49]. As informatisation increases, society experiences less friction in interacting with information. Here, less friction means that due to smarter and more advanced technological artifacts it becomes easier, more accessible, and more efficient to handle information.

Informatisation provides potential for the GSIS, in the sense that it can qualitatively contribute to Globality, Sustainability, and also Informationality by reducing friction towards Information. However, there are also antagonisms of Informatisation. Practical systems that are in use today, including social media, are at least partially used in ways detrimental to the GSIS. Phenomenas such as fake news, disinfotainment, and large language models are just a few examples of misguided developments that Informatisation plays a role in. This contrast between potential and antagonisms makes Informatisation ambiguous.

Hofkirchner introduces the term *Informationalisation* [33, p. 144] as the capability of a system to create information such that it can be used for self-organization. Informationaliation is the capacity to raise intelligence and achieve a collective intelligence. If Informationalisation is contrasted with Informatisation, it can be seen that Informatisation is the basis that allows for Informationalisation. However, Informatisation needs to be tamed in a way such that the negative aspects get reduced and the positive ones remain. If Informatisation is tamed in this way, it can be said that Informatisation is being harnessed for Informationalisation. In doing so, unreflected or misguided artifacts of Informatisation are removed and the Informationalisation helps to achieve a collective intelligence. Hofkirchner [33, p. 200] states: “Informatisation refers then to the Is, and Informationalisation to the Ought.”

Note that the term *Informationality* [33, p. 144] is used to describe the ongoing state or progress of Informationalisation of a system.

## 4.4 Towards Global Consciousness

In order to introduce the concept of *Global Conciousness*, first, the concept of a *Global Brain* is introduced. “Informatisation can be viewed as the process of laying the material connections of a Global Brain” [33, p. 200]. This means that Informatisation provides technical grounds for the Global Brain, representing the technical connectedness of the planet similar to that of a nervous system in biology. This Global Brain, guided by a process of Informatisation, is already partly in place. Information Technologies like the Internet Of Things, Big Data, Artificial Intelligence or even just the Internet as a global network in general provide a level of Informatisation that provides grounds for a Global Brain to emerge.

The Global Brain, however, can be viewed as the technological foundation required for a Global Conciousness. Hofkirchner [33, p. 201] states: “However, a Global Brain is not

the same as a global "consciousness", a global "conscience", or a global "mind" – which are rather the process and product of informationalisation".

Informatisation alone can only be sufficient for a Global Brain, but not for a Global Consciousness. The Global Brain is built on the underlying Informatisation process. It does not bring about an informed society or raise congitive intelligence on its own: Along with the potential benefits, unreflected and potentially unwanted "information" is provided and spread via this global brain. A Global Brain can be viewed as the (ongoing) Is, while a Global Consciousness would be the Ought. In order to achieve Global Conciousness, the Global Brain needs to be designed in a certain way such that the underlying Informatisation process is being harnessed for Informationalisation, keeping the goal in mind to raise collective intelligence. By harnessing the Global Brain and thereby providing it with an ethical foundation itself, it becomes a Global Conciousness. It can also be said that due to Informationalisation, collective intelligence emerges to the next level. This in return represents the topmost path in the Great Bifurcation, resulting in a rise of complexity and systemic evolution.

## 4.5 Reorganisation, Self-Regulation, Systemic Evolution and Sustainability

As Global Conciousness emerges, the underlying systems emerge as well, bringing about systemic evolution. It is important to note that there is no concrete blueprint for this new *Meta-/Suprasystem*, which is the GSIS. This Suprasystem however subsumes and *reorganises* all currently existing social (sub-) systems. The resulting Suprasystem represents a new unit, containing many sub-systems which are reorganised in a way to serve the Suprasystem. Due to the reorganisation, not only the Suprasystem itself is in an emergent state, but all of the containing subsystems are.

Even though there is no definitive blueprint for such a Suprasystem, it is important that it still needs to be designed respecting adequate ethical values. This is where Informationalisation (going beyond the unharnessed Informatisation) is needed and ultimately contributes to.

Figure 4.3 shows such a transition to a new Suprasystem and how Reorganisation is designed. Agents represent systems that can self-regulate through the up and down arrows within them. These agents further communicate with co-systems using the Suprasystem. Self-regulation in this way provides Sustainability, which is an emergent level of organisation [cf. Section 4.2]. Furthermore, unity through diversity is achieved with this approach: Diverse systems are united via the Suprasystem. This aspect alone would only comprise unity in diversity, but the design also makes sure that diversity in unity is present: Through reorganisation and self-regulation, the diverse subsystems shape the existence of the Suprasystem, so neither can exist without the other. More in-depth explanations and references are found in Hofkirchner's work [33, p. 141].

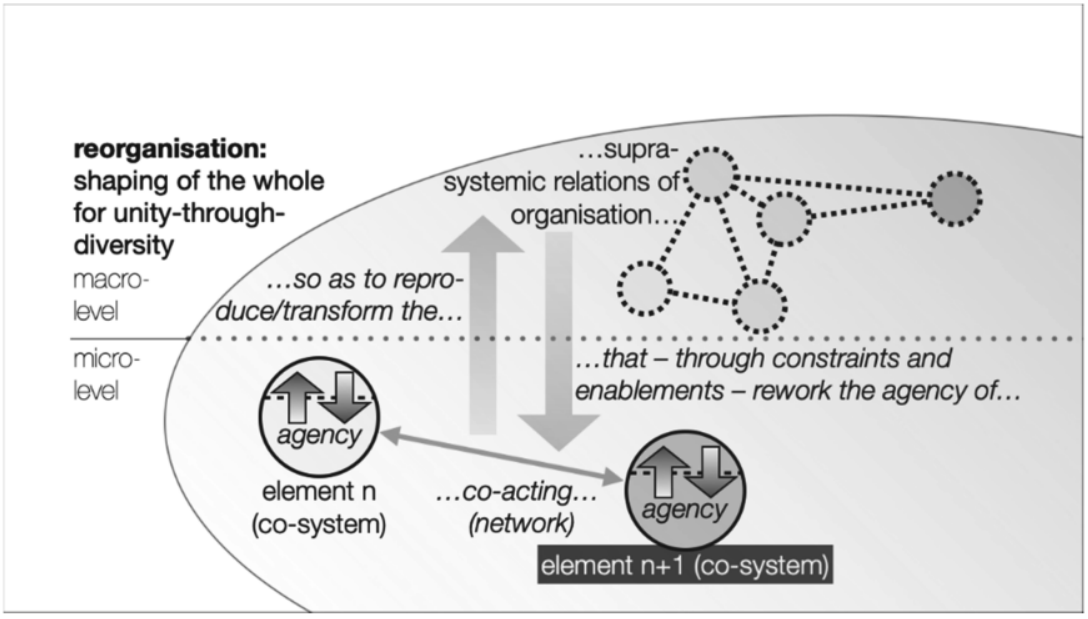


Figure 4.3: Reorganisation [33, p. 141]

## 4.6 Global Sustainable Information Society and Information and Communications Technology

Within the vision of the GSIS, fulfilling the properties that society demands in order to master the global challenges is the most important task. However, in the context of this thesis, special focus on Information and Communications Technology (ICT) is set. Therefore, a connection between the concrete moral concerns of Software Engineering and the GSIS is drawn. Note that when using ICTs in this thesis, we include Software Engineering, Artificial Intelligence, etc., but also general Computer Science. Recognizing the GSIS as a foundational and necessary concept to advance to the evolutionary step in the Great Bifurcation leads to ethical responsibilities in the Software Engineering domain.

First, it is established why and how a connection between the GSIS and concrete applied ethics in the field of ICTs is drawn: When going back to the crisis 4.1, reasons for this connection are already found in the third social relation, that is the human-to-technology relation. This relation proposes replacing power-centricity with ethically aligned technology, i.e. Digital Humanism, as a solution to advance to the next step in the sociogenesis.

Next, it is discussed how the ICTs can best be integrated into the GSIS: Here, it seems apparent that the GSIS property that ICTs can contribute to best is Informationality. Note that it would be wrong to dismiss concrete responsibility towards Globality and

Sustainability, but Informationality is where ICTs play a direct role in contributing to the GSIS, as ICT professionals actively produce artifacts which shape Informatisation. These artifacts then in a secondary step can shape Globality and Sustainability. Deciding on developing concrete artifacts can positively harness this Informatisation towards Informationalisation. Furthermore, designing and implementing artifacts that are not ethically aligned are in direct conflict with the GSIS. This means that there is a present responsibility to reconsider, redesign or even reject developing artifacts that are not ethically aligned towards Informationality.

Since one of the three properties of the GSIS is Informationality, it can be said that if one subscribes to the GSIS as a concept and a vision, there is a moral responsibility for any ICT professional to contribute to Global Conciousness. This can only be achieved by designing ethically aligned technology, in the spirit of Digital Humanism, utilizing the Global Brain in a way that contributes to the emergent state of collective intelligence.

Therefore, it is concluded that existing concrete ethical foundations in the field of ICTs demand a link to the GSIS, especially contribution to Informationalization (and therefore Global Conciousness). Disregarding foundational concepts of the GSIS in these areas should be seen as a breach of ethical contract. Even just opting out of this contribution to a GSIS might result in a systemic breakdown. Therefore, on this basis, opting out is not an ethical option.

What is proposed here represents a paradigm shift. ICTs, Software Engineering in particular, have contributed to Informatisation since its existence. Processing data, automation, information sharing, connectivity and much more have positively contributed to humanity's well-being and shape our lives radically. However, privacy and security concerns, misinformation, surveillance, environmental impact, etc., represent real challenges to humanity's future.

If we stay on the path of Informatisation (as opposed to Informationalisation), the potential for a serious change is greatly diminished. Ethical concerns towards a Global Conciousness should be taken seriously. ICT professionals should not be able to opt out of the responsibility to contribute towards a positive resolution in the Great Bifurcation. Aiming towards well-being for the presence and future of humanity has to be the foundational goal and an ethical stance that humans should take. In particular relevance for this work, every ICT professional should consider contribution to Informationality and the Global Conciousness as a foundational ethical principle within their profession rather than blindly contributing to Informationality and the Global Brain. Moreover, because of the techno-eco-social transformative appraoch of the GSIS we argue that responsibility to contribute to Informationality alone falls short. All three properties of the GSIS should be respected in order to positively resolve the crisis. As software artifacts offer potential for large-scale implications in regards to all of the three properties, we argue that breaching GSIS properties is not an ethical option for the software professional.

CHAPTER 5

# Ethical Guidelines

The objective of this chapter is to equip the reader with ethical *guidelines* and *pillars*. They are derived from GSIS concepts and serve as ethical criteria that can be applied in real life scenarios by professions in information technology. In this way, this chapter addresses Blind Spots in software ethics. Furthermore, they aim to promote awareness towards interdisciplinarity skills for software engineers.

This chapter is structured as follows. In a first step we introduce the core concepts *common good* and *alliance* [cf. Section 5.1]. Then, we consider the prevailing social relations and their issues [cf. Section 5.2]. By examining how to overcome them, we formulate guidelines to follow in order to achieve the next step in the sociogenesis. We proceed by considering the properties of GSIS [cf. Section 5.3]. They represent key ethical pillars to adhere to in times of the Great Bifurcation. Based on these pillars, we then formulate guidelines to follow in order to achieve mega-evolution. Finally, we revisit Blind Spots [cf. Chapter 3] and discuss ways that the introduced guidelines overcome existing ethical paradigms of software engineering [cf. Section 5.4].

## 5.1 Introducing Core Concepts

We begin by introducing the concept of the *common good* as follows.

*Common good*s (also referred to as *the commons*) [33, p. 49] are societal entities that are produced and consumed by common actors within the society.

Note that literature offers multiple definitions for commons or the common good. However, for the purposes of this work, we adopt Hofkirchner's concept of the commons as in his vision of the GSIS [33, p. 49]. This is largely a matter of practicality, as the argumentation for the ethical guidelines is closely related to GSIS concepts regardless.

We now outline important properties and ideas of the commons.

Common goods

- include, but are not limited to physical goods
- can also be ideas, systems, or processes
- represent what is best for society
- should not be kept private: They are common in the sense that there is common access to them
- are produced by actors within a society for the society, and this makes them common.

Examples of common goods that we see in our current time include air quality, access to electricity, a country's legal sytem, healthcare, etc.

Furthermore, the concept of an *alliance* between certain actors is critical for this Section and therefore introduced here. There is no intention to provide a clear blueprint or plan on how such an alliance is structured. What can be said though is that as a basis, the foundation of the GSIS, including structures like the Global Consciousness [cf. Section 4.4] is used to utilize these alliances.

Much rather than a complete blueprint it is important to recognize the emergent synergy from such an alliance and its benefits. An alliance represents a recognition of establishing a mutually beneficial and harmonious relationship between the actors. Characteristics of an alliance include, but are not limited to:

- Recognition of an aim towards harmony and equality of participants in an alliance.
- Unity through diversity. Not only does the unity (of actors) benefit from the diversity (of actors), but the diversity (of actors) also benefits from the unity (of actors).
- Reduction of power of one actor over the other. In an alliance, actors work towards common goals instead of abusing power over each others. Such a reduction of a power-centric relationship reduces friction within the system, making the system more efficient.
- Collaboration in order to work towards common goals, such as a sustainable future.
- Shared responsibilities and a division of labor
- Knowledge exchange and learning

When this thesis considers alliances, it is suggested that participation in the alliance is mandatory. Neither party can simply opt out of the alliance and thereby opting out of the (moral) responsibility to respect the respecting partner of the alliance. Opting out of such an alliance would breach emergence of the next level in the sociogenesis [cf. Figure 4.1].

## 5.2 Social Relations Criteria

With regard to making progress in the social relations, the three prevailing social relations of self-centricity, anthropocentricity, and power-centricity [cf. Section 4.1] can no longer be supported as such. Instead, they should be replaced by pan-humanism (united humanity), anthropo-relational humanism (alliance with nature), and Digital Humanism (ethically aligned technology), respectively. This replacement enables and promotes progress towards the next level in the sociogenesis. We now present three guidelines, one related to each of the three prevailing social relations. For each guideline, a brief list of issues that are being addressed is given. Furthermore, benefits and core ideas of these guidelines are listed.

### 5.2.1 Human-Human Relations

We now introduce the human-human guideline.

**Guideline 1** *Human-human relations need to be supported in a way that discourages self-centricity and encourages a united humanity (pan-humanism). In particular, denial of affiliation to a social group is to be discouraged, while an inclusive logic that encourages living together for the greater societal common good is to be encouraged.*

Negative effects that get addressed:

- Human exploitation
- Human discrimination
- Social gaps such as rich and poor

Positive effects:

- The societal common goods are encouraged

This guideline positions itself against self-centricity and aims for a united humanity. As self-centricity is being reduced and unity is encouraged, common goods are emphasized as a product of this process. Although this guideline is not a concrete blueprint, it

addresses ethical ends and serves as a general concept to fight human exploitation and social inequality. This guideline can be seen as an alliance between humans in society.

In particular for this guideline, common goods related to human-human interaction are encouraged, i.e. laws, a government, equality of rights, non-discrimination, etc.

### 5.2.2 Human-Nature Relations

We now introduce the human-nature guideline.

**Guideline 2** *Human-nature relations need to be supported in a way that discourages colonization of nature and encourages an alliance with nature. In particular, usage of nature without intending to recognize and preserve nature as a functioning system itself is to be discouraged, while recognizing and encouraging a harmonious interaction with the nature as a system is to be encouraged.*

Negative effects that get addressed:

- Global warming
- Depletion of resources
- Reduction of biodiversity

Positive effects:

- Alliance with nature as a common good is encouraged

This guideline requires recognition that colonization and exploitation of nature can not lead to a fruitful future of humanity. Humans are dependant on the planet and its resources and therefore need to exit the idea that we solely use nature to our benefits. Although this guideline is not a concrete blueprint, it addresses ethical ends and serves as a general concept to fight exploitation of nature. A recognized alliance with the nature has to be our goal, i.e. a synergy that benefits both actors of the relation. Both, humanity and nature are benefitting from this guideline in order to provide a sustainable future - for the planet, and for humanity.

In particular for this guideline, common goods related to human-nature interaction are encouraged, e.g. fresh, breathable air, access to healthy food, sustainable agriculture, long-term climate stability, and ecological balance.

### 5.2.3 Human-Technology Relations

We now introduce the human-technology guideline.

**Guideline 3** *Human-technology relations need to be supported in a way that discourages power-centricity and encourages Digital Humanism. In particular, promoting technology without recognizing humanity and nature as a functioning systems itself is to be discouraged. Recognizing and encouraging a harmonious interaction between technology with humans and nature as systems is to be encouraged.*

Negative effects that get addressed:

- Misguided information technology
- Unethical technological advances

Positive effects:

- Ethically aligned technology as a common good is encouraged

This guideline outlines that technology which enables power-centricity is to be opposed. It is not a sustainable approach to design technology that over-benefits those who are in charge and in power. Technology built for humans means that technologies themselves become commons. Technology that discriminates against humans in any shape or form needs to be prevented and replaced by technology in the best interest of the society. Although this guideline is not a concrete blueprint, it addresses ethical ends and serves as a general concept to fight unethical technological advances.

In particular for this guideline, common goods related to human-technology interactions are encouraged, e.g. access to communication technology such as the internet, infrastructure, encouragement of information rather than misinformation, and technology designed for the benefit of human society.

### 5.2.4 Transitivity Property of Relations

When these three guidelines are in place, from the general understanding about the characteristics of an alliance between actors, the guidelines also follow the rules of transitivity. This is based on the recognition of an aim towards harmony and equality of actors in an alliance: No participant in an alliance should be in power of the other, so they mutually benefit from each other. If humanity and nature are actors on a level, and humanity and technology are actors on a level, it therefore follows that nature and technology are also in an alliance. This means that technology needs to benefit a sustainable nature and nature's resources need to be used to design such technology (which in return benefits nature, or humanity).

## 5.3 Pillars of GSIS Criteria

The three pillars of Globality, Sustainability and Informationality [cf. Chapter 4.2] that society demands, can not be opted out of. In the context of ethical responsibility we consider it a breach of guidelines to actively undermine support for any of these three pillars.

As for relevance in information technology: Software should actively be developed with the goal in mind to adhere to pillars of the GSIS. The pillars need to be recognized as key principles that define desirable ethical ends. Software development should not breach Globality, Sustainability, and Informationality goals of the suprasystem GSIS, as working towards the GSIS is required for humanity to reach mega-evolution rather than systemic breakdown. Therefore, in this section we formulate ethical guidelines from these pillars.

First, each of the three pillars is briefly examined in regards to its historical evolution and possible continuations or goals. Each of these pillars is supported with a figure [cf. Figures 5.1, 5.2, 5.3] showing the evolution of time (horizontal axis) and the evolution of the respective pillar (level of Globality, Sustainability, or Informationality on the y-axis). Note that the concepts of the y-axis in the figures are understood to be inclusive in an emergent way: Every level includes the underlying concepts, in a modified and improved way. Furthermore, every new level builds a new concept on top of the existing paradigms from a systemic point of view.

From there, guidelines in regards to the respective pillars are stated.

### 5.3.1 Globality Pillar

Figure 5.1 shows the evolution of the globality, including the proposed third step, a new *world system* that builds on nomadism and territoriality with a *new cosmopolitanism*, or *glocality*. The key systemic evolution called glocality is a system that connects locality with globality. Within this glocality there is still nomadism. This nomadism can however be improved by supporting it with information technology, traffic and travel technology, etc. As for territoriality, it is also still existant and is improved by glocality. This can be achieved by integrating the existing city-states, empires, nation states, unions, etc., into a new world system.

The solution to the next step towards globality is not to simply replace existing nation-states with one world state. Glocality proposes an integration of nation states on a truly global level.

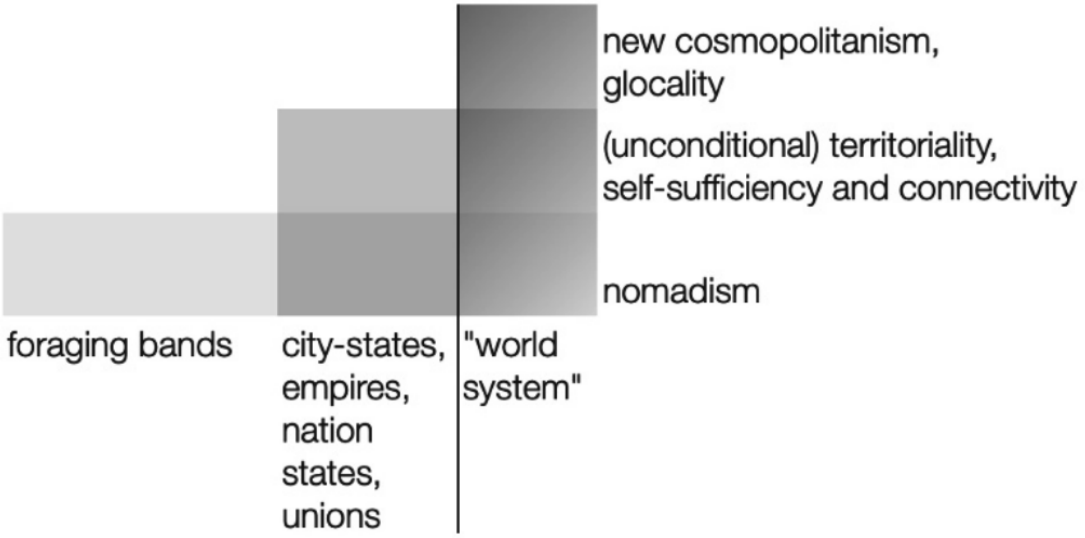


Figure 5.1: Evolution of Globality. [33, p. 140]

We now introduce the globality guideline.

**Guideline 4** *Reaching the next step in Globality needs to be supported in a way that aims for a new glocality by linking the local with the global. Efforts to disrupt or slow down technological advance towards Globality need to be actively opposed and represent a breachment of this Guideline.*

### 5.3.2 Sustainability Pillar

Figure 5.2 shows the evolution of sustainability, including the proposed third step, a *wise society* that builds on undifferentiated mythical community and diversity of identities with a *unitas multiplex*, or integrating differences. The key systemic evolution, which is unitas multiplex, is a new organisational level that elevates the means-end relationship further [46]. This organisational level requires integrating all of humanity with all of its differences, strengths, weaknesses, and visions for a fruitful future. Within the unitas multiplex, unity in diversity is encouraged in order to agree upon common means and ends. Neither means nor ends are set in stone, and an integration of differences of actors leads to concrete ends and means. This approach enables a reorganisation of the social systems and holds social dysfunctions at bay such that they do not hinder humanity from reaching a positive outcome in the Great Bifurcation. We achieve Self-Organisation on a systemic level.

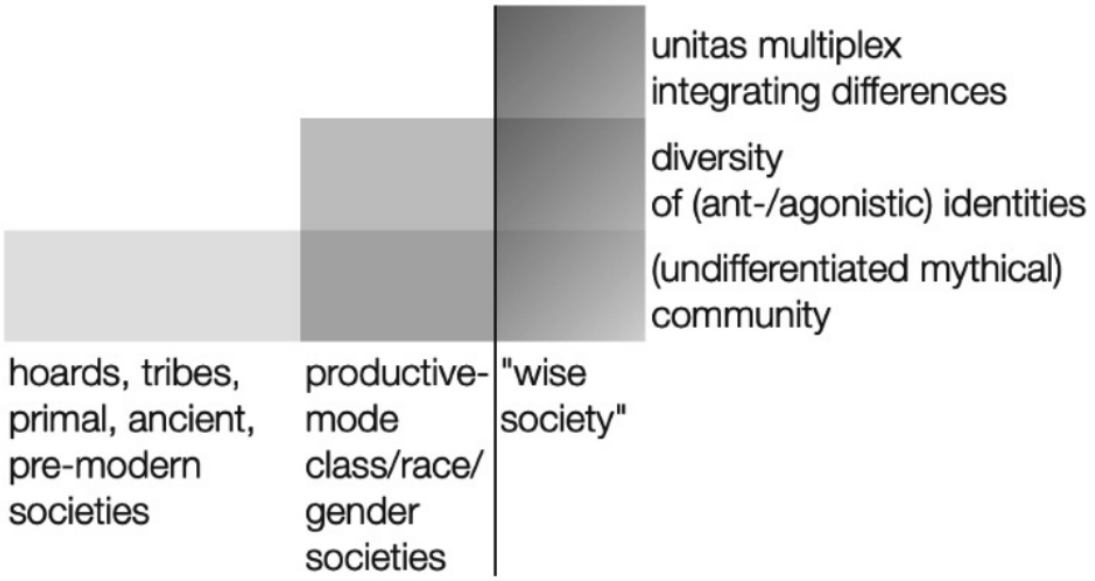


Figure 5.2: Evolution of Sustainability. [33, p. 143]

We now introduce the sustainability guideline.

**Guideline 5** *Reaching the next step in Sustainability needs to be supported in a way that encourages diversity and differences of actors. This unity in diversity supports systemic evolution in a way that commonly has its actors agree upon means and ends. Efforts to disrupt or slow down technological advance towards this new level of reorganisation and integrating all of humanity need to be actively opposed and represent a breachment of this Guideline.*

### 5.3.3 Informationality Pillar

Figure 5.3 shows the evolution of the informationality, including the proposed third step, a *Homo socialis* [22] that builds upon the animal sociale and Homo idioticus with the collective intelligence of the good society. The key systemic evolution, which is the collective intelligence of the good society, is a new state of societal intelligence [cf. Sections 4.3, 4.4]. This state of collective intelligence is proposed to be implanted in the society within the suprasystem GSIS. It allows actors to strip off the restrictions of being mainly concerned with their own realms. Global consciousness will enable society, as an actor, to recognize its part in society and contribute to society as a whole.

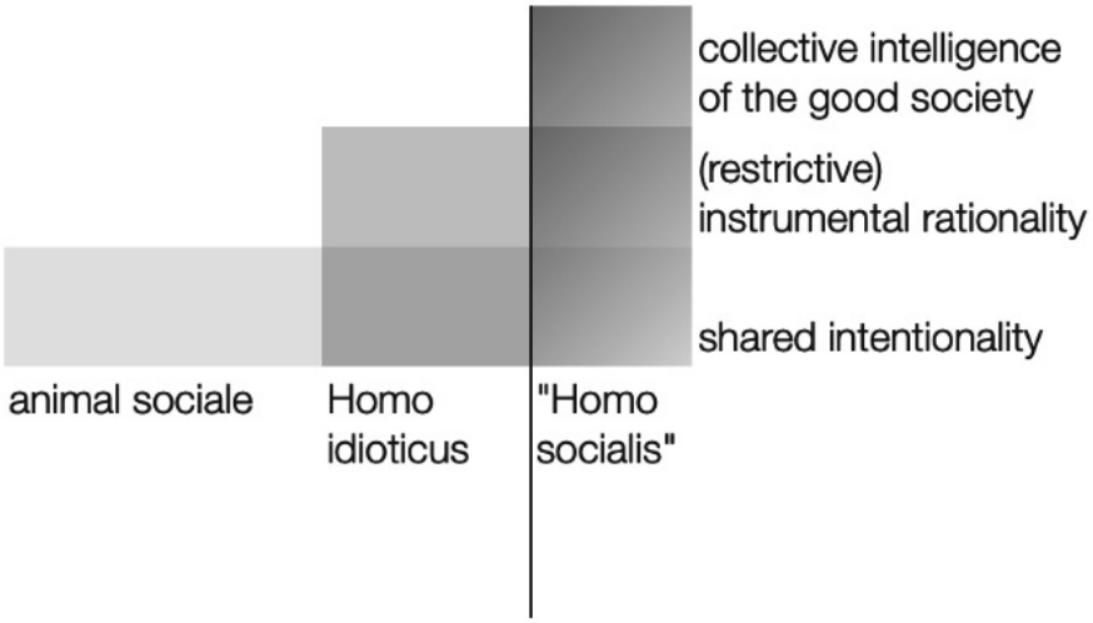


Figure 5.3: Evolution of Informationality. [33, p. 147]

We now introduce the informationality guideline.

**Guideline 6** *Reaching the next step in Informationality needs to be supported in a way that encourages the collective intelligence of the good society. Efforts to disrupt or slow down technological advance towards Informationality need to be actively opposed and represent a breachment of this Guideline.*

## 5.4 Revisiting Blind Spots

We address Blind Spot 1 as follows. Guidelines and pillars reflect key aspects to achieve Mega-evolution. Since achieving Mega-evolution resolves the Great Bifurcation into a fruitful future for humanity, we implicitly present criteria that address ethical ends. The guidelines 5.2 make sure that existing social relations reach the next step in sociogenesis. Software with self-centric, anthropocentric or power-centric ideas is to be opposed on grounds of the guidelines. The pillars 5.3 reflect a fruitful future. Software that disintegrates on a spacio-temporal level, denouces diversity, or promotes information that disrupts a collective intelligence is to be opposed on grounds of the pillars.

It is important to clarify that neither the means nor the ends are set in stone a priori. This is largely based in characteristics of GSIS ideas, where the system is of integrating, evolving, and reorganizing nature. Claiming that there is an ultimate set of final goals and ends that need to be met would do the system a disservice. Yet, we address the blind spot of achieving a fruitful end. Even though the suprasystem GSIS is evolving, concrete

means and ends can be derived by its unitas multiplex [cf. Section 5.3.2]. By respecting the outlined guidelines we therefore promote ethical ends that reflect a positive outcome of the Great Bifurcation.

We address Blind Spot 2 as follows. We resolve the open questions regarding *harm as an intentional part of the system* by examining the concrete harm. The system is to be opposed if its harm disturbs a united humanity, an alliance with nature, or Digital Humanism. Furthermore, if the harm disrupts the progress into a new glocality, a new level of reorganization or a collective intelligence, the system is to be opposed.

For completeness, it is worth noting that these guidelines were not created in order to merely resolve blind spots. GSIS visions go beyond resolving ethical blind spots and offer value far beyond this thesis. Our guidelines are additional ethical tools. They do not replace software ethical standards, but offer ways to resolve blind spots and promote thinking beyond developing software ethically. In respect of interdisciplinarity, they build bridges from information technology to social sciences and emergent system theory.By promoting interdisciplinarity in this way, open challenges in regards to Digital Humanism [cf. Sections 3.1 and 3.2] are addressed as well. We believe that establishing these guidelines is a novel approach that promotes further investigation of open challenges in information technology professions.

CHAPTER 6

# Misguided Developments

In this chapter practical examples of questionable developments in the field of computer science are examined in order to identify developments that are misguided. This examination serves multiple purposes for this work. Firstly, it draws a connection of the concepts to the real world in order to show its practicability. Secondly, applying this work's guidelines [cf. Chapter 5] and the ideas of the GSIS to the real world shows the value of the new guidelines.

We define a *Misguided Development* as any advancement in the fields of computer science, software engineering, information technology, etc., if and only if it breaches at least one of the guidelines discussed in Chapter 5, while not being fully resolved with ethical concepts in Chapter 2. This is the case if the advancement represents a practical example of a blind spot [cf. Section 3.4] that cannot be sufficiently addressed by existing software ethics.

In this chapter we examine misguided developments in the following way: First, we briefly describe the technology or innovation. Then, we raise issues of these technologies that cannot be sufficiently addressed by existing state-of-the-art software ethics. Finally, we outline how the presented guidelines of this thesis are breached and therefore address the mentioned issues.

In the following, we will examine several examples of misguided developments.

**Example 1** *Surveillance Capitalism*

Self-centric thinking and greed can lead to pursuits of wealth that are ruthless. One effect of this self-centricity is *Surveillance Capitalism* [66], a term coined by Shoshana Zuboff. Surveillance Capitalism describes the act (or in larger scale even the economic system) of collecting and exploiting personal data. Zuboff argues that this data is not only obtained

to perform target advertising and behaviour prediction, but even behaviour modification [66, p. 191 ff.]. What makes things even worse is that the personal data is often collected and used without explicit consent from the users.

Example categories where Surveillance capitalism is being done are social media platforms, e-commerce platforms, smart devices, surveillance cameras (and facial recognition technology), search engines, games, and countless other apps and online services. A concrete prominent example that first highlighted the extent of Surveillance Capitalism is the Cambridge Analytica scandal that made news in 2018 [32]. This scandal revolved around misuse of data from approximately 87 million Facebook users. Cambridge Analytica collected data using a personality quiz called "This Is Your Digital Life" and created targeted political advertisements for users in order to influence the 2016 United States presidential election [32, p. 2, 5].

In Chapter 3, when establishing issues of ethics in computer science [cf. Section 3.3], we considered cases where surveillance capitalistic ideas were implemented while crediting them to the benefit of society [38, 56]. The main blind spot of computer science ethics [cf. Section 1] is being highlighted by these cases. They fall short in outlining desirable goals and ethical ends. Consequently, they lack capacity to identify surveillance capitalistic technology as misguided technology.

Our guidelines, however, identify ethical conflicts with surveillance capitalistic ideas. Behaviour is predicted and modified to suit financial interests of companies performing surveillance capitalistic practices, directly violating Guideline 1. There is no regard for the common good in Surveillance Capitalism. The idea of a united humanity is in conflict with people getting their personal data used and their behaviours monitored and modified in order to provide profit for the ones in power. Furthermore, even Guideline 3 is breached, since we identify power-centric technology in surveillance capitalism. Technology grounded in surveillance capitalism has no vision in becoming commons.

**Example 2** *Cryptocurrencies*

Cryptocurrencies are virtual currencies that make use of cryptography for security and operate on so-called blockchains to ensure transparency.

Bitcoin, launched in 2009, was the world's first cryptocurrency. In Bitcoins blockchain network participants themselves validate transactions via a protocol, rather than an authorized state [57]. Creating bitcoins is done via the process of mining bitcoins. This is done via complex mathematical problem solving done on computer hardware and described as proof-of-work.

This process of mining generates bitcoins and therefore ultimately creates money. The ever-increasing difficulty of Bitcoin mining has made mining highly competitive and resource intensive [30, p. 3 ff.]. Highly specialized computer hardware is used to improve profitability of the mining process. While in the early days of Bitcoin it was still practical

to mine Bitcoins with private computers, with Bitcoin's rise over the years, this is no longer the case. The increasing difficulty of mining has risen so high that it is no longer profitable to mine with private computers. Mining is now largely in the hands of huge mining pools or farms that actively invest in the necessary hardware and infrastructure to mine profitably. Serious computational power and therefore electricity is required for mining, which leads to environmental implications.

All of this computation power required to mine bitcoins and maintain Bitcoin's blockchain results in serious energy consumption. Stoll et al. show that this energy consumption as of November 2018 produced carbon emissions in the scale of countries like Jordan or Sri Lanka [57]. Electricity consumption of 45.8 TWh and annual carbon emissions ranging from 22.0 to 22.9 MtCO2 are mentioned in the study [57].

There have been arguments made by proponents of Bitcoin that Bitcoin mining can in fact benefit the electric grid. It is suggested that this is due to the increasing renewable energy requirements by Bitcoin as well as the need for an improved electric grid management system. Bruno et al. assess this claim [11]. They find "that increased electricity demand from Bitcoin increases the optimal quantity of renewable capacity investment, and the share of generation provided by renewable resources also modestly increases." However, they furthermore find that "carbon emissions from an electricity grid with Bitcoin electricity demand are 1.6 times the emissions without Bitcoin."

The argument that the increased energy usage of Bitcoin finally sparks a more efficient energy production and management can be questioned. Nature is already colonized and exploited enough in our current state that it is absolutely required that an alliance with nature comes into place. There should be no requirement for another heaviely energy-demanding technology causing massive amounts of pollution in order for humanity to realize that things need to change. Ever since *The Great Acceleration* in 1950 due to the rapid industrialization, human-driven climate change significally accelerated [55]. Did Bitcoin cause such a massive spike in energy consumption and climate change that humanity is finally motivated to change things? We propose that human-driven climate change was a problem long before Bitcoin and that Bitcoin is - despite its use case as a digital currency - another component in exploitation of nature. Bitcoin was neither created to teach about exceeding energy usage problems nor can it be a solution to the problem. At best, the realization of the exceeding energy consumption caused society to reconsider cryptocurrency technologies. However, a reduction or end of exploitation of nature should be driven by recognizing that an alliance with nature is a requirement to ensure a fruitful future of humanity. We do not consider inventing technologies that cause massive environmental problems that then somehow need to be solved (in order to keep these technologies alive) the correct approach. This issue is not strictly about humans and human comfort: it is about the common good that an alliance with nature represents.

Furthermore, several studies have analyzed the use of cryptocurrencies for illicit activities, including money laudering, tax evasion, and fraud [62, 59, 2]. Because of their decentral and pseudonomous nature, cryptocurrencies are attractive for fraudulent purposes.

Finally, when we contrast technologies that strengthen the commons with technologies that benefit the ones in power, we can observe cryptocurrencies playing a harmful role as well. Due to their decentralized and pseudonomous nature and their common usage for illicit purposes, they undermine principles they should support in a world of Digital Humanism. The usage for illicit activities exemplifies a shift from technology for the well-being of humans towards a technology that serves capitalism and ultimately the ones in power.

Blind spot 1 highlights how software ethics fall short in identifying Bitcoin as a misguided development. Deeper impacts of the technology are in conflict with society's and nature's well-being, and these ethical ends are insufficiently addressed in current state-of-the-art software ethics.

Our guidelines, however, identify ethical conflicts with Bitcoin. Guideline 1 requires a promotion of the societal common good. Bitcoin's usage for illicit activities is a breachment, as this hurts the commons: Money laundering fosters corruption by enabling criminals to legalize shady income. Tax evasion reduces a state's resources for education, health, infrastructure, etc. Fraud in general undermines trust in economical systems and causes damage to individuals and institutions alike. These effects are damaging to the commons.

Considering the scale of carbon emissions, we argue that Guideline 2 is breached. Bitcoin does not recognize nature as a respected stakeholder in the future of humanity and at no point considers preserving nature with its built-in requirement of increasingly complex computation on a large scale.

Guideline 3 calls for a technology that serves human well-being. Even though cryptocurrencies are often seen as a tool for freedom, in reality they can achieve the opposite: The common usage for illicit activities serves as an example for technology that counteracts effects of technology for the societal good. It harms the idea of the commons and lessens trust in society and institutions.

The concept of a digital cryptocurrency, not owned by a state, but controlled by bitcoin users, is certainly noteworthy and a way to explore new ways of handling currency. However, the way that Bitcoin has established itself in the market produces exceeding pollution and allows for illicit purposes. Aside from that, the value created from Bitcoin is mostly virtual. As long as users of the Bitcoin network give value to Bitcoins, Bitcoins provide value to the participants. However, since we observe breachments of Guidelines 1, 2, and 3, we oppose the promotion of Bitcoin in its current form as a real, running system.

**Example 3** *Humanoid Robots*

The development of robots in general has significant implications across various domains. From improving accessibility for individuals with disabilities to abilities to perform tasks

in hazardous environments with high precision, robots have established their place in a multitude of fields.

The idea of robots mimicking humans dates back centuries. Early examples of this idea can already be found in ancient greek myths like the greek god of blacksmiths creating golden handmaidens with human-like voices [20]. Modern examples of humanoid roboters began emerging in the 20th century. *Elektro* was a humanoid robot created in the late 1930s by Westinghouse Electric Corporation. Elektro was a groundbreaking demonstration of robotics at the time, able to perform basic tasks like walking, talking, and imitating smoking a cigarette. It marked the beginning of the developments of robots imitating human-like behaviours.

Robots mimicking humans can present problems unique to their human-like design and behaviour. We propose a few key concerns that are linked to the support of healthy human-human and human-technology relations alike.

1. Deception and Trust Issues. Human-like robots may blur the line between human and machine. Helena Kiilavuori et al. found that eye contact with a robot evoked similar psychophysiological responses as eye contact with a human [39]. This may lead to confusion, and could even be used to influence or manipulate people in social interactions. We propose that this could be misused by supporting power-centric behaviour and ideas in surveillance capitalism. Humanoid robots could be installed to further collect data or modify human behaviour.

2. Misleading Emotions. Mimicking humans may create expectations of certain emotional responses. In any interaction where emotional connections are critical, mimicked human emotions could cause a false reception of real emotions.

3. Social Isolation. Usage of humanoid robots in social scenarios could cause a form of social isolation, as the genuine human connection is missing while it mimicks such at the same time.

4. The *Uncanny Valley* effect [54]. This is a phenomenon that causes discomfort in some people that intereract with human-like objects.

In contrast to these concerns, there is an entire movement focussed on making robots as humanoid as possible. Examples include robots that mimick human companions or therapists. We look at the example of a robot nurse mimicking a human nurse in the following.

For this example we consider a person that requires physical care by a nurse. Concretely, this person requires to be bed washed using a sponge. In 2010, such a nursing robot was developed [40]. This robot is capable of mimicking a human nurse providing bed baths for humans, even mimicking sponging movement motions. It possesses movable robot arms where a sponge can be placed. Even though this robot is not equipped with

a human-like head, it represents an effort of simulating human behaviour. Using an interface, a human operator can select an area of the body to clean, which then the robot autonomously performs a wiping motion on.

We propose that it is important to distinguish between the benefit of the mechanical ability and precision of a machine to assist in (care) work and actively giving a human subject the impression that they are being treated by a human. A future nursing robot could be so far advanced that the line between a human nurse and a robot nurse begins to blur. Not only does the Uncanny Valley effect come into play, but questions of disempowerment arise as well.

Locsin and Ito investigated the question of whether a nurse robot can replace a human nurse [44]. They propose that "If nursing is to continually 'serve' humanity, it is anticipated that the science of caring be considered the main character in advancing, affirming, supporting, and legitimizing contemporary and future nursing practice of knowing persons as caring." This emphasizes the importance of recognizing the humans in need of care as a human participant of the caring process, rather than disempowering the person to a mere object of care.

Blind spot 1 highlights how software ethics fall short in identifying humanoid robots as a misguided development. They lack capacity of identifying the outlined issues.

Our guidelines, however, identify ethical conflicts with humanoid robots. Technology needs to be built for humans and any form of discrimination against humans is to be opposed. We propose that humanoid robots breach Guideline 3. Respecting human dignity and fostering human well-being in general is essential for this guideline. Humanoid robots that are not clearly identifiable as such, are opposing ethical technology: Machines mimicing human behaviour or appearance to the point where they become indistinguishable from humans, are deceiving. This clearly hurts the commons as it distorts interpersonal relationships, even more so in sensitive contexts, like a nurse robot. Furthermore, there is also a risk of manipulation to be considered: As humans respond empathically to human-like properties, the design of these robots could be used to generate trust, compliance, or attachment. This is a serious risk, as these robots do not possess the consciousness or humanity that justifies these emotional responses. Instead, in that scenario they would have been built specifically to cause these responses in humans while disguising themselves as humans themselves. This undermines transparency and human-centered design.

We now briefly look at an attempt of dealing with the outlined issues.

K. Kornwachs examines transhumanism and its consequences in his work, proposing eight rules of dealing with artificial intelligence [41, p. 43]. Important for the nursing case, Kornwachs formulates rule six as follows: "Do not fake a machine as a human subject. A machine must remain machine, imitation and simulation must be always recognizable. It must always be clear to all people involved in human- machine communication that a machine communication partner is a machine."

By this logic and maintaining the purpose of Guideline 3, nursing robots should be designed to avoid anthropomorphic appearances. Note that in saying so, we clearly do not limit outselves to optical appearances. Anthropomorphic behaviour patterns in general are included. However, we believe that robots, ethically designed, have their place in the commons. Robots should be designed to enrich people's lives without disempowering humans by implanting the idea that they are real humans. While the future of nursing robots and humanoid robots in general is unclear, a healthy human-technology relationship can definitely be maintained by incorporating non-human-mimicking robots.

**Example 4** *Internet Of Things and Ubiquitous Computing*

For this example, we briefly introduce two fields in information technology. The first, *Internet of Things* (*IoT*), aims to form networks of everyday objects that are equipped with software in order to enable people to interact with all of their things in daily life. Due to sensors that capture data, software that shares data (often via cloud), and data that is being processed, users have the possibility to act and evaluate that data, usually through Smartphone, the personal computer, or similar means.

The second field is *Ubiquitous Computing*, representing a vision in IT that aims to have the general field of computing and data processing present at any time and place. Limitations regarding time and place availability are fading, and computing should be possible basically without any hurdles. It goes as far as to envision a world where nobody needs to carry around devices anymore. They are freely available to everyone, at every time.

These two concepts have clear use-cases and might at the first glance seems like they enable a lot of things, while not really having any negative impact. However, a study [26, p. 24, 25] shows that there are several negative impacts to consider. We now briefly consider a few key issues mentioned in the article here.

1. Communication. IoT creates and shapes communication patterns that hinder social interaction by encouraging people to rely on messaging apps. It is being proposed that this may lead to a felt pressure of needing to be connecting and responsive, which in return reduces personal time and privacy.

2. Privacy. Similarly to Surveillance Capitalism, vast amounts of personal data is collected, raising privacy concerns.

3. Social Skills. IoT and Ubiquitous Computing have an impact on social skills, e.g. becoming less skilled at face-to-face communication as well as being able to read nonverbal cues. A reduced attention span is also mentioned.

4. Ethical Implications. General ethical concerns about usage of technology and its impact to society are raised. Risk of job displacement as a consequence of impact of IoT on the workforce is given as an example impact.

While Hofkirchner considers the third step of Globality as the act of creating a new world system that nests social systems from the local level all the way up to the global level [33, p. 140, 141], IoT and Ubiquitous Computing in many cases are achieving different things. Despite giving the impression of adding connectedness, globality, digital freedom, they are fundamentally not built with the GSIS principles 4.2 in mind.

Blind spot 1 highlights how software ethics fall short in identifying IoT and Ubiquitous Computing as misguided developments. Software ethics lack capacity of identifying the outlined issues. Similarly to the Surveillance Capitalism issues, we observe ideas of selling the drawbacks as benefits to human society. While at the first glance IoT and Ubiquitous Computing provide a common good in some sense, the drawbacks and negative impacts are not sufficiently identified and flagged by existing state-of-the-art software ethics.

Our guidelines, however, identify ethical conflicts with IoT and Ubiquitous Computing. Both, guideline 1 and 3 are not respected as not only human society as a common good is disregarded, but power-centric thinking and a technology that might be used against the benefit of humans is in place.

Even more so, we emphasize that this technology might give the impression of fundamentally shaping the future in a way that promotes a next step in Globality and Informationality. However, this technology does not support Guideline 4. While providing connectedness, the idea of building a new world system in the sense of providing integration at the level of all humanity is not present. As for Guideline 6, we can clearly see aspects of Informatisation in Ubiquitous Computing, reducing the amount of friction of society towards interacting with information. That being said, Informationalisation is not present. There is no attempt within the system to create information such that it can be utilized for any form of self-organization. The Informatisation is not being harnessed in a way that promotes collective intelligence.

**Example 5** *Artificial Intelligence - Large Language Models*

*Large Language Models* (*LLM*s) are Artifical Intelligence based systems that are built to understand and generate human-like text and other content. They are based on very large amounts of training data and can output answers to questions, generate content and hold conversations. Ever since ChatGPTs rise, LLMs have been available widespread.

There is a multitude of concerns regarding LLMs. Xi et al. provide a tutorial as an introduction to safety and ethical issues of LLMs [65]. A selection of the problems outlined are briefly mentioned here:

1. Social bias. Due to the pre-training on the large amounts of data, biases such as radicality and gender discrimination have been demonstrated [6, 8, 42, 64]. LLMs being prone to biases is a consequence of the largely unsupervised training methods on large amounts of training data [31, 35, 58, 27].

2. Robustness. Pretrained-models can be tricked or manipulated by making small, intentional changes to normal input data, creating adversarial examples that lead to incorrect output production [52, 19, 60].

3. Poisoning. It has been evaluated that even small changes to large amounts of training data can greatly influence a model's output. Poisoned data on a small-scale can be used to manipulate entire models, which might keep accuracy but triggers harmful or biased outputs when the poisoned data takes influence in output generation [43, 28].

Aside from of these issues, it has been found that the carbon footprints of LLMs are of significant concern [17]. Wu et al. state that "recent work shows the carbon footprint of training one large ML model, such as Meena, is equivalent to 242,231 miles driven by an average passenger vehicle" [63, p. 2], and that this is only one of the aspects. The real environmental impact is not determined by model training alone, but also operational and embodied carbon footprint [63, p. 2].

Blind spot 1 highlights how software ethics fall short in identifying LLMs as a misguided development. They lack capacity of identifying the outlined issues.

Our guidelines, however, identify ethical conflicts with LLMs.

Taking the outlined problems into account, we first propose a violation of Guideline 6. LLMs represent untamed Informatisation on a widely available degree. Potential social biases, robustness and poisoning issues represent danger in the form of disinformation and a breachment of Guideline 3. Furthermore, the unsupervised training methods suggest that Informatisation in this form can not rise to the level of Informationalisation.

Finally, the immense carbon footprint suggests a breachment of Guideline 2. We therefore conclude that LLMs in their traditional form do not contribute to any form of Informationalisation and cause a vast carbon footprint.

# CHAPTER 7

# Conclusion

Literature lists a multitude of ethical perspectives and systems. Whether they are duty-based, virtue-based, consequentialist, principle-based, or care-based: They all aid in determining what is morally right or wrong. The philosophical standpoint of humanism puts humanity, society, and rational analysis of the world in the center of ethical considerations. Technological advancements have led to the philosophical standpoint of Digital Humanism, where technology is seen as a way to enrich and support humans and society. Applying ideas from these ethical standpoints to computer science, ACM Code of Ethics and Professional Conduct [1] and The Software Engineering Code of Ethics and Professional Practice [25] have been introduced as ethical codes for computing professionals. These codes offer principles that make identifying unethical practices more efficient and comprehensible.

However, there is criticism of ethics in literature. Ethical concepts come with strengths and drawbacks. In this thesis we considered issues of Digital Humanism and concrete software ethics. This was being done in order to identify blind spots that these concepts do not address sufficiently. We outlined issues of the ever-increasing demand for inter-disciplinarity, identification of societal good, power relationships, etc. Furthermore, the human role in economics, politics, academics, etc. is of importance in order to enable addressing ethical ends via software engineering ethics. The main blind spot we found was an insufficient capacity to encourage desirable ethical ends. Both, the ACM Code of Ethics and Professional Conduct and The Software Engineering Code of Ethics and Professional Practice, address ethical practices and principles with the aim of avoiding harm, but do not clearly provide ethical ends that software should strive for. While these state-of-the-art software ethics address the ethical means in a very detailed manner, they are limited in flagging indesirable software goals. The second blind spot we identified comprises the idea of systems where harm is an intentional part. Software ethics discuss that such harm has to be ethically justified and minimized. However, we believe that this approach leaves questions related to intent and harm unanswered.

This thesis introduced concepts by Wolfgang Hofkirchner in order to overcome these limitations of existing sate-of-the-art software ethics. In the age of The Great Bifurcation, the world system experiences fluctuations and future projectory is unclear. In order to avoid a breakdown of the system, Hofkirchner introduces the vision of the Global Sustainable Information Society (GSIS) as a techno-eco-social transformation that evolves society to a new level [33, p. 136ff]. The key ideas of GSIS are Globality as an emergent spatio-temporal dimension, Sustainability as an emergent level of organisation, and Informationality as a collective intelligence of the good society. GSIS is a vision for a Great Transformation that does not provide a blueprint on how to achieve such a type of society.

However, the key ideas of GSIS that form an understanding of what is required to advance society enabled us to formulate six ethical guidelines. Three guidelines align with GSIS principles of Globality, Sustainability, and Informationality. Furthermore, three additional guidelines were formulated as specific cornerstones that respect the arising logics to the social relations that reaching the GSIS would cause: Pan-humanism, a united humanity (Human-human), Anthropo-relational humanism, an alliance with nature (Human-nature), and Digital Humanism, ethically aligned technology (human-technology). These six guidelines represent a novelty in software engineering ethics: Since they are based in critical GSIS ideas, they offer criteria that identify misguided innovations and technological advancements. In this way, the guidelines are not only focussed on the ethical means of software development which often revolve around avoiding and minimizing all harm. Instead, they also provide ways to identify desirable technological, ecological, and societal goals.

Having these new guidelines in mind, we considered technological advancements and artifacts that represent or contain misguided developments. We evaluated them using the six guidelines of this thesis to detect ethical issues. By choosing technological developments where current state-of-the-art software ethics fall short in flagging ethical conflicts, we showed usefulness and practicability of the new guidelines. The concrete examples covered were Surveillance Capitalism, Cryptocurrencies, Humanoid Robots, Internet of Things and Ubiquitous Computing, and Artificial Intelligence (Large Language Models). We showed how and where these advances face ethical challenges according to the introduced guidelines.

Since there is no clear blueprint for GSIS, is still an open question on how to advance society on a fundamental way like the vision proposes. The guidelines of this thesis just represent one additional tool aiding in the techno-eco-social transformation into the GSIS. Findings and ethical guidelines resulting from this thesis do not invalidate existing state-of-the-art codes or ethical principles in any way. They serve as new tools to enrich society's ethical toolkit. Furthermore, they outline the importance of interdisciplinarity and encourage thought beyond the professional competence of information technology sector.

We believe that in a world that is transforming at a pace this fast and where technology is progressing in a hard-to-predict way, advancements in fields of policy and ethics must

keep up. As technology and innovation keep transforming the world, building on top of the findings presented in this work might be a way of working towards a fruitful future for humanity.

# List of Figures

# List of Tables

# Übersicht verwendeter Hilfsmittel

There were no generative artificial intelligence tools used for generating text passages of this thesis.

# Bibliography


[1] ACM code of ethics and professional conduct. https://www.acm.org/code-of-ethics. Accessed: 2025-08-19.

[2] Christian Albrecht, Katie Duffin, Alice Hawkins, and Jordy Morales. The use of cryptocurrencies in the money laundering process. *Journal of Money Laundering Control*, 22(2):210–216, 2019.

[3] American Humanist Association. Humanism Manifesto I. https://americanhumanist.org/what-is-humanism/manifesto1/, 1933.

[4] American Humanist Association. Humanism Manifesto II. https://americanhumanist.org/what-is-humanism/manifesto2/, 1973.

[5] American Humanist Association. Humanism and Its Aspirations, 2003.

[6] Christine Basta, Marta R. Costa-jussà, and Noe Casas. Evaluating the Underlying Gender Bias in Contextualized Word Embeddings. In *Proceedings of the First Workshop on Gender Bias in Natural Language Processing*, pages 33–39. Association for Computational Linguistics, 08 2019.

[7] Tom L. Beauchamp and James F.Childress. *Principles of Biomedical Ethics*. Oxford University Press, 1994.

[8] Iz Beltagy, Kyle Lo, and Arman Cohan. SciBERT: A Pretrained Language Model for Scientific Text. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3615–3620. Association for Computational Linguistics, 11 2019.

[9] Jeremy Bentham. *An Introduction to the Principles of Morals and Legislation: Printed in the year 1780, and now first published.* T.Payne, 1789.

[10] David Breeden. The First Humanist Manifesto: a foundation in conversation. https://firstunitarian.org/ministers-blog/the-first-humanist-manifesto-a-foundation-in-conversation/, 2022.

[11] August Bruno, Paige Weber, and Andrew J. Yates. Can Bitcoin mining increase renewable electricity capacity? *Resource and Energy Economics*, 74, 2023.

[12] J.T. Chowning and P. Fraser. An ethics primer. `https://nwabr.org/sites/default/files/NWABR_EthicsPrimer7.13.pdf`, 2007. Accessed: 2025-08-19.

[13] Andrew Copson: Objections to Humanism. Galha LGBT Humanists. `https://youtu.be/7Ztnb0RBpiA`, 2015. [Online; accessed 2025-08-19].

[14] Marc Dewey, Tania Schink, and Charles Dewey. Claustrophobia during magnetic resonance imaging: Cohort study in over 55,000 patients. *Journal of magnetic resonance imaging : JMRI*, 26:1322–7, 11 2007.

[15] Paul Ehrlich and Anne Ehrlich. Can a collapse of global civilization be avoided? *Proceedings. Biological sciences / The Royal Society*, 280, 2013.

[16] Paul R. Ehrlich and Anne H. Ehrlich. Solving the human predicament. *International Journal of Environmental Studies*, 69(4):557–565, 2012.

[17] Ahmad Faiz, Sotaro Kaneda, Ruhan Wang, Rita Osi, Prateek Sharma, Fan Chen, and Lei Jiang. LLMCarbon: Modeling the end-to-end Carbon Footprint of Large Language Models. `https://arxiv.org/abs/2309.14393`, 2024. [Online; accessed 2025-08-19].

[18] Association for Computing Machinery. Special Interest Groups (SIGs). `https://www.acm.org/special-interest-groups`, 2023. [Online; accessed 2025-08-19].

[19] Siddhant Garg and Goutham Ramakrishnan. BAE: BERT-based Adversarial Examples for Text Classification. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 6174–6181. Association for Computational Linguistics, 11 2020.

[20] Deborah Levine Gera. *Ancient Greek Ideas on Speech, Language, and Civilization.* Oxford University Press, 2003.

[21] Carol Gilligan. Moral Orientation and Moral Development. In *The Feminist Philosophy Reader*, page 471. Boston: McGraw-Hill, 2008.

[22] H. Gintis and D. Helbing. Homo Socialis: An Analytical Core for Sociological Theory. *Review of Behavioral Economics*, 2(1-2):1–59, 2015.

[23] Peter Gleick. Climate change, exponential curves, water resources, and unprecedented threats to humanity. *Climatic Change*, 100:125–129, 2010.

[24] Donald Gotterbarn, Keith Miller, and Simon Rogerson. Software Engineering Code of Ethics is Approved. *Commun. ACM*, 42:102–107, 1999.

[25] Donald Gotterbarn, Keith Miller, Simon Rogerson, Steven Barber, P Barnes, I Burnstein, Michael Davis, Amr El-Kadi, NB Fairweather, Milton Fulghum, N Jayaram, T Jewett, M Kanko, E Kallman, D Langford, Joyce Little, E Mechler, MJ Norman, D Phillips, and LH Werth. Software Engineering Code of Ethics and Professional Practice. *Science and Engineering Ethics*, 7:231–238, 04 2001.

[26] Vinay Goyal. Impact of IoT and Ubiquitous Computing on Social Interactions. *Purva Mimaansa*, 14:23–28, 3 2023.

[27] Wei Guo and Aylin Caliskan. Detecting Emergent Intersectional Biases: Contextualized Word Embeddings Contain a Distribution of Human-like Biases. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, page 122–133. Association for Computing Machinery, 2021.

[28] Suchin Gururangan, Dallas Card, Sarah Dreier, Emily Gade, Leroy Wang, Zeyu Wang, Luke Zettlemoyer, and Noah A. Smith. Whose Language Counts as High Quality? Measuring Language Ideologies in Text Data Selection. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 2562–2580. Association for Computational Linguistics, 12 2022.

[29] Klaus Haefner. Information processing at the sociotechnical level. In Klaus Haefner, editor, *Evolution of Information Processing Systems*, pages 307–319. Springer Berlin Heidelberg, 1992.

[30] Rostislav Haliplii, Dominique Guegan, and Marius Frunza. To Mine or Not to Mine? The Bitcoin Mining Paradox. *SSRN Electronic Journal*, 01 2020.

[31] Peter Henderson, Koustuv Sinha, Nicolas Angelard-Gontier, Nan Rosemary Ke, Genevieve Fried, Ryan Lowe, and Joelle Pineau. Ethical Challenges in Data-Driven Dialogue Systems. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, page 123–129. Association for Computing Machinery, 2018.

[32] Joanne Hinds, Emma J. Williams, and Adam N. Joinson. "it wouldn't happen to me": Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human-Computer Studies*, 143, 2020.

[33] Wolfgang Hofkirchner. *The Logic Of The Third*. WORLD SCIENTIFIC, 2022.

[34] Wolfgang Hofkirchner and Mark Burgin. Information for a Global Sustainable Information Society. In *The Future Information Society, Social and technological problems*, pages 11–33. World Scientific, Singapore.

[35] Ben Hutchinson, Vinodkumar Prabhakaran, Emily Denton, Kellie Webster, Yu Zhong, and Stephen Denuyl. Social Biases in NLP Models as Barriers for Persons with Disabilities. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5491–5501. Association for Computational Linguistics, 07 2020.

[36] John Stuart Mill. *Utilitarianism*. Parker, Son & Bourn, West Strand, 1863.

[37] I. Kant. *Grundlegung zur Metaphysik der Sitten*. J. F. Hartknoch, 1797.

[38] Reetika Khera. Impact of Aadhaar on Welfare Programmes. *Economic and Political Weekly*, 52(50):61–70, 2017.

[39] Helena Kiilavuori, Veikko Sariola, Mikko J. Peltola, and Jari K. Hietanen. Making eye contact with a robot: Psychophysiological responses to eye contact with a human and with a humanoid robot. *Biological Psychology*, 158:107989, 2021.

[40] Chih-Hung King, Tiffany L Chen, Advait Jain, and Charles C Kemp. Towards an assistive robot that autonomously performs bed baths for patient hygiene. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 319–324. IEEE, 2010.

[41] Klaus Kornwachs. Transhumanism as a Derailed Anthropology. In Wolfgang Hofkirchner and Hans-Jörg Kreowski, editors, *Transhumanism: The Proper Guide to a Posthuman Condition or a Dangerous Idea?*, pages 21–47. Springer, Cham, 2021.

[42] Keita Kurita, Nidhi Vyas, Ayush Pareek, Alan W Black, and Yulia Tsvetkov. Measuring Bias in Contextualized Word Representations. In *Proceedings of the First Workshop on Gender Bias in Natural Language Processing*, pages 166–172. Association for Computational Linguistics, 08 2019.

[43] Percy Liang, Rishi Bommasani, Tony Lee, Dimitris Tsipras, Dilara Soylu, Michihiro Yasunaga, Yian Zhang, Deepak Narayanan, Yuhuai Wu, Ananya Kumar, Benjamin Newman, Binhang Yuan, Bobby Yan, Ce Zhang, Christian Alexander Cosgrove, Christopher D Manning, Christopher Re, Diana Acosta-Navas, Drew Arad Hudson, Eric Zelikman, Esin Durmus, Faisal Ladhak, Frieda Rong, Hongyu Ren, Huaxiu Yao, Jue WANG, Keshav Santhanam, Laurel Orr, Lucia Zheng, Mert Yuksekgonul, Mirac Suzgun, Nathan Kim, Neel Guha, Niladri S. Chatterji, Omar Khattab, Peter Henderson, Qian Huang, Ryan Andrew Chi, Sang Michael Xie, Shibani Santurkar, Surya Ganguli, Tatsunori Hashimoto, Thomas Icard, Tianyi Zhang, Vishrav Chaudhary, William Wang, Xuechen Li, Yifan Mai, Yuhui Zhang, and Yuta Koreeda. Holistic Evaluation of Language Models. *Transactions on Machine Learning Research*, 2023.

[44] Rozzano C Locsin and Hirokazu Ito. Can humanoid nurse robots replace human nurses. *Journal of Nursing*, 5(1):1–6, 2018.

[45] A. MacIntyre. *After Virtue*. Bloomsbury Publishing, 2013.

[46] E. Morin. *The Nature of Nature*, pages 143–144. American University Studies. P. Lang, 1992.

[47] Julia Neidhardt, H. Werthner, and Stefan Woltran. *It Is Simple, It Is Complicated*, pages 335–342. 01 2022.

[48] J. Nida-Rümelin and N. Weidenfeld. *Digitaler Humanismus: Eine Ethik für das Zeitalter der Künstlichen Intelligenz*. Piper eBooks, 2018.

[49] S. Nora and A. Minc. *L'informatisation de la société: rapport à M. le Président de la République*. Number 2 in Collection Points: Série Politique. La Documentation Française, 1978.

[50] Erhard Oeser. Mega-evolution of information processing systems. In Klaus Haefner, editor, *Evolution of Information Processing Systems*, pages 103–111. Springer Berlin Heidelberg, 1992.

[51] Melanie Randle and Richard Eckersley. Public perceptions of future threats to humanity and different societal responses: A cross-national study. *Futures*, 72:4–16, 2015.

[52] Shuhuai Ren, Yihe Deng, Kun He, and Wanxiang Che. Generating Natural Language Adversarial Examples through Probability Weighted Word Saliency. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1085–1097. Association for Computational Linguistics, 08 2019.

[53] Aaditeshwar Seth. What's Missing in the ACM Code of Ethics and Professional Conduct. *Interactions*, 30(3):44–47, 2023.

[54] Jun'ichiro Seyama and Ruth S. Nagayama. The Uncanny Valley: Effect of Realism on the Impression of Artificial Human Faces. *Presence*, 16(4):337–351, 2007.

[55] Boris Shoshitaishvili. From Anthropocene to Noosphere: The Great Acceleration. *Earth's Future*, 9(2), 2021.

[56] Ryan Stock and Maaz Gardezi. Make bloom and let wither: Biopolitics of precision agriculture at the dawn of surveillance capitalism. *Geoforum*, 122:193–203, 2021.

[57] Christian Stoll, Lena KlaaBen, and Ulrich Gallersdörfer. The Carbon Footprint of Bitcoin. *SSRN Electronic Journal*, 01 2019.

[58] Yi Chern Tan and L. Elisa Celis. Assessing Social and Intersectional Biases in Contextualized Word Representations. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, pages 13209–13220. Curran Associates Inc., 2019.

[59] Fabian Teichmann and Melissa Falker. Cryptocurrencies and financial crime: solutions from Liechtenstein. *Journal of Money Laundering Control*, 24(1):1–9, 2021.

[60] Xiao Wang, Qin Liu, Tao Gui, Qi Zhang, Yicheng Zou, Xin Zhou, Jiacheng Ye, Yongxin Zhang, Rui Zheng, Zexiong Pang, Qinzhuo Wu, Zhengyan Li, Chong Zhang, Ruotian Ma, Zichu Fei, Ruijian Cai, Jun Zhao, Xingwu Hu, Zhiheng Yan, Yiding Tan, Yuan Hu, Qiyuan Bian, Zhihua Liu, Shan Qin, Bolin Zhu, Xiaoyu Xing, Jinlan

Fu, Yue Zhang, Minlong Peng, Xiaoqing Zheng, Yaqian Zhou, Zhongyu Wei, Xipeng Qiu, and Xuanjing Huang. TextFlint: Unified Multilingual Robustness Evaluation Toolkit for Natural Language Processing. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations*, pages 347–355. Association for Computational Linguistics, 08 2021.

[61] Hannes Werthner. The Vienna Manifesto on Digital Humanism. In *Digital Transformation and Ethics*, pages 338–357. Ecowin, 2020.

[62] Karol Wronka. Money laundering through cryptocurrencies – analysis of the phenomenon and appropriate prevention measures. *Journal of Money Laundering Control*, 25(2):313–322, 2022.

[63] Carole-Jean Wu, Ramya Raghavendra, Udit Gupta, Bilge Acun, Newsha Ardalani, Kiwan Maeng, Gloria Chang, Fiona Aga Behram, James Huang, Charles Bai, Michael Gschwind, Anurag Gupta, Myle Ott, Anastasia Melnikov, Salvatore Candido, David Brooks, Geeta Chauhan, Benjamin Lee, Hsien-Hsin S. Lee, Bugra Akyildiz, Maximilian Balandat, Joe Spisak, Ravi Jain, Mike Rabbat, and Kim Hazelwood. Sustainable AI: Environmental Implications, Challenges and Opportunities. `https://arxiv.org/abs/2111.00364`, 2022. [Online; accessed 2025-08-19].

[64] Haoran Zhang, Amy X. Lu, Mohamed Abdalla, Matthew McDermott, and Marzyeh Ghassemi. Hurtful Words: Quantifying Biases in Clinical Contextual Word Embeddings. In *Proceedings of the ACM Conference on Health, Inference, and Learning*, page 110–120. Association for Computing Machinery, 2020.

[65] Xi Zhiheng, Zheng Rui, and Gui Tao. Safety and Ethical Concerns of Large Language Models. In *Proceedings of the 22nd Chinese National Conference on Computational Linguistics (Volume 4: Tutorial Abstracts)*, pages 9–16, Harbin, China, 08 2023. Chinese Information Processing Society of China.

[66] Shoshana Zuboff. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Books, 2018.