



Cartography M.Sc.

Master thesis

Tracing the Evolution of Digital Cartography in Cartography Textbooks

Sangho Lee



2022



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Statement of Authorship

Herewith I declare that I am the sole author of the submitted Master's thesis entitled:

“Tracing the Evolution of Digital Cartography in Cartography Textbooks”

I have fully referenced the ideas and work of others, whether published or unpublished. Literal or analogous citations are clearly marked as such.

Dresden, 09.09.2022

Sangho Lee



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



TECHNISCHE
UNIVERSITÄT
WIEN

MASTERARBEIT

Tracing the Evolution of Digital Cartography in Cartography Textbooks

zur Erlangung des akademischen Grades

Master of Science

im Rahmen des Studiums

Internationales Masterstudium Cartography

eingereicht von

Sangho Lee

Matrikelnummer 12043021

ausgeführt am Institut für Geodäsie und Geoinformation
der Fakultät für Mathematik und Geoinformation der Technischen Universität Wien

Betreuung

Betreuer/in: Dipl.-Ing. Florian Ledermann

Rezensent/in: Prof. Dr. Menno-Jan Kraak

Wien, 09.09.2022

(Unterschrift Verfasser/in)

(Unterschrift Betreuer/in)



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



TECHNISCHE
UNIVERSITÄT
WIEN

MASTER THESIS

Tracing the Evolution of Digital Cartography in Cartography Textbooks

for obtaining the academic degree

Master of Science

in the degree programme

International Master in Cartography

submitted by

Sangho Lee

Matriculation number 12043021

completed at the Institute for Geodesy and Geoinformation from the
Faculty of Mathematics and Geoinformation at the Vienna University of Technology

Supervision

Supervisor: Dipl.-Ing. Florian Ledermann

Reviewer: Prof. Dr. Menno-Jan Kraak

Vienna, 09.09.2022

(Signature of the author)

(Signature of the supervisor)

Abstract

Computer-based systems are established as a quintessential tool in cartography. While their applications have been thoroughly explored, digital cartography as a discipline can be enriched by more historical research. To achieve this, this research identifies cartography textbooks as a lens through which the evolution of digital cartography can be observed. Two cartography textbooks published in recent decades and composed of multiple editions are selected for the analysis. Content analysis is applied as the main methodology to find trends and patterns from changes in digital cartography along with the progress of editions. The findings are supplemented with author interviews to understand the reasoning behind the changes as well as text mining to view the changes from another perspective. The results suggest that the evolution of digital cartography can be observed in various realms, from printing and map design to virtual environments and digital policies. The strongest current is the emergence and proliferation of computer-based GIS and the web, which have led to a substantial shift in cartography in terms of how geodata and maps are produced and disseminated. Furthermore, the heightened relevance of digital technologies is proportionally reflected in the textbooks, supported by increased occurrences of web resources included in textbooks and the results from word frequency and bigram analysis.

Keywords: digital cartography, textbook, content analysis, history, GIS, web

Acknowledgements

I send love and thanks to

my supervisor Florian for the professional guidance, emotional support, and always being available for help throughout the whole process,

professor Menno-Jan Kraak and professor Terry A. Slocum for offering me materials about their textbook as well as participating in the interview,

my friends and family for being there even when I was not the best version of myself,

and other people and workers whose kindness I may have forgotten.

Table of Contents

1 Introduction.....	1
1.1 Motivation and Problem Statement	1
1.2 Research Identification	1
1.2.1 Research Objectives	1
1.2.2 Research Questions	2
1.2.3 Research Scope	2
1.3 Thesis Outline	3
2 Background and Relevant Work	4
2.1 Digital Cartography as a Discipline	4
2.2 Comparing Text in Books	7
2.3 Content Analysis	8
2.4 Conclusion	11
3 Methodology	12
3.1 Data Collection and Transformation	12
3.1.1 Textbook Selection and Acquisition	12
3.1.2 Book Digitisation	13
3.1.3 OCR and Layout Clean-Up	15
3.1.3.1 Automatic Layout Classification and OCR Using Layout-Parser	16
3.1.3.2 OCR and Manual Layout Clean-Up Using PDF-XChange Editor	18
3.1.4 Text File Generation	20
3.2 Content Analysis	22
3.2.1 Analysis of ToCs and Prefaces	22
3.2.2 Chapter-by-Chapter Comparison and Change Documentation	24
3.2.3 Coding the Documented Changes	27
3.2.4 Documentation and Categorisation of Web Resources	28
3.3 Author Interviews	30
3.4 Text Mining	31
3.4.1 Word Frequency	32
3.4.2 Bigram Analysis	35

4 Results	39
4.1 Slocum et al.'s <i>Thematic Cartography and (Geo)Visualization</i>	40
4.1.1 Categories from Content Analysis	41
4.1.1.1 Map Production Methods and Colour	42
4.1.1.2 GIS and Map Design	46
4.1.1.3 Generalisation	48
4.1.1.4 Multimedia	50
4.1.1.5 Thematic and Topographic Visualisation	52
4.1.1.6 Virtual Environments	55
4.1.1.7 Data Exploration and Animation	60
4.1.1.8 Electronic Atlases and Web Maps	66
4.1.2 Evolution of Web Resources	72
4.1.3 Derived Information from Text Mining	73
4.1.3.1 Top-15 Words and Top-15 Digital Keywords	73
4.1.3.2 Relationships among Digital Keywords and Their Neighbours	77
4.2 Kraak & Ormeling's <i>Cartography: Visualization of (Geo)Spatial Data</i>	81
4.2.1 Categories from Content Analysis	83
4.2.1.1 GDI and Geodatabases	83
4.2.1.2 Electronic Atlases	90
4.2.1.3 Animation, Multimedia, and Web Mapping	92
4.2.1.4 Map Production	95
4.2.1.5 Remote Sensing Technologies and Applications	98
4.2.1.6 Semantics and Policies	99
4.2.2 Evolution of Web Resources	102
4.2.3 Derived Information from Text Mining	104
4.2.3.1 Top-15 Words and Top-15 Digital Keywords	104
4.2.3.2 Relationships among Digital Keywords and Their Neighbours	107
5 Synthesis and Discussion	111
6 Conclusion and Outlook	119
Bibliography	121
Appendix A	125
Appendix B.....	126
Appendix C.....	127

List of Figures

Figure 1. Map of the US generated by Tobler and data preparation procedures in CGIS.....	5
Figure 2. Geological map of Abingdon, UK	5
Figure 3. Comparison of three editions of the <i>Book of Mormon</i>	7
Figure 4. Text comparison engine (TEI-Comparator)	8
Figure 5. Scheme of the content analysis leading to higher levels of abstraction	9
Figure 6. Text (Slocum et al., 2004) scanned in two different settings	14
Figure 7. Image (Slocum et al., 2004) scanned in two different settings	14
Figure 8. Colour plate (Kraak & Ormeling, 2010) scanned in 300 dpi colour setting	15
Figure 9. Layout detection performed by Layout-Parser	16
Figure 10. Text extraction of Kraak & Ormeling’s book using Layout-Parser	17
Figure 11. Layout detection of Kraak & Ormeling’s book using Layout-Parser.....	18
Figure 12. Text extraction of Kraak & Ormeling’s book using <i>pdfreader</i>	19
Figure 13. Layout clean-up of Slocum et al.’s book	20
Figure 14. Layout clean-up of Slocum et al.’s book	20
Figure 15. Extracted text from Kraak & Ormeling’s book	21
Figure 16. Diagram representing the data transformation procedures	21
Figure 17. Example of the flow diagram showing the relationship between two editions.....	24
Figure 18. Result of using the ‘Compare Document’ tool in PDF-XChange	25
Figure 19. Summary pane of the ‘Compare Document’ tool	25
Figure 20. Change documentation of <i>Cartography: Visualization of (Geo)Spatial Data</i>	27
Figure 21. Binary matrix where documented changes can be filtered by keywords	28
Figure 22. Documentation of web resources in <i>Thematic Cartography and (Geo)Visualization</i> ...	30
Figure 23. Algorithm of the manual stemming process to calculate top-15 words.....	34
Figure 24. Number of unique n-grams by increasing N	35
Figure 25. Excerpt from a directed network graph of an edition	36
Figure 26. Excerpt from a directed network graph of a book	37
Figure 27. Network graph algorithms that are tested on a sample dataset.....	38
Figure 28. Flow diagram of chapters in <i>Thematic Cartography and (Geo)Visualization</i>	40
Figure 29. Flow diagram of chapters from <i>Thematic Cartography and (Geo)Visualization</i> involved in the formation of the category ‘Map Production Methods and Colour’	42
Figure 30. Flow diagram of chapters from <i>Thematic Cartography and (Geo)Visualization</i> involved in the formation of the category ‘GIS and Map Design’	46
Figure 31. Flow diagram of chapters from <i>Thematic Cartography and (Geo)Visualization</i> involved in the formation of the category ‘Generalisation’	49
Figure 32. Flow diagram of chapters from <i>Thematic Cartography and (Geo)Visualization</i> involved in the formation of the category ‘Multimedia’	50

Figure 33. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* involved in the formation of the category ‘Thematic and Topographic Visualisation’ 52

Figure 34. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* involved in the formation of the category ‘Virtual Environments’ 56

Figure 35. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* involved in the formation of the category ‘Data Exploration and Animation’ 60

Figure 36. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* involved in the formation of the category ‘Electronic Atlases and Web Maps’ 66

Figure 37. Classification of hyperlinks in *Thematic Cartography and (Geo)Visualization*. 72

Figure 38. Flow diagram of bar charts showing changes in the rank of digital keywords from *Thematic Cartography and (Geo)Visualization* 76

Figure 40. Directed network graph of bigrams in each edition of *Thematic Cartography and (Geo) Visualization* 80

Figure 39. Directed network graph of bigrams in *Thematic Cartography and (Geo)Visualization* 80

Figure 41. Timeline of *Cartography: Visualization of (Geo)Spatial Data* 81

Figure 42. Flow diagram of chapters in *Cartography: Visualization of (Geo)Spatial Data* 81

Figure 43. Flow diagram of chapters from *Cartography: Visualization of (Geo)Spatial Data* involved in the formation of the category ‘GDI and Geodatabases’ 90

Figure 44. Flow diagram of chapters from *Cartography: Visualization of (Geo)Spatial Data* involved in the formation of the category ‘Electronic Atlases’ 92

Figure 45. Flow diagram of chapters from *Cartography: Visualization of (Geo)Spatial Data* involved in the formation of the category ‘Animation, Multimedia, and Web Mapping’ . 95

Figure 46. Flow diagram of chapters from *Cartography: Visualization of (Geo)Spatial Data* involved in the formation of the category ‘Map Production’ 98

Figure 47. Classification of hyperlinks in *Cartography: Visualization of (Geo)Spatial Data* 102

Figure 48. Example of a figure where the hyperlink to the original source is not provided ... 103

Figure 49. Flow diagram of bar charts showing changes in the rank of digital keywords from *Cartography: Visualization of (Geo)Spatial Data* 106

Figure 50. Directed network graph of bigrams in each edition of *Cartography: Visualization of (Geo)Spatial Data* 109

Figure 51. Directed network graph of bigrams in *Cartography: Visualization of (Geo)Spatial Data* 110

Figure 52. Venn diagram of codes extracted from two textbooks 111

Figure 53. Venn diagram of digital keywords from two textbooks 115

List of Tables

Table 1. Frequency distribution of coding categories for men and women from the content analysis of <i>Rolling Stone</i> magazine covers	10
Table 2. List of candidate textbooks and their editions	13
Table 3. Changes that are excluded from the content analysis	26
Table 4. Dates of change documentation, coding, and re-coding of the main body text	28
Table 5. Dates of change documentation, coding, and re-coding of web resources.....	30
Table 6. Categories of web resources in two textbooks	30
Table 7. Textbook-specific stop words and their examples in textbooks	32
Table 8. Custom in-text citation format applied to the <i>Results</i> section.....	39
Table 9. Comparison of descriptions regarding the sixth consequence of technological changes in <i>Thematic Cartography and (Geo)Visualization</i>	44
Table 10. Continuities of CMS examples in <i>Thematic Cartography and (Geo)Visualization</i>	46
Table 11. Comparison of descriptions regarding the fifth consequence of technological changes in <i>Thematic Cartography and (Geo)Visualization</i>	56
Table 12. Comparison of descriptions regarding Virtual Los Angeles	59
Table 13. Continuities of VE examples in <i>Thematic Cartography and (Geo)Visualization</i>	59
Table 14. Comparison of descriptions of Vis5D in <i>Thematic Cartography and (Geo)Visualization</i>	61
Table 15. Comparison of descriptions regarding the data exploration software	62
Table 16. Continuities of data exploration software examples in <i>Thematic Cartography and (Geo)Visualization</i>	65
Table 17. Continuities of animation examples in <i>Thematic Cartography and (Geo)Visualization</i>	65
Table 18. Continuities of electronic atlas and web map examples in <i>Thematic Cartography and (Geo)Visualization</i>	71
Table 19. Occurrences of hyperlinks in <i>Thematic Cartography and (Geo)Visualization</i>	72
Table 20. Top-15 words of <i>Thematic Cartography and (Geo)Visualization</i>	75
Table 21. Comparison of objectives of the book from prefaces of <i>Cartography: Visualization of (Geo)Spatial Data</i>	83
Table 22. Comparison of descriptions regarding the acquisition of geodata.....	87
Table 23. Continuities of geodatabase examples in <i>Visualization of (Geo)Spatial Data</i>	89
Table 24. Continuities of electronic atlas examples in <i>Cartography: Visualization of (Geo)Spatial Data</i>	92
Table 25. Comparison of descriptions regarding the realisation of animation.....	94
Table 26. Occurrences of hyperlinks in <i>Cartography: Visualization of (Geo)Spatial Data</i>	102
Table 27. Top-15 words of <i>Cartography: Visualization of (Geo)Spatial Data</i>	105

List of Acronyms

ATKIS	Amtliches Topographisches Kartographisches Informationssystem
CMM	Colour Management Module
CMS	Colour Management Systems
CRT	Cathode Ray Tube
DCW	Digital Chart of the World
DEM	Digital Elevation Model
dpi	Dots per inch
DTM	Digital Terrain Models
DTP	Desktop Publishing Software
ESDA	Exploratory Spatial Data Analysis
GDI	Geodata Infrastructure
GIS	Geographical Information Systems
GML	Geographic Markup Language
GPS	Global Positioning System
ICA	International Cartographic Association
ICC	International Colour Consortium
LBS	Location-based Services
LCD	Liquid Crystal Display
ME	Mixed Environments
NHGIS	National Historical Geographic Information System
OCR	Optical Character Recognition
OGC	Open Geospatial Consortium
OS	Ordnance Survey
OSM	OpenStreetMap
PDF	Portable Document Format
SABE	Seamless Administrative Boundaries Europe
SVG	Scalable Vector Graphics
ToC	Table of Contents
USGS	The United States Geological Survey
VE	Virtual Environments
VGI	Volunteered Geographical Information
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Map Service
WWW	World Wide Web

1 Introduction

1.1 Motivation and Problem Statement

Computer-based systems are established as a quintessential tool in cartography. Monmonier (1982a, 1982b) stated that cartography is undergoing its 'digital transition', with more and more geographic data available in machine-readable form and with a rapidly diminishing need to draw and print maps. 40 years after this diagnosis, the evolution of digital cartography is still in progress. It started by using a computer in map production and now it is extended to AI and machine learning. It started from visualising terrains and now it encompasses multi-dimensional data analytics and virtual environments. And due to this ever-changing characteristic and fluid boundaries of digital, cartography has been continuously concerned with the use of cutting-edge technologies. While this trend is inevitable and often beneficial, it can also result in chasing the tail without a sense of orientation. Meng (2003) criticised the phenomenon that a majority of cartographic processes simply aim at making things work by providing default solutions.

The first step in resolving this issue is to take a step back and reflect on the history of digital cartography. And to instantiate this, this research suggests analysing the content of cartography textbooks adopted in universities. University textbooks are a good source of tracing history because they reflect the academic paradigm and reinforce them by instilling values and priorities into future scholars. Although their content may not be up to date, they are a credible source of techniques that are widely accepted and used in cartography. Thus, analysing textbooks over different times can offer insights into how a certain topic has evolved.

Despite its potential, content analysis of cartography textbooks is yet scarcely done. Most of the content analyses revolve around specific types of maps or geography textbooks used in schools. Although they provide general insights into the science domain, they fall short of tracing a specific subfield that is constantly changing. On the other hand, there have been plenty of studies that defined trends in digital cartography to strengthen its genealogy. However, they stay in the theoretical realm without connections to actual learning materials. Thus, tracing the evolution of digital cartography through the content analysis of textbooks pinpoints this niche and can contribute to enriching the discipline of digital cartography.

1.2 Research Identification

This section poses the objectives of this research and questions that help achieve the respective goal. Furthermore, the scope will be set to define what can be expected from this research.

1.2.1 Research Objectives

This project aims to trace the evolution of digital cartography and map the history of digital technologies used in cartography. To reach this goal, the following research objectives are set:

- RO 1** Analysing the changes related to digital cartography found in different editions of cartography textbooks.
- RO 2** Identifying in which ways the increased relevance of digital technologies has been reflected in cartography textbooks.
- RO 3** Exploring how content analysis can be used as a tool for mapping changes in textbooks and understanding the evolution of a theme.

1.2.2 Research Questions

To achieve the **RO 1**, the following research questions will be answered:

- RQ 1a** Which codes related to digital technologies can be derived from cartography textbooks?
- RQ 1b** What kind of patterns and trends are observed concerning digital technologies?

The following research questions concern the **RO 2**:

- RQ 2a** How have the usage and occurrences of web resources in cartography textbooks changed over time?
- RQ 2b** How have the occurrences and neighbours of digital keywords changed over time?

And the following research questions help us achieve the **RO 3**:

- RQ 3a** How useful is content analysis in comparing and understanding changes in the textual content of cartography textbooks?
- RQ 3b** What other methods can be applied to complement the results of content analysis?
- RQ 3c** Can content analysis and other methods be further used to compare different books?

1.2.3 Research Scope

This research proposes how textbooks can be utilised as a window through which the evolution of digital cartography can be observed. Therefore, the findings of this research are confined to the scope of selected textbooks and the time they were written. While it was considered during the selection process (3.1.1) that textbooks should span a broad range of subjects in cartography so that the application of digital technologies can be observed over various topics, it should be noted that the writing of a book is a subjective act, and therefore, not all aspects of digital cartography are expected to be covered in the selected textbooks.

‘Digital cartography’ will include all mapping techniques and technologies that involve computer-based systems or the internet in map production. It can also embrace other analogue mapping techniques; provided that they are important in understanding the lineage of their successive digital technologies. The brief history of digital cartography is shown in 2.1.

‘Evolution’ does not imply linear development with a single directionality but is rather a tree with multiple branches that transforms itself over time. It follows the concept of genealogy by Michel Foucault; a form of history that patiently documents change over time in a non-linear fashion (Foucault, 1984). As a result, any changes, deletions, or additions observed over different editions of textbooks will be counted as a part of evolution.

‘Content analysis’ involves the examination of various documents and texts, which can be quantitative, focused on coding data into predetermined categories in a systematic and easily replicable manner, or qualitative, seeking to uncover deeper meanings in the materials (Bryman & Bell, 2019). The background information of content analysis as a research method is introduced in 2.3, and 3.2 demonstrates how content analysis is applied in this project.

1.3 Thesis Outline

This thesis continues with the second chapter which contextualises the background of this research. After examining previous works in three broader fields to which this research is related, the necessity of this project will be summarised based on what is left unknown and underutilised.

This is followed by methodologies used in this research to answer the research questions. It is divided into four subsections: data collection and transformation, content analysis, author interviews, and text mining. Except for the author interview, each subsection (3.1, 3.2, and 3.4) is further broken down into specific methods that belong under the same research method.

After introducing the methodologies, the results are shown in the fourth chapter. It is organised in a block-by-block format; each textbook is assigned a subsection (4.1 and 4.2) which has an identical structure consisting of three parts. The first part exhibits categories derived from the content analysis through which the evolution of digital cartography can be observed, which are complemented with author interviews. The second part displays the evolution of web resources, and the third part shows the results from text mining.

In the fifth chapter, the results that were presented on a textbook level will be combined together to answer the research questions. Also, it will be discussed how and to which extent the research objectives have been achieved and what the limitations are.

Finally, in the sixth section, the conclusion and the outlook are given to outline what this research has achieved and what other realms can be explored in further research.

2 Background and Relevant Work

This research aims to trace the evolution of digital cartography through content analysis of textbooks. Therefore, it is related to three broader topics: digital cartography, (qualitative) text comparison, and content analysis. This section aims to provide background information regarding those topics, examine what research has been conducted, and summarise their findings. First, studies that concerned the internal history of digital cartography will be covered, as well as a brief introduction to the early development of digital cartography. Second, studies that dealt with comparing different editions of text will be shown. Third, the concept of content analysis and its application in cartography will be introduced. Finally, a small conclusion is provided to recapitulate what is known so far and to identify the unknown.

2.1 Digital Cartography as a Discipline

In 1981, Brian Harley and David Woodward founded the project ‘The History of Cartography’. The purpose of this project was to create a comprehensive encyclopaedia about maps across cultures from prehistoric origins to the twentieth century and consists of six volumes of book series that are divided by time and culture. Among these, the sixth volume *Cartography in the Twentieth Century* was published in 2015, which defined the twentieth century as the pivotal period in map history when the transition from paper to digital formats led to previously unimaginably dynamic and interactive maps (The History of Cartography Project, 2015). Here, they listed the ‘electronic transition’ as one of six themes that characterise the cartography of the twentieth century. With articles and research collected by hundreds of experts, the book describes how GIS and the internet radically altered cartographic institutions and lowered the skill required to be a map author, and satellite positioning and mobile telecommunications revolutionised map-based wayfinding (Monmonier, 2015).

While the entire twentieth century could be labelled as the pivotal period from a broader historical perspective, it was the beginning of the latter half of the century when the start of digital cartography was signalled. Ground zero can be traced back to one of the earliest maps generated with a machine by Tobler (1959). It was drawn by recording the end point of each coastline segment as punched holes in IBM cards and then feeding them into the plotting machine that drew the map (*Figure 1, left*). Several years later, the first form of GIS using computer-based systems was developed by Roger Tomlinson. Designed for the Canadian Land Inventory, it was named ‘Canada Geographic Information System (CGIS)’. While CGIS was not equipped with actual mapping functionalities, it could perform simple measurements and extract information such as distance and area. And in the subsequent paper published in 1968 where the capabilities and working principles of CGIS were introduced (*Figure 1, right*), the term ‘geo-information system’ appeared for the first time; It was defined as “a computer-based information system for the storage and manipulation of map-based land data” (Tomlinson, 1968).

2 Background and Relevant Work

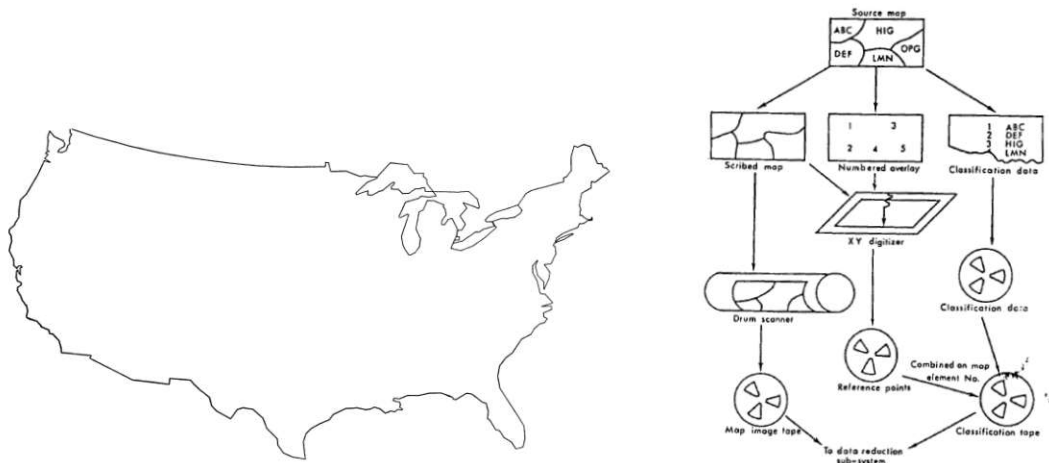


Figure 1. (Left) Map of the US drawn by a plotter and punched cards (Source: Tobler, 1959). (Right) Data preparation procedures in CGIS (Source: Tomlinson, 1968).

The invention of CGIS sparked other countries to examine the effectiveness of utilising computers for managing geodata and producing maps. Some of the institutions that joined this early wave are the US Geological Survey, Ordnance Survey of the UK, and Institut Géographique National of France. In 1971, the UK Experimental Cartography Unit and the British Geological Survey produced a geological map of Abingdon, UK (Figure 2). This is known as the world's first map produced by automated means and published in a standard map series to established cartographic standards (Longley et al., 2015). This event was followed by the commercial software GIS developed by companies such as ESRI, MapInfo, Autodesk, and Intergraph in the 1980s.

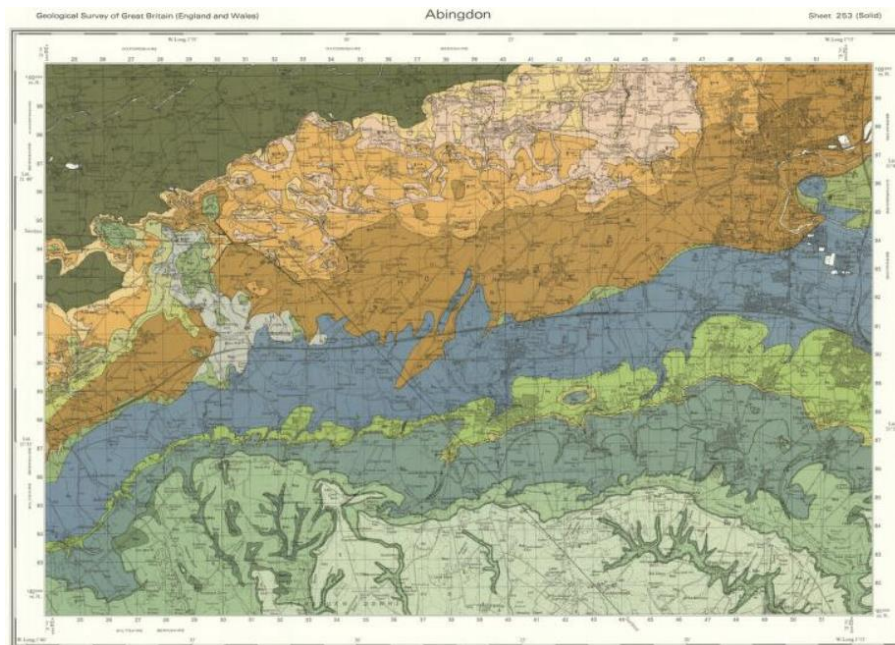


Figure 2. Geological map of Abingdon, UK (Source: British Geological Survey).

Since computer-based systems started to be actively incorporated in the production of maps, many cartographers reflected on the rapid growth of GIS and reviewed the technical difficulties of the time. For example, Tomlinson (1987) provided an overview of the developments and applications of GIS in North America. Here, it was pointed out that developments of GIS have been confined due to technical restrictions as well as a lack of professional personnel and suggested several principles to encourage further development. On the other hand, Goodchild (1991) contended that a lot of technical problems of handling spatial data have been solved over the past years and that GIS and digital technologies will offer solutions to longstanding issues in cartography such as automatic map generalisation.

With the rapid growth of GIS and the increased use of computer-based systems in map production, ‘digital cartography’ manifested itself as a theoretical subfield. However, despite its increased relevance, the majority of the research revolved around the application of novel technologies, and less light was shed on ‘digital cartography’ as a discipline. Therefore, some suggested restructuring the discipline and re-examining fundamental questions such as ‘What is a map?’ and ‘What is digital?’. In 1984, at the ICA meeting in Perth, the term ‘a new cartography’ was introduced to draw attention to the impact of information technologies on the discipline. Part of the concluding argument was that “The greatest . . . challenge facing cartography does not lie in teaching or learning new techniques, but in creating a radically new concept for our discipline” (Taylor, 1985). Similarly, Harley & Woodward (1989) have pointed out the phenomenon of the internal history of cartography being deprioritised as academic cartography has been increasingly redefined by digital technologies. Since cartography is a discipline—whereas GIS and automated cartography are techniques—the scholars saw the need to develop and maintain an intellectual non-applied core which was largely lacking in the discipline and to consider ‘processes’ as well as ‘products’ (Taylor, 1991).

One of the rare definitions that pinpointed digital cartography is the one provided by Visvalingam (1990); It was defined as “the fast-growing sub-discipline concerned with the use of information technology in the making and use of maps”. They also provided three major trends of digital cartography; formal theorising about cartographic processes through automated production of maps, reliable and efficient management of integrated but changeable spatial databases, and appropriate packaging of technology for geographical information systems. Other research focused on the meanings and ramifications of digital cartography. For example, Jackson (1987) noted the adjacent development of remote sensing and image processing technologies along with computer-based systems and diagnosed that the revolutionary consequence of digital cartography is the possibility to produce an ‘ephemeral display’ on a computer screen at the scale and projection that can be adjusted flexibly. This was a dominant viewpoint of the time; that the paradigm shift has only begun and will holistically affect the production, dissemination, and use of maps (Goodchild, 2000). On the other hand, Pickles (1999) argued that this transition is not linear nor single-directed, and it will create uneven patterns of development, resulting in multiple and open paths to future worlds of geoinformation.

Furthermore, the development of digital technologies caused a whole host of ‘x-GIS’ subfields and activities to emerge (Wilson, 2015). As a result, contrary to earlier studies that treated GIS as a single unit, the research trend turned to interdisciplinary applications and tracing their respective developments. One of the phenomenal studies among numerous intersections of digital cartography and other subfields is by Milson & Alibrandi (2008). They explored the integration of GIS in social studies education in middle- and high schools in the US. They analysed curricula, workshops, and practices in schools and examined how geospatial technologies are applied for teaching and learning history, geography, economics, and environmental science. They succeeded in mapping the modern history of the diffusion of geospatial technologies and internet-based GIS in an educational context.

2.2 Comparing Text in Books

Since this research focuses on the textual content of a book, studies that compared different texts in books will be introduced. Depending on the material, a comparison of books can be divided into two types: cross-book and inter-edition. Cross-book comparison is to compare independent sets of text such as two separate books. For example, Korat & Shamir (2004) compared Hebrew versus Dutch electronic books for young children in terms of text and multimedia. They identified components that help children’s literacy development and evaluated the difference between books from the two countries. Cobb (1987) reviewed 16 new introductory statistics textbooks and analysed how far each topic in statistics is carried out, and how well the exercises are designed. In both studies, all of the involved books were independent and non-reactive; None of the books was directly related to another book (e.g., a sequel) nor were they written in consideration of the compared books. Here, the analysis focuses on the similarities and differences of the books; their scopes and theses, materials used, how each book approaches the topic, et cetera.

On the other hand, inter-edition comparison concerns multiple editions of a book that is published under the same title. In this case, it is presupposed that while the goal and overall structure of the book remain constant, there are changes and updates in the content in each edition. Consequently, the focus of the comparison will be set on these differences. For example, Larson (1974) compared textual variations in the two manuscripts and the first three editions of the *Book of Mormon* (Figure 3). While they disclosed that some genuine corrections and editorial revisions were made in 1837 and 1840, it was also concluded that some alterations of the text arose from accidental causes such as scribal or errors in the printing stage.

1830 Edition	1837 Edition	1840 Edition
1. NEPHI 13:40 (32:10)		
the Lamb of God is the [] Eternal Father and the Saviour	the Lamb of God is the <u>Son of the</u> eternal Father, and the Savior	+
		the Lamb of God is the <u>Son of the</u> eternal [J] Father, and the Savior

Figure 3. Comparison of three editions of the *Book of Mormon* (Source: Larson, 1974).

There is an ambiguity surrounding the definition of the word ‘edition’; It is broadly defined as ‘the form in which a text is published’ (Merriam-Webster, n.d.; Oxford Dictionary, n.d.). Therefore, while new editions are usually written due to the elapse of time, they can also mean translated forms or any other versions of the book published by different editors and publishers. For example, Wendling (2005) compared two editions of the collected works of Karl Marx and Friedrich Engels: *the Marx–Engels Collected Works* and *the Marx–Engels Gesamtausgabe*. Although they were published by two different projects, both were referred to as an ‘edition’ because the original authors were the same. Here, each edition’s volume, the coverage of works, and the level of detail were compared and discussed. Furthermore, Salemans (1994) devised a method to discern differences in editions called the ‘fingerprint formula’, by additionally considering scanned images or texts. This idea was extended by ‘The Holinshed Project’ where paragraphs from two editions of *Chronicles of England, Scotland, and Ireland* were matched and compared by a text comparison engine with a web interface (Cummings & Mittelbach, 2010) (Figure 4).

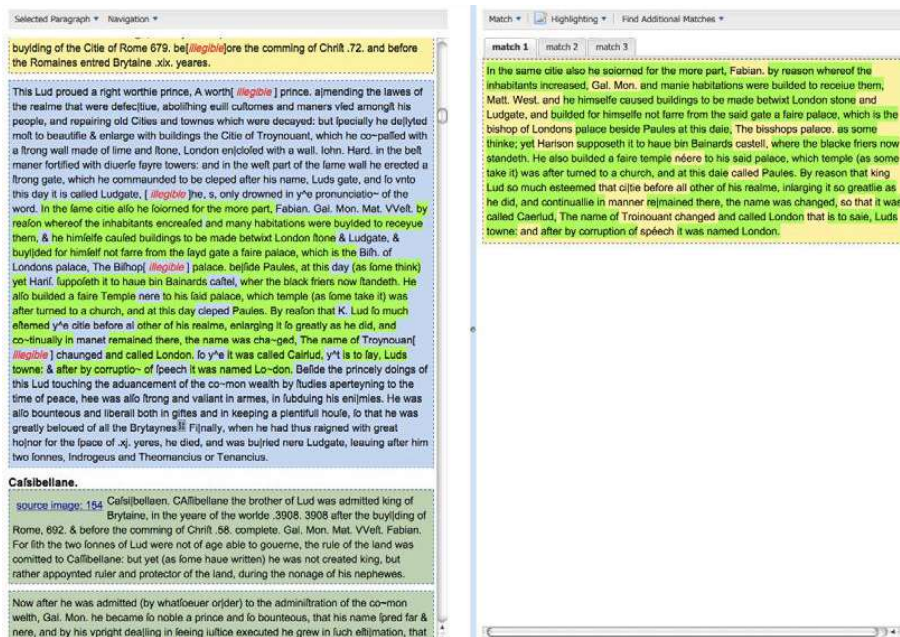


Figure 4. Text comparison engine (TEI-Comparator) displaying a match between two editions of text (Source: Cummings & Mittelbach, 2010).

2.3 Content Analysis

Content analysis is a research technique for making replicable and valid inferences from texts, pictures, audio, video, and other sources to examine patterns in communication (Bryman & Bell, 2011; Krippendorff, 2018). This encompasses a broad range of means of communication such as personal or official documents, literature, newspapers, photos, online posts, and interviews. For the analysis of text-based materials, it is usually performed by compressing many words of text into fewer content categories based on explicit rules of coding (Berelson, 1971).

And this ‘category-based abstraction’ is a distinct feature of content analysis. Categories refer to aspects within the text, which put the meaning of those aspects in a nutshell (Mayring, 2019). In other words, they are groups where sentences or paragraphs are assigned, based on their meanings and characteristics. This abstraction helps content analysts find patterns from a huge dataset, which can be performed on multiple levels (Figure 5). First, condensation is a process of shortening the text while still preserving the core meaning. Second, a code can be thought of as a label; a name that most exactly describes what this particular condensed meaning unit is about. Third, a category is formed by grouping together these codes that are related to each other through their content or context. Finally, a theme can be seen as expressing an underlying meaning, found in two or more categories (Elo & Kyngäs, 2008; Graneheim & Lundman, 2004; Hsieh & Shannon, 2005; Schilling, 2006).

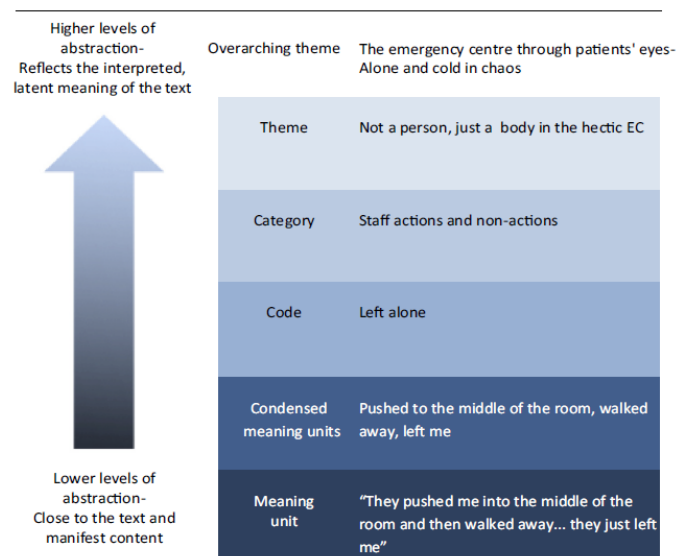


Figure 5. Scheme of the content analysis leading to higher levels of abstraction (Source: Erlingsson & Brysiewicz, 2017).

One of the main goals of content analysis is to find trends by observing a topic over time, which is referred to as a ‘longitudinal approach’. For example, Hatton & Trautner (2011) analysed more than four decades of *Rolling Stone* magazine covers and measured the frequency and intensity of women and men being sexualised. They created a comprehensive scorecard for the degree of sexualisation based on criteria (e.g., body parts, poses, text, and sexual roles) (Table 1) and found out that women are increasingly likely to be ‘hypersexualised’ in comparison to men when represented in popular media. Greer & Mensing (2006) analysed online newspaper websites in the US between 1997 and 2003 and analysed their changes in content, multimedia, and interactive functions. Similarly, Herring et al. (2012) analysed random samples of weblogs from 2003 to 2004, and documented the changes such as structural characteristics of blog entries, blogger name types, and blogging software.

Table 1. Frequency distribution of coding categories for men (M) and women (M) from the content analysis of *Rolling Stone* magazine covers (Source: Hatton & Trautner, 2011).

	Coded as "0"		Coded as "1"		Coded as "2"	
	M	W	M	W	M	W
Clothing/nudity	n = 554 (77%)	n = 78 (28%)	n = 70 (10%)	n = 44 (16%)	n = 56 (8%)	n = 20 (7%)
Touch	n = 496 (69%)	n = 141 (50%)	n = 206 (28%)	n = 93 (33%)	n = 20 (3%)	n = 39 (14%)
Pose	n = 659 (91%)	n = 149 (53%)	n = 50 (7%)	n = 99 (35%)	n = 15 (2%)	n = 34 (12%)
Mouth	n = 595 (82%)	n = 154 (55%)	n = 116 (16%)	n = 105 (37%)	n = 13 (2%)	n = 23 (8%)
Breasts	n = 653 (90%)	n = 154 (55%)	n = 54 (7%)	n = 59 (21%)	n = 17 (2%)	n = 69 (24%)
Genitals	n = 666 (92%)	n = 213 (76%)	n = 48 (7%)	n = 42 (15%)	n = 10 (1%)	n = 27 (10%)
Buttocks	n = 718 (99%)	n = 254 (90%)	n = 3 (< 1%)	n = 16 (6%)	n = 3 (< 1%)	n = 12 (4%)
Text	n = 652 (90%)	n = 177 (63%)	n = 52 (7%)	n = 79 (28%)	n = 20 (3%)	n = 26 (9%)
Head vs. body shot	n = 258 (36%)	n = 40 (14%)	n = 466 (64%)	n = 242 (86%)	–	–
Sex acts	n = 720 (99%)	n = 277 (98%)	n = 4 (1%)	n = 5 (2%)	–	–
Sexual role play	n = 719 (99%)	n = 259 (92%)	n = 5 (1%)	n = 23 (8%)	–	–

Other studies are less interested in observing materials over a long period, and rather focus on interpreting a specific social phenomenon. For example, Bell et al. (2007) analysed local newspapers, census data, and voting records to account for the 2004 election result in Alberta, Canada. They suggested that various factors including populism, leadership, and public sentiment played a role in their win and those standard theories of one-party dominance are incomplete to account for the phenomenon. Burrows et al. (2000) studied various forms of interaction in online forums in the UK and analysed what kind of social support is exchanged on topics such as disability, parenting, and housing debt. Besel et al. (2009) examined the advice contained in books targeted at brides as well as the style of writing, pictures, and comments. They found out that the messages in bridal books reproduce patriarchal ideologies and typical gender roles, peripheralising the participation of grooms and putting an imbalanced strain on brides.

In cartography, content analysis has been mainly applied to maps, seeking to understand changes and patterns in maps over time—both quantitatively and qualitatively. For example, Robinson (1982) evaluated various historical maps and demonstrated that thematic mapping is a technique that has a longer history than common belief. Robinson's method was later adopted by Muehlenhaus (2011); They conducted a content analysis for persuasive maps by breaking down each map among a total of 190 contextual, data, design, and layout variables. Kessler & Slocum (2011) quantitatively and qualitatively analysed map designs in thematic maps that were published in two major geographical journals in the twentieth century. They showed that qualitative measures (e.g., improvement of map design over time) can be quantified, and generated summary statistics for map design elements. Furthermore, Putu et al. (2014) extended the content analysis to story maps created by participatory mapping and sought possibilities of using content analysis as a tool for understanding imminent problems of a local community.

Although less common, some researchers have dealt with the content of cartography textbooks. For example, Tyner (2005) analysed the evolution of two classic cartography textbooks: six editions of Arthur Robinson's *Elements of Cartography* and two editions of Erwin Raisz's *General Cartography*. They compared the changes in the book content from pre-war to post-war cartography and documented the evolution of mapping techniques and philosophies along with the editions. They also found out that the introduction of GIS was the largest change in the sixth

edition of Robinson's book. Here, they noted that this was the first time when the term 'visualisation' was used in the text, which was due to the need to treat databases as a separate concept from visualisation techniques. Kessler (2018) analysed 24 cartography textbooks to explore the development of map projections. They parameterised various map projection characteristics and summarised the evolution of map projection depicted in textbooks as well as examined them if they reflect actual developments in each era. They further tied the results of the content analysis and proposed which projection material should be included in future textbooks.

Mittelstädt (1989) analysed the prefaces of three German atlas books used in schools. Since each book was published at different times (1853, 1944, and 1958), each book was treated independently, and they did not make a comparison among these books. Instead, they analysed each preface in-depth to comprehensively understand the geography education of the time through the authors' sentences. Scholz et al. (2014) evaluated four university-level world geography textbooks in the US and compared them with the former research by Jo & Bednarz (2011) that analysed high school textbooks. They coded the components of spatial thinking by taxonomies and evaluated the percentage of questions that foster students' spatial thinking.

2.4 Conclusion

The daybreak of digital cartography can be dated to the late 1950s and early 1960s when cartographers sought to make maps using machines and the earliest form of GIS appeared. As the relevance of computer-based systems in map production increased, a lot of research focused on the application and prospects of new digital technologies. Consequently, studies of the history of digital cartography have been relatively scarce, resulting in "the lack of intellectual non-applied core in the discipline" (Taylor, 1985).

Comparing two sets of text can be performed across different editions of a book (i.e., inter-edition comparison) or across different books (i.e., cross-book comparison). There have been researchers who compared several editions of a book, most of which concerned spotting changes in historical publications in each edition. However, since each study dealt with specific sets of text and they focused on documenting the very differences, there is yet no clear consensus in terms of how to compare, sort, and analyse the changes in various versions of the text.

Here, content analysis can function as a methodological guideline. Content analysis is a research technique for making replicable and valid inferences from texts and other sources to examine patterns in communication (Bryman & Bell, 2011; Krippendorff, 2018). While some have applied content analysis to cartography textbooks, the longitudinal approach of content analysis has been mainly adopted in cartography for map contents, instead of textual content.

In conclusion, digital cartography as a discipline can be enriched by more historical research, to which this project aims to contribute by analysing different editions of cartography textbooks published in recent decades. Here, content analysis will be adopted to find trends and patterns, which I identify as an underutilised tool in cartography that can help comprehend the evolution of textual information in textbooks.

3 Methodology

This section concerns the methods that are applied to meet the research objectives. First, two textbooks are selected, collected, and transformed into a digital format that can be utilised for analysis. Second, the content analysis is applied to the main text data as well as the web resources of the books. Third, the authors of the textbooks are interviewed to inquire about the reasoning behind the detected changes. Fourth, text mining is performed using R to support and extend the findings from the content analysis. Each step is explained in a separate subsection with input and output summarised at the beginning to help readers track the workflow.

3.1 Data Collection and Transformation

The data preparation process is explained in four steps. First, two textbooks are selected and all of their editions are loaned from libraries. Second, the hardcopy textbooks are digitised into *.pdf* files. Third, optical character recognition (OCR) is performed on the documents and their layouts are cleaned up to exclude parts that are considered irrelevant for the analysis. Finally, the main body text is exported as *.txt* files to be used as input for the analysis in 3.2 and 3.4.

3.1.1 Textbook Selection and Acquisition

Input. University-level cartography textbooks in the market.

Output. Two textbooks with a total of seven editions.

Five university-level cartography textbooks were listed as candidates (*Table 2*). The criteria of the initial selection were as follows; First, a textbook should span a broad range of subjects in cartography so that the application of digital technologies can be observed over various topics. Second, it should have at least three editions to ensure that the book has been continuously updated to keep track of digital technologies. Third, all editions of a book should be accessible via university libraries in Germany or Austria.

Here, considering the ever-changing characteristic of digital technologies, an additional criterion was applied that the latest edition of a book must be published no longer than 10 years ago from the date of selection. With this filter, the list was further narrowed down to two textbooks: Slocum et al.'s *Thematic Cartography and (Geo)Visualization* and Kraak & Ormeling's *Cartography: Visualization of (Geo)Spatial Data*. Since the fourth edition of *Thematic Cartography and (Geo)Visualization* was released on the 18th of August of 2022 and was not available at the time of analysis, the first three editions of Slocum et al.'s book and four editions of Kraak & Ormeling's book were eventually selected for the analysis. Physical copies of a total of seven editions were loaned from two public libraries: the university library of the Technical University of Vienna and Staats- und Universitätsbibliothek Dresden (SLUB).

Table 2. List of candidate textbooks and their editions.

Textbook Title (Edition)	Year	Author(s)
Thematic Cartography and Visualization (1st ed.)	1999	Slocum
Thematic Cartography and Geographic Visualization (2nd ed.)	2004	Slocum et al.
Thematic Cartography and Geovisualization (3rd ed.)	2014	Slocum et al.
Thematic Cartography and Geovisualization (4th ed.)	2022	Slocum et al.
Cartography: Visualization of Spatial Data (1st ed.)	1996	Kraak & Ormeling
Cartography: Visualization of Geospatial Data (2nd ed.)	2003	Kraak & Ormeling
Cartography: Visualization of Spatial Data (3rd ed.)	2010	Kraak & Ormeling
Cartography: Visualization of Geospatial Data (4th ed.)	2020	Kraak & Ormeling
Kartographie I (5th ed.)	1975	Hake
Kartographie II (2nd ed.)	1976	Hake
Kartographie I (6th ed.)	1982	Hake
Kartographie II (3rd ed.)	1985	Hake
Kartographie (7th ed.)	1994	Hake & Grünreich
Kartographie (8th ed.)	2002	Hake et al.
Cartography: Thematic Map Design (1st ed.)	1985	Dent
Cartography: Thematic Map Design (2nd ed.)	1990	Dent
Cartography: Thematic Map Design (3rd ed.)	1993	Dent
Cartography: Thematic Map Design (4th ed.)	1996	Dent
Cartography: Thematic Map Design (5th ed.)	1999	Dent
Cartography: Thematic Map Design (6th ed.)	2009	Dent
Map Use and Analysis (1st ed.)	1990	Campbell
Map Use and Analysis (2nd ed.)	1993	Campbell
Map Use and Analysis (3rd ed.)	1998	Campbell
Map Use and Analysis (4th ed.)	2001	Campbell

3.1.2 Book Digitisation

Input. Seven hardcopy textbooks.

Output. Seven *.pdf* files.

Since the textbooks are acquired in hard copies, they should be transformed into a digital format. Although this is not an absolute prerequisite for content analysis—reading and analysing a book can be done in an analogue setting—all textbooks are digitised in this project for the following reasons. First, digitisation is necessary for text mining which requires machine-readable text as input. Second, the maximum rental period that libraries allow did not cover the entire time the analyses would need. Third, since the focus of this project is to compare multiple editions of a book, the content analysis requires constant searching and comparison. Not only would this process be overly tedious in an analogue setting, but it also drastically hinders the precision of the analysis. As a result, considering both characteristics of this research and the realistic demands, all books were scanned from the first page to the end and exported in *.pdf* format.

The digitisation is initiated by scanning every page of a textbook using a book scanner. This is done in three different configurations of a scanner: 300 dpi greyscale, 400 dpi black & white, and 300 dpi colour. Whereas the first option is sufficient for regular books with larger typefaces, the chosen textbooks were published in very small serif typefaces and often contained italic typefaces, causing frequent blurs in scanned images (*Figure 6*). Lower scanning quality consequently hinders OCR, a process where the text is recognised from scanned images. Therefore, the books were scanned again in black & white instead of greyscale, assigning each pixel with binary values (i.e. black or white) so that no blurred grey pixels exist in the scanned area. Also, the dpi was increased from 300 to 400 (The higher dpi of 400 was only allowed in black & white in the book scanner used, and for other modes, the maximum was 300). This resulted in crispier images and the quality of OCR was satisfactory after testing it with several parts of books that contained italic typefaces and smaller text such as footnotes. While this is beneficial in terms of text recognition, it has the downside of distorting the grey values of photos (*Figure 7*). Therefore, to conserve the images, each book was scanned in greyscale as well.

categories of sites relevant to cartography.

In addition to providing a wealth of maps and related information, the Web also changes the way in which maps are produced and used. Prior to the existence of the Web, cartography could be characterized as *supply-driven*—cartographers created maps and mapping software that they and other experts thought would be useful. With the advent of the Web, cartography can be considered

(a)

categories of sites relevant to cartography.

In addition to providing a wealth of maps and related information, the Web also changes the way in which maps are produced and used. Prior to the existence of the Web, cartography could be characterized as *supply-driven*—cartographers created maps and mapping software that they and other experts thought would be useful. With the advent of the Web, cartography can be considered

(b)

Figure 6. Text (Slocum et al., 2004) scanned in two settings: (a) 300 dpi greyscale, (b) 400 dpi black & white.

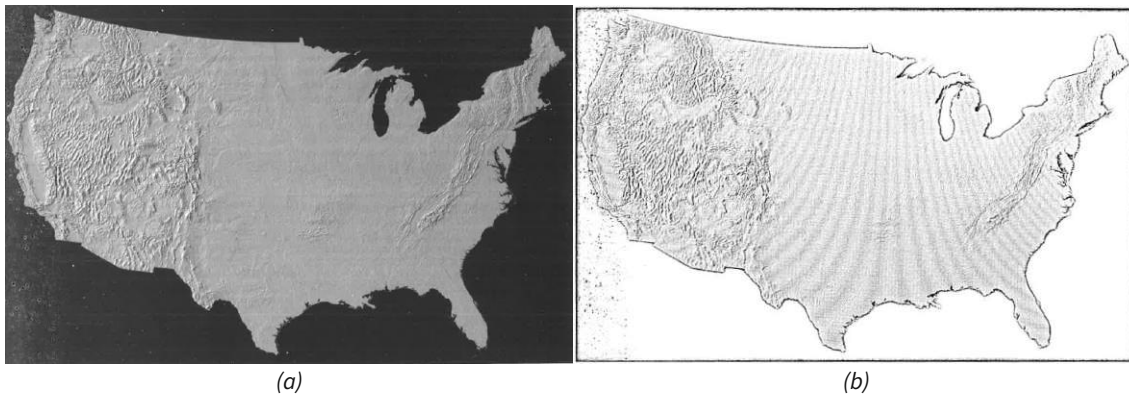


Figure 7. Image (Slocum et al., 2004) scanned in two settings: (a) 300 dpi greyscale, (b) 400 dpi black & white.

Also, while the analysed books were all mainly printed in greyscale, some of the more recent editions contained a few pages printed in colours, attached at the end or episodically inserted in the middle. These are called ‘colour plates’—pages dedicated to images that must be printed in colours—and are exclusively scanned in 300 dpi colour (*Figure 8*). Whereas this separation of the greyscale main text and a handful of colour plates has been the de facto standard for a lot of

textbooks, it is changing due to the constant development in the printing software and hardware. For example, both Kraak & Ormeling (2020) and Slocum et al. (2022) published their fourth edition in full colour, which is exciting news for students since they have access to colour images on every page without having to search for colour images on the back. On the other hand, for research purposes, this would require all pages to be scanned in colour, especially when images in the book are a crucial part of the analysis. Not only do colour scanned images take up more memory space than greyscale, but their OCR quality is also the lowest compared to black & white and greyscale images because their background—which is supposed to be completely white in an ideal scanning condition—also contains colour information, therefore making it more difficult to be cleanly separated from the text.

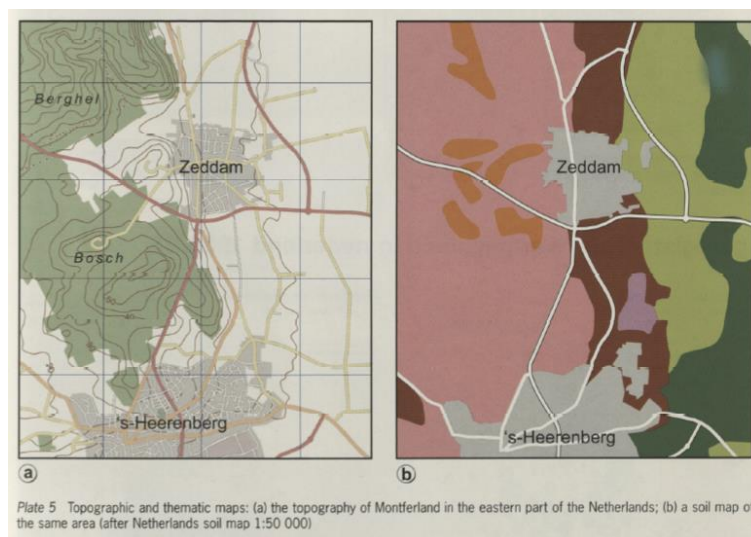


Figure 8. Image from a colour plate (Kraak & Ormeling, 2010) scanned in 300 dpi colour setting.

3.1.3 OCR and Layout Clean-Up

Input. Seven *.pdf* files.

Output. 112 *.pdf* files, divided by edition and chapter.

Once the book is exported in a *.pdf* format, the next step is to recognise the text using an OCR engine. The OCR engine converts the scanned images into machine-encoded text and embeds a transparent text layer underneath the text area of the PDF document, making the document a ‘searchable image’. This is essential in this project where searching and cross-checking the keywords throughout multiple books is required. After the OCR, the layout has to be cleaned up to remove unnecessary regions so that only text regions are left. Since OCR engines search every page from corner to corner, they recognise letters and numbers that do not help construct a clean text file, such as page numbers and contents inside tables. Also, it is highly likely that some parts of diagrams or images are mistakenly recognised as letters. Because cartography textbooks

have a particularly complex layout where text, tables, charts, maps, and diagrams are densely arranged, the OCR results in a messy PDF document that requires rearrangement.

There are a myriad of OCR engines in the market and their performance, accuracy, platform, and prices highly vary. Among the most well-known ones is the open-source OCR engine Tesseract. On the other hand, there are high-end commercial engines that promise the best accuracy as well as fast processing time. An example of this is ABBYY FineReader Engine, an artificial-intelligence-based OCR engine equipped with various image processing and document analysis functionalities which are run on their cloud server. The obvious drawback of using these cutting-edge technologies is that their prices are often not affordable for individual researchers. Thus, this subsection describes two options that were tested for the OCR and the layout clean-up in terms of efficiency and accuracy. These are, first, Python library 'Layout-Parser' and second, the commercial PDF software application 'PDF-XChange Editor'.

3.1.3.1 Automatic Layout Classification and OCR Using Layout-Parser

Shen et al. (2021) developed a unified toolkit for deep learning-based document image analysis called Layout-Parser. It currently supports two OCR engines: Tesseract and Google Cloud Vision, both of which are proven to be sufficiently good at detecting and extracting text. Aside from the OCR functionality, the breakthrough of Layout-Parser is its ability to automatically detect complex layouts of a document and to classify each part into various regions such as text, title, figure, table, and list (Figure 9). This is enabled by a spectrum of pre-trained models on different datasets such as journal articles and newspapers. In other words, instead of applying OCR to the entire document and cleaning up afterwards, Layout-Parser first defines document areas and then selectively runs OCR. Because of this more efficient workflow, Layout-Parser was initially selected as the main method for extracting text, minimising manual intervention in cleaning up the layout after the OCR.



Figure 9. Layout detection performed by Layout-Parser (Source: Shen et al., 2021).

However, models utilised by Layout-Parser do not distinguish different text regions such as captions and footnotes, and they are all classified into a single category. As a result, although the

orders and the topology of the main body text were correct, a lot of captions were inserted between the text columns when the text was extracted (*Figure 10; Figure 11*). In theory, these captions could be removed with the help of regular expressions (regex), however, not all captions consist of a single line or a single sentence, making it difficult to come up with a regex pattern that applies to all textbooks. This becomes even more complicated when dealing with footnotes that do not begin with a clear cue such as ‘Figure’ but are preceded by small symbols such as an asterisk (*) or a cross symbol (†) that are often wrongly recognised as a plus sign (+), the lowercase T (t), or other dot-shaped symbols (°). This depends on small differences in the scanning quality of each page and is highly unpredictable. Also, other minor recognition errors were found when testing with sample pages, such as mistaking the lowercase L in italic (*l*) with a slash (/), which happened less frequently in other OCR engines of commercial software. Consequently, after some test runs, it turned out that Layout-Parser would not ensure manual-intervention-free results for this project which seeks to extract specifically the main body text.

Furthermore, almost all of the pre-trained models supported by Layout-Parser are based on ‘Detectron2’—an object detection algorithm developed by Facebook AI Research—which is difficult to install in a Windows-based environment due to some known errors. Because of this, Windows users should configure a workaround to install Detectron2, which requires substantial knowledge of computer programming. For the test runs, another workaround of using an online Python executing platform ‘Google Colab’ was utilised in order to avoid this operating system dependency issue. Although this temporarily resolved the problem, running the model on an online platform was highly unstable and often resulted in an error that it is unable to find the model inside the library. Also, the processing time was significantly slower compared to the commercial software since the model could not be installed locally and had to be run on an online server. As a result, despite its ability to sort major layout groups, Layout-Parser was not adopted as a tool for OCR and cleaning up the layouts.

```

▶ for txt in text_blocks.get_texts():
  print(txt, end='\n--\n')

```

can be noticed that the tools for presentation are the most highly developed (Robinson et al., 1995). While producing maps to communicate spatial information, cartographic rules (together called ‘cartographic grammar’) are available to make the maps effective. However, as these rules are not part of the GIS software, GIS allows users to produce their own maps even when they are unaware of cartographic grammar. In other words, there is no guarantee that the maps will be effective. These cartographic rules could also be applied in the analysis phase, but the necessity to do so would be less strong here. When cartographers and analysts discuss this matter, the second group would always claim ‘who cares about your

Figure 1.1 The map domain in a changing world of spatial data handling: influences of technology, map user environments and internal developments lead to a change in functions

rules, as long as one understands one’s own maps?”

Figure 10. Text extraction of Kraak & Ormeling’s book using Layout-Parser. The caption (highlighted in blue) is wrongly recognised as the main body text.

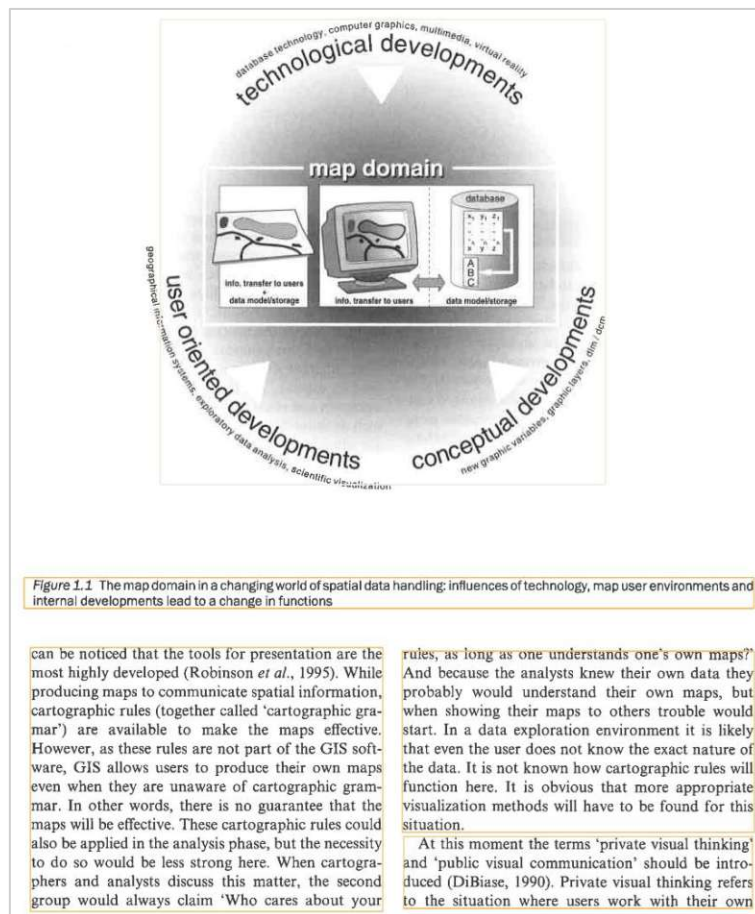


Figure 11. Layout detection of Kraak & Ormeling's book using Layout-Parser. Only diagrams (marked in beige) and the text (marked in yellow) are differentiated.

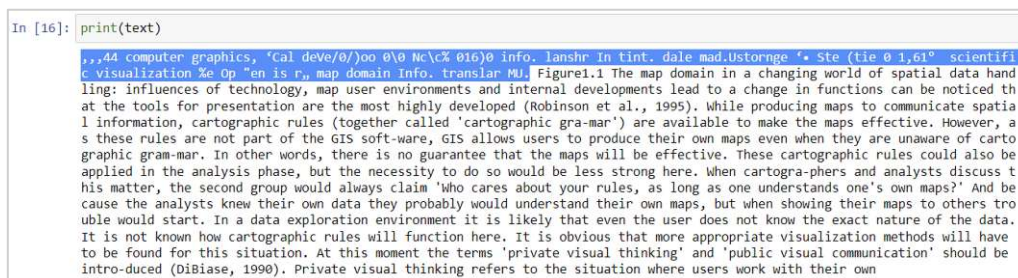
3.1.3.2 OCR and Manual Layout Clean-Up Using PDF-XChange Editor

PDF-XChange Editor is a commercial software application distributed by Tracker Software that is designed to read, create, and edit PDF documents. The distinct advantage of using software applications compared to Python libraries is that they offer pre-treatment steps (e.g., deskewing, background removal, and text sharpening) that help enhance the quality of the consequent OCR. Deskewing makes sure that paragraph blocks are aligned as perpendicular as possible. This is important because image processing would be computationally more expensive if the rotation has to be taken into account. Also, since OCR places the recognised letters in a row so that they form a sentence, it might result in unwanted line changes if the paragraph is severely tilted. Background removal and text sharpening are useful in case the paper is not purely white or the letters are not purely black, where OCR cannot easily separate the background from the text. These two steps were irrelevant for this project since the books were scanned in black & white and the OCR was accurate even only with the deskewing. However, depending on the scanning method and the colour composition of a book, it may be necessary to remove the background and sharpen the text.

Another advantage of using a software application that offers graphic user interfaces is that it makes it easier to spot pages that are missing or that require rescanning. These errors are usually found after the OCR by visually debugging the OCRred areas. If a surrounding text box does not cover the entire text region or is deformed, it means that a certain portion of a paragraph is not properly scanned. Therefore, several pages of textbooks were rescanned at this stage. While all of these functionalities are included in Adobe's Acrobat DC as well, PDF-XChange Editor produced more accurate OCR results. As a result, it was used for the OCR of the textbooks.

As mentioned earlier, OCR does not result in a pure text file that is ready to be used for the analysis. It only makes the PDF document a 'searchable image' with a lot of wrongly recognised letters such as inside diagrams and images, which results in an additional step of cleaning up the layout. Layout-Parser was anticipated to make this process simpler by first sorting the document areas before the OCR. However, since it could at best distinguish major layout components (e.g., text, title, figure, table, and list) and not further such as captions and footnotes, the OCR was performed using PDF-XChange Editor. This results in a PDF document containing both relevant and irrelevant text areas, which means that the irrelevant ones should be manually deleted.

Since this is a repetitive and tedious process, three Python libraries that handle PDF files were tested against several OCRred pages if they can detect main text areas without having to manually delete the irrelevant ones. These are *pdfplumber*, *pdfreader*, and *pdfminer.six*. The testing proved that they simply copy and paste all of the OCRred regions and that none of the libraries understands the layouts (Figure 12). In conclusion, the manual layout clean-up was inevitable.



```
In [16]: print(text)
',,44 computer graphics, 'Cal deve/0/0o 0\0 Ne\c% 016)0 info. lanshr In tint. dale mad.Ustorage '+ Ste (tie 0 1,61° scientifi
c visualization %e Op "en is r, map domain Info. translar MJ. Figure1.1 The map domain in a changing world of spatial data hand
ling: influences of technology, map user environments and internal developments lead to a change in functions can be noticed th
at the tools for presentation are the most highly developed (Robinson et al., 1995). While producing maps to communicate spatia
l information, cartographic rules (together called 'cartographic gra-mar') are available to make the maps effective. However, a
s these rules are not part of the GIS soft-ware, GIS allows users to produce their own maps even when they are unaware of carto
graphic gram-mar. In other words, there is no guarantee that the maps will be effective. These cartographic rules could also be
applied in the analysis phase, but the necessity to do so would be less strong here. When cartogra- phers and analysts discuss t
his matter, the second group would always claim 'who cares about your rules, as long as one understands one's own maps?' And be
cause the analysts knew their own data they probably would understand their own maps, but when showing their maps to others tro
uble would start. In a data exploration environment it is likely that even the user does not know the exact nature of the data.
It is not known how cartographic rules will function here. It is obvious that more appropriate visualization methods will have
to be found for this situation. At this moment the terms 'private visual thinking' and 'public visual communication' should be
introduced (DiBiase, 1990). Private visual thinking refers to the situation where users work with their own'
```

Figure 12. Text extraction of Kraak & Ormeling's book using *pdfreader*. Some of the contents inside a diagram are wrongly recognised as text, resulting in a faulty text extraction (highlighted in blue).

The manual layout clean-up was performed by visually checking every page of an OCRred document using PDF-XChange Editor and deleting irrelevant or wrongly recognised text areas. For example, the left half of Figure 13 shows the initial OCR result performed by PDF-XChange Editor. While only the region highlighted in yellow is the main body text, a lot of other parts are also recognised such as the chapter name, page number, figure caption, and letters and symbols inside the figures. In other cases, although less common, there were text boxes that were wrongly connected because the letters inside the figures interfered with them. The left half of Figure 14 shows this error, which required deleting all OCRred areas of the page and rerunning the OCR in a manually selected region.

After this process, the PDF documents are only left with titles, subtitles, and the main body text that are needed for the text analysis. Finally, the cleaned-up PDF documents were divided into multiple files by chapters. This is a necessary step for the chapter-by-chapter comparison described in 3.2.2. As a result, a total of 112 .pdf files are generated from multiple editions of two textbooks, divided by edition and chapter.

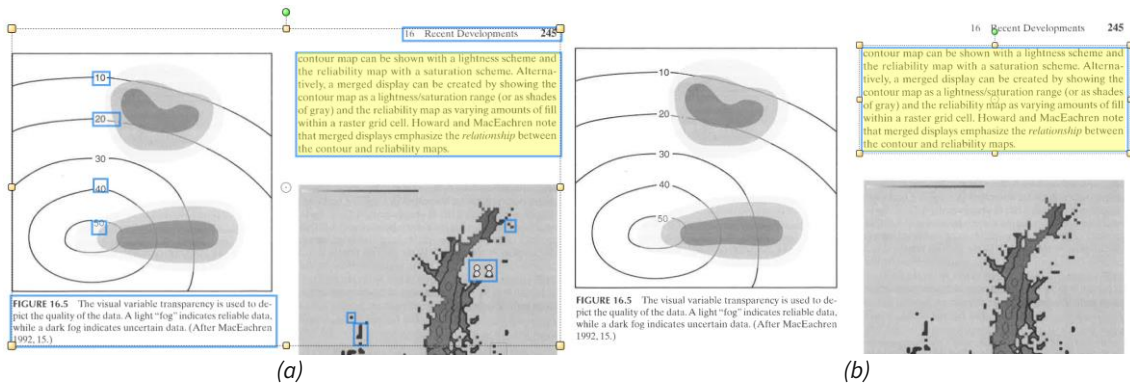


Figure 13. Layout clean-up of Slocum et al.'s book. The main text region is highlighted in yellow. (a) The initial OCR results where many regions are wrongly recognised as text, (b) After the clean-up.

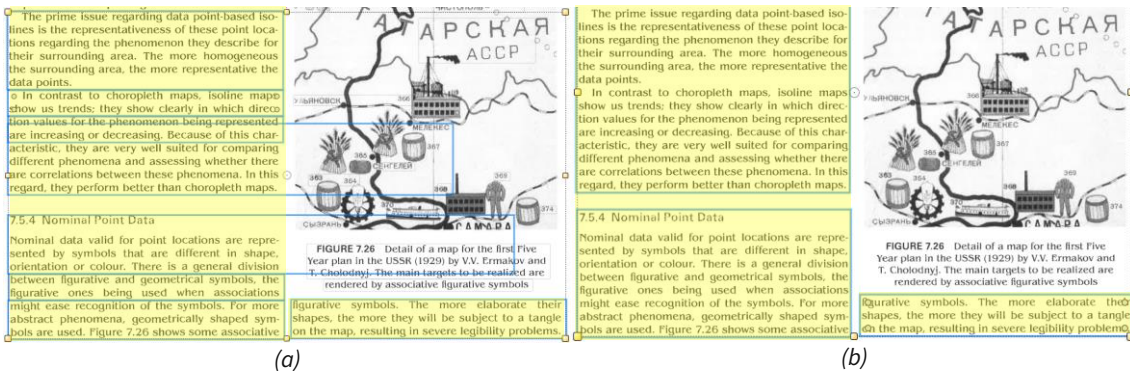


Figure 14. Layout clean-up of Slocum et al.'s book. The main text region is highlighted in yellow. (a) The initial OCR results where text boxes are wrongly connected, (b) After the clean-up and the re-OCR.

3.1.4 Text File Generation

Input. 112 .pdf files, divided by edition and chapter.

Output. 112 .txt files, divided by edition and chapter.

With the manual clean-up, all irrelevant text areas are deleted except for titles, subtitles, and the main body text. Thus, the remaining text can be extracted as a .txt file and be used as input for text mining and chapter-by-chapter comparison. This is done by selecting all text in PDF-XChange Editor and then copying and pasting it to the text editor. This results in a text file with some unnecessary hyphenations and line breaks. An empty line indicates that the paragraph

was interrupted by a page that does not contain any text, and hyphenations are the result of directly copying and pasting the text from a PDF document. Whereas empty lines will usually be automatically ignored by text analysis packages, hyphenations should be removed because they can interfere with the analysis, depending on the package. For instance, Tidytext replaces hyphens with an empty space to separate two connected words—the word ‘computer-based’ will be broken down into ‘computer’ and ‘based’. This means that for example in *Figure 15*, it will separate ‘visualiza’ and ‘tion’, which is not the desired outcome. To avoid this, all hyphens followed by a line break are replaced with an empty space using regex (*Figure 15*). Also, all line breaks are replaced with a single space so that the text file can resemble the look of a written article instead of a narrow block. As a result, a total of 112 *.txt* files are generated from multiple editions of two textbooks, divided by edition and chapter. The entire data transformation procedure from 3.1.2 to 3.1.4 is schematised into a diagram in *Figure 16*.

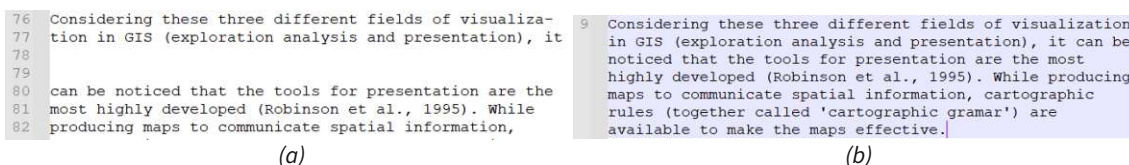


Figure 15. Extracted text from Kraak & Ormeling’s book.
 (a) The initial result containing empty lines and hyphenations, (b) After the removal and tidying.

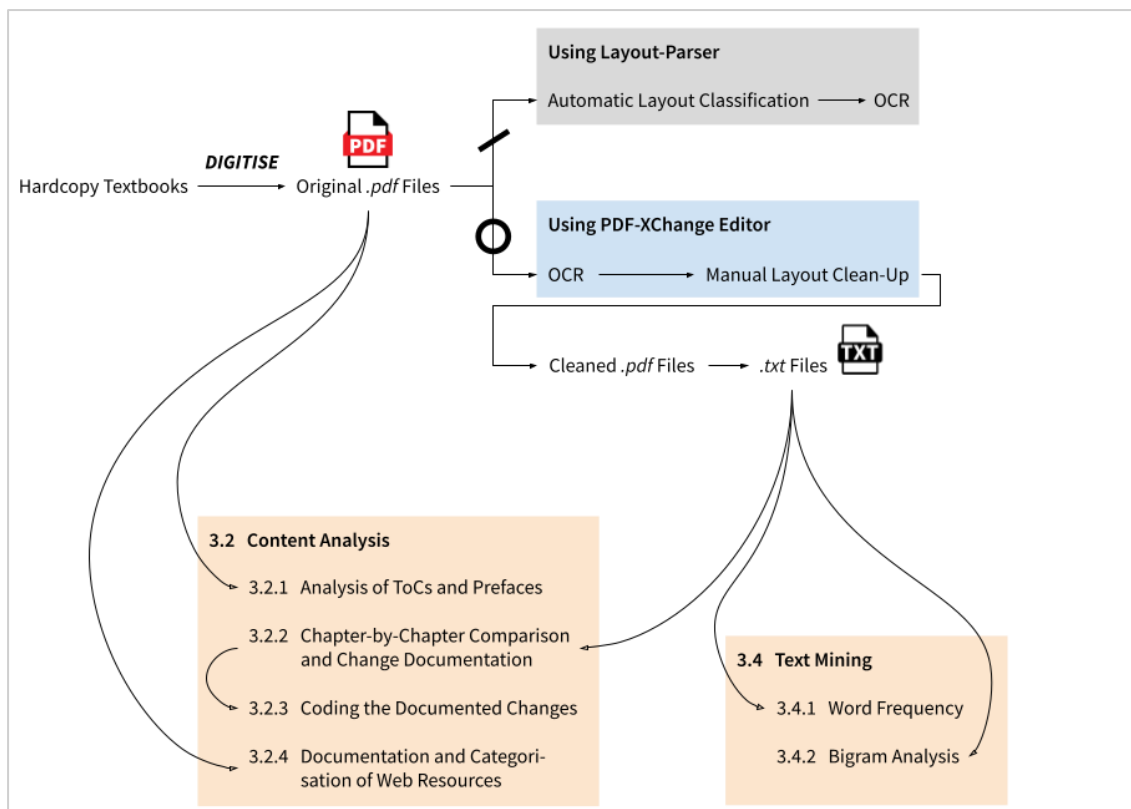


Figure 16. Diagram representing the data transformation procedures.

3.2 Content Analysis

Content analysis involves the examination of various documents and texts, which can be quantitative, focused on coding data into predetermined categories in a systematic and easily replicable manner, or qualitative, seeking to uncover deeper meanings in the materials (Bryman & Bell, 2019). Since this research codes the text from textbooks into several categories, it can be seen as quantitative. On the other hand, it also has qualitative aspects, because it seeks to exhibit meanings by detecting changes within developed categories. Because this project adopts both qualitative and quantitative and they are not separable, the distinction between the two will not be made throughout the research and the term ‘content analysis’ will uniformly be used.

This research is a combination of three approaches suggested by Hsieh & Shannon (2005): conventional, directed, or summative. It is conventional because categories are directly derived from the text data instead of predetermined, but it is also directed to an extent because the overarching theme ‘digital cartography’ limits the categories that can be developed and changes irrelevant to this theme are excluded from the analysis. Also, it has a summative aspect because the number of web resources used in textbooks is counted and compared.

Despite adopting these typical approaches, this research simultaneously deviates from traditional content analysis for two reasons. First, the content analysis of this research requires less abstraction. This is because it deals with textbooks where the text is more structured, explanatory, and descriptive—with little emotion or judgment—than interviews and other sources of natural language. Thus, there is less room for interpretation as well as less need for abstraction. Second, this research does not actively extract the latent meaning from the text; The focus lies on the comparison of different editions. Should the reasoning behind the change be known to better understand the change, it is complemented by author interviews. Therefore, unlike traditional content analysis, meanings are not extracted from the text but are rather exhibited through rearranging and restructuring the changes, making them more pronounced.

And because of its heavy emphasis on changes, the research objectives of this project cannot be met with the traditional coding method, represented by a series of abstractions that contributes to figuring out the latent meaning of the text. Here, the editions of a textbook are not treated as an independent text corpus but are rather a set of corpora that are loosely related in chronological order. In addition, the very content from a single edition of a textbook does not possess much significance. What becomes meaningful is its derivative—when the content is situated in comparison to other editions and its changes and consistencies are examined. Therefore, the research requires additional processing of the text which enables the detection and documentation of changes along the editions. Its detailed procedures are described in the subsections below.

3.2.1 Analysis of ToCs and Prefaces

Input. Tables of contents and prefaces from eight textbooks.

Output. Flow line diagrams (Based on codes originally created by Florian Ledermann).

The content analysis is initiated at the macro level by examining the table of contents (ToC) and the preface of each edition. A ToC is a list usually located in the very first part of a book, itemising titles of chapters and subchapters along with their commencing page numbers. A preface is an introduction to a book, usually found between the ToC and the main written part of a book. Here, authors provide information such as their motivation for writing the book, topics covered, and distinctive features and characteristics of their book. If a book has more than one edition, this is where the authors will list chapters that are updated in comparison to the last edition with some explanation to justify the changes.

Although the role of ToCs and prefaces are preliminary and they do not belong to the main body text, they both play a key role in the content analysis of mapping changes. First, ToCs help researchers grasp the overall structure of a book, enabling them to initially sort more relevant chapters and the less relevant ones. Even when the entire book is to be analysed, it is still helpful to have an image of which sections of a book are the most likely to contain important information. For example, in this project, it could be anticipated that the chapters handling map projections are likely to be relatively static without many changes, especially regarding digital cartography. On the other hand, chapters such as «*Data Exploration*» and «*Map Animation*» were to be approached with heightened attention since they are likely to involve computer-based systems. Equipping oneself with this initial idea helps in planning the consequent content analysis and structuring the change documentation. Second, the emergence, deletion, and restructuring of chapters are observed the most clearly by analysing the ToCs. Since chapters of a textbook are indicative of topics that the authors found the most relevant at the time of writing the book, their changes provide meaningful insights into tracing the evolution of a certain theme.

Along with ToCs, prefaces can assist in understanding the overall updates of a book. This is because, in prefaces, the authors introduce major changes made in this edition and demonstrate the reasoning behind each decision. Comprehending the authors' intention is crucial so that researchers do not make a rash claim concerning updates of an edition based on their assumptions. This is especially important for (sub)chapters that are deleted because whereas newly introduced chapters are likely to reflect the increased relevance of the topic, it is more difficult to infer the reason for the deletion of chapters. While a deletion can occur due to the decreased relevance (e.g., outdated technologies), it can reversely mean the increased relevance or could be due to other editorial reasons; For example, in the preface of the fourth edition, Slocum et al. explain why they decided to delete the chapter «*Web Mapping*» that “A dedicated chapter on Web Mapping seemed unnecessary given that we now make heavy use of web links throughout the textbook” (Slocum et al., 2022). In this case, the deletion of a chapter does not imply that web maps are deemed less important, but is rather the opposite; They have become ubiquitous, permeating every part of the new edition instead of existing as a separate chapter. Therefore, reading prefaces can help understand the overall changes of a new edition.

The analysis of ToCs and prefaces is summarised into a flow diagram that is created using HTML, Javascript, and D3. The purpose of this visualisation is to give both readers and the researcher a graphic image of the evolution of the book structure. Each box represents a chapter

of a book and its height is proportional to the total pages of the chapter. The boxes are stacked, constituting an edition of a book. The boxes are linked to boxes of the subsequent edition using flow lines, based on the analysis of prefaces. The flow lines were later refined with findings from the content analysis; For example, while it was mentioned in the preface of the third edition that the chapter «*Web Mapping*» is newly introduced, connections of its content with the chapter «*Electronic Atlases and Multimedia*» were identified during the content analysis. Therefore, the chapter «*Web Mapping*» is linked with the chapter «*Electronic Atlases and Multimedia*». Chapters without sufficient inheritance of contents from the earlier edition are considered ‘completely new’ and are marked with a small ring on the left side of the box. Similarly, chapters, where the majority of contents were deleted in the following edition, are considered ‘discontinued’ and are marked with a vertical bar on the right side of the box. Also, due to the restructuring of a book, some chapters are derived from multiple chapters from the earlier edition and others are divided into multiple chapters in the following edition. These changes are represented by joining or ramifying flow lines (Figure 17). These diagrams are shown at the beginning of 4.1 and 4.2, along with the written description of major changes that took place throughout the editions.



Figure 17. Example of the flow diagram showing the relationship between two editions.

In addition to the main diagram that visualises the overall change of editions, another set of diagrams is derived from the main diagram. These display only some chapters that are relevant to keywords extracted from the content analysis. The purpose of these zoomed-in diagrams is to help readers precisely grasp which chapters are involved in explaining the evolution of each keyword. They are inserted in every subsection under 4.1.1 and 4.2.1.

3.2.2 Chapter-by-Chapter Comparison and Change Documentation

Input. 112 .txt files, divided by edition and chapter.

Output. Two spreadsheets documenting relevant changes per book.

As mentioned at the entrance of 3.2, this project requires additional processing of the text before the coding in order to focus on the changes. This part describes how the text is processed to extract the changes and how the changes are documented.

With the help of diagrams that show the connection among chapters, it can be roughly estimated which chapters should be compared against each other. The comparison is done on a chapter level, using .txt files divided by edition and chapter that were generated in 3.1.4. Also, to accurately detect all the changes, the comparison is assisted by computer-based systems. The difficulty of detecting changes in literature is that the changes take place in an organic and

complex manner. Unlike lines of programming codes, the textbook content does not merely get deleted or added; A lot of sentences become paraphrased or relocated to other paragraphs. This makes it nearly impossible for existing Python libraries (e.g., *difflib*, *changedetection.io*) to intelligently consider all these possible revisions—they would mark almost everything as deleted and newly added. Therefore, a more sophisticated tool is required for change detection.

PDF-XChange Editor has a ‘Compare Documents’ tool which accepts both *.pdf* and *.txt* formats. The comparison of two *.pdf* files was unsatisfactory because they contained other information such as layouts and figures, and they interfered with the change detection. Therefore, the *.txt* files were chosen. The mechanism of this tool is to compare the new document against the old one and mark the detected changes, classified into three types: replacements, insertions, and deletions (Figure 18). An interesting feature aside from showing all changes is the result sheet summarising total changes (Figure 19). While this could provide interesting quantitative insights, it was not included in the analysis for the following reasons; First, the comparison tool detects all possible subword-level changes (e.g., spelling differences, articles, and spaces) which fall outside the scope of this research. Second, a lot of detected changes are not numerically comparable with one another. For example, an insertion of an entire paragraph is counted as a single change, as well as the change in the spelling from ‘thirdly’ to ‘third’.

The introduction of on-screen maps and their corresponding databases resulted in a split between these functions. To cartographers it brought the availability of database technology and computer graphics techniques that resulted in new and alternative presentation options such as three-dimensional and animated maps. In a GIS environment, geospatial analysis often begins with maps; maps support judging intermediate analysis results, as well as presenting final results. In other words, maps play a major role in the process of geospatial analysis. They have acquired an additional role with the advent of geospatial data infrastructures (GDI). These new infrastructures for accessing geospatial data are being developed all over the world in order to allow access to the geospatial datafiles created and maintained in order to monitor the population, resources and environment. Access to the data needed requires complex querying procedures that are simplified when using maps to pinpoint the areas and themes for which data are needed (Figure 1.1).

Figure 18. Result of using the ‘Compare Document’ tool in PDF-XChange. It displays word-level changes (highlighted in yellow) and an inserted paragraph (highlighted in blue).

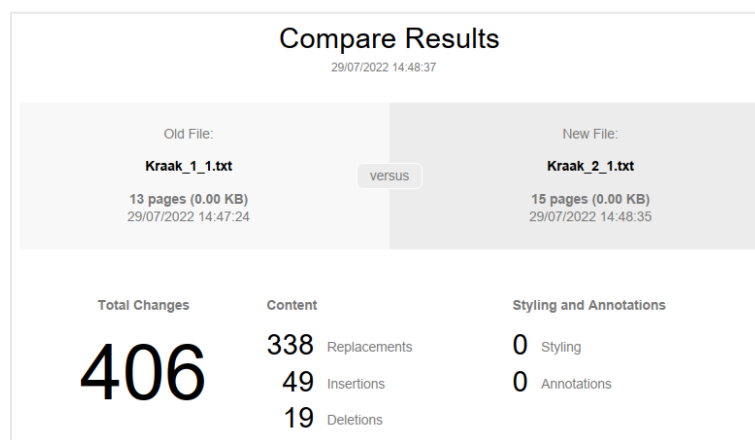


Figure 19. Summary pane of the ‘Compare Document’ tool. It shows the total number of replacements, insertions, and deletions between two documents.

And since the comparison tool displays every difference between two *.txt* files, the majority of them are irrelevant to explaining the evolution of digital cartography. Therefore, a set of rules was established to exclude these changes (Table 3). All other detected differences are documented in a spreadsheet where each sheet is one chapter and each column represents an edition. Figure 20 shows how the comparison of the first chapter of Kraak & Ormeling's book is documented. Here, column [B] contains the newly added sentences and subchapters in the second edition. Cells marked in red represent parts that were deleted in the subsequent edition. For example, the cells [A9:B9] indicate that the entire subchapter *«The Relation Between GIS and Cartography»* was deleted in the second edition and a new subchapter *«The Relation Between GDI and Cartography»* was introduced. Since it was difficult to propose a universal threshold from which point it is considered that a subchapter is replaced with a new one, this involved carefully reading and comparing the individual chapters to make a decision, in case the comparison tool marked a substantial amount of a subchapter as deleted or updated.

Unfortunately, a one-to-one comparison of two *.txt* files was not feasible in some cases, especially in Slocum et al.'s book which went through a substantial restructuring of chapters. Many subchapters were relocated to a different chapter, some chapters were merged to form a larger chapter, and some ramified to smaller chapters. Thus, chapters that have gone through these major restructuring were additionally identified and broken down into a subchapter level. Here, subchapters whose connections were found in other chapters were again put into the comparison tool with their corresponding subchapter. The remaining ones whose connection could not be identified were marked as either discontinued or newly added (Figure 20).

Table 3. Changes that are excluded from the content analysis.

Types of Changes	Example
Changes in terms that rather reflect a larger paradigm shift than the evolution of digital cartography.	“spatial” ~ “geospatial”, “GIS” ~ “GIScience”
Spelling changes and paraphrased sentences.	“geographical” ~ “geographic”, “fulfill” ~ “fulfil”
Descriptions added or subtracted that are highly unlikely to reflect the evolution in digital cartography.	~ “The notion of time is not as straightforward as one might think.”
Updated examples and websites due to the passage of time and of which description has not changed significantly.	“... the population of the Netherlands in 1970-1980” ~ “1990-2000”
Changes in the description of file formats or data types that are NOT repetitively emphasised throughout the chapter.	“GIF” ~ “PNG”
<i>Changes in embedded hyperlinks. (Documented separately in 3.2.4)</i>	

	A	B	C	D
1	CHAPTER 1 Geographic information systems and maps			
2	Kraak_1	Kraak_2	Kraak_3	Kraak_4
3				The ever more detailed satellite imagery available, the increasing number of sensor networks and new
4			The rise of Internet brought the next revolution in mapping. Access to interactive maps is no longer	
5		They have acquired an additional role with the advent of geospatial data infrastructures (GDI). These new		
6		Possibilities for interaction are boosted by the advent of the Internet and its potential for querying the		
7		Currently GIS is used in virtually all disciplines that require geospatial data to execute their tasks or solve		
8		The potential for analysis is already greatly enhanced by the possibilities of applying GIS processing on the		
9	1.5 The relation between GIS and cartography	1.5 The relation between GDI and cartography (whole new)	1.5 The spatial data infrastructure and maps	
10		Companies and government departments at different levels (municipal, provincial/county, state, or national)		
11		In order to enable data users to find out whether data sets from different information systems can be		
12		A next step is the development of distributed geospatial warehouses, central repositories where not		
13	GIS and cartography Many of the concepts and functions of GIS were first conceived by cartographers.			
14			In Europe the GDI implementation is guided by the EU INSPIRE initiative which, based on legislation, will	

Figure 20. Change documentation of the first chapter of *Cartography: Visualization of (Geo)Spatial Data*.

3.2.3 Coding the Documented Changes

Input. Two spreadsheets documenting relevant changes per book.

Output. Two spreadsheets with an additional column containing codes.

Once all the relevant changes are extracted and documented, they are ready to be categorised. While there are many levels of abstraction utilised in the content analysis (Figure 5), in this project, the documented changes are first categorised at the code level. A code can be thought of as a label; a name that most exactly describes what this particular condensed meaning unit is about—usually one or two words long (Erlingsson & Brysiewicz, 2017). Codes are selected as the appropriate level of abstraction considering the objective of this research, which is to map the evolution of digital cartography through various keywords.

Codes are developed by carefully reading the changes documented in the spreadsheet. Also, since they were not predefined, they were constantly revised and updated along the reading. Each change was allowed to have multiple codes; For example, the following sentences that were newly added to the fourth edition of Kraak & Ormeling's book were assigned with three codes: *web*, *remote sensing*, and *participation*.

This approach is also used by volunteers who want to update map data for humanitarian purposes after disasters such as hurricanes, flooding or earthquakes (in the so-called Missing Maps projects). The volunteers use recent satellite imagery to, for instance, indicate where the damage is most severe, and the new maps, often part of the OpenStreetMap, can be used to restore the basic infrastructure (M. J. Kraak & Ormeling, 2020).

To minimise the mistake of misassigning or omitting codes, it is recommended that the material is re-coded after a 10-14 day hiatus if a single coder is involved (Schreier, 2012). Following this advice, the dates of documenting and coding the changes were tracked, and they were revisited and revised after 10-14 days (Table 4).

Table 4. Dates of change documentation, coding, and re-coding of the main body text.

	Change Documentation	Coding	Re-Coding
Slocum et al.	6/1 ~ 6/10	6/18	6/29
Kraak & Ormeling	6/11 ~ 6/15	6/19	6/30

After the re-coding, a binary matrix was created so that the changes can be filtered by codes. Since the changes are documented along the flow of the textbooks, rearranging them in a different order provides an alternative view of observing the changes across the book. This is especially beneficial when a topic is broadly covered throughout multiple chapters. For example, in Slocum et al.'s books where GIS is not assigned with an independent chapter, it would be useful to show only the changes that are related to GIS. To achieve this, first, the comma-separated codes were split to form a list of all recorded codes. Second, duplicates were removed from the list. Then, it was transposed to form a header row where each column is a unique code. Finally, the matrix was filled out by searching for the header inside the original comma-separated cell; The value is 1 if the code exists, if not 0. As a result, a sparse binary matrix is created through which only the rows that contain a certain code can be filtered (*Figure 21*).

The result of the main text content analysis is exhibited under 4.1.1 and 4.2.1.

	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC
1	anima	atlas	cost	critica	datab	defini	desigr	displa	distrib	educa	esda	file for	gdi	gis	hardw	intern	lbs	multir	partici	printir	rs
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
11	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
12	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
22	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
24	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

Figure 21. Binary matrix where documented changes can be filtered by keywords.

3.2.4 Documentation and Categorisation of Web Resources

Input. Seven .pdf files, OCRed.

Output. Two spreadsheets documenting all hyperlinks and their categories.

Web resources—often signified as hyperlinks when referenced—help readers accurately access the examples of the book without having to search for them. This has gained importance with the growing relevance of the web because the majority of maps are published on the internet

and their interactive functionalities cannot be fully delineated in a printed textbook. Thus, they are a direct proxy of how digital technologies are integrated into the textbook.

Although the internet and the web are covered through the methods described above, this research documents the web resources of textbooks separately for the following reasons. First, hyperlinks are not the creation of the textbook authors and a lot of them refer to external materials. Thus, they cannot be handled in the same bucket as the main text content analysis which focuses on the internal changes of the written text. Second, introducing web resources inside a hardcopy textbook is a specific phenomenon and it provides other insights from observing the general development of the web. Third, the analysis of web resources is largely quantitative and does not follow the same procedure described in 3.2.1 to 3.2.3. Therefore, this subsection concerns the content analysis method of web resources included in the textbooks.

To make a complete list of web resources, all hyperlinks that appear in textbooks are collected. Here, unlike the other content analysis, the captions of figures and tables are included in the analysis. This is because a lot of figures in textbooks are acquired from the internet, and therefore, their captions are linked to their original sources. The search is done by simply searching the OCRed PDF documents using two keywords: *http* and *www*. Both keywords are utilised because, first, some hyperlinks only contain either one of *http* or *www*, and second, there could have been an error during the OCR. Then, each hyperlink was copied and pasted into a spreadsheet along with its context (*Figure 22*). These were initially documented in a similar way as the former analysis—side-by-side change documentation—in order to track the changes along the editions. However, because web resources are frequently updated and relocated, the majority of web resources have become obsolete and could not be accessed. Therefore, as it turned out to be impossible to document the changes, the analysis method was steered to a more quantitative approach; focusing on the occurrence of hyperlinks.

The web resources are grouped into three large categories per textbook which were revisited after a 10-14 day of hiatus for re-categorising (*Table 5*). There are three points to be noted here. First, for web resources, a higher abstraction level ‘category’ is used instead of ‘code’. This is because the population of the web resources are significantly smaller than the text data, therefore, a code is too specific a unit to group the web resources. Second, unlike the main text content analysis where each row could be assigned with multiple codes, the web resource categories are exclusive. Third, two out of three categories developed from the two books are not identical because of their different scopes (*Table 6*).

The result of the analysis of the web resources is exhibited under 4.1.2 and 4.2.2.

3 Methodology

	A	B	C
1	Slocum_1	Slocum_2	Slocum_3
2	CHAPTER 1 Introduction	CHAPTER 1 Introduction	CHAPTER 1 Thematic Cartography and Geovisualization
3	Longitude and latitude of major U.S. cities http://www.bcca.org/misc/qibih/latlong_us.html	http://www.census.gov/cgi-bin/gazetteer	http://www.census.gov/cgi-bin/gazetteerAttribute
4	Population estimates for counties https://www.census.gov/population/www/estimates/county.html	http://census.gov/popest/counties/	http://www.census.gov/popest/counties/Maps
5	3D representation of pollution in New York harbor https://www.rpi.edu/locker/69/000469/dx/harbor.www/maxus.dieldrin.cr	http://www.rpi.edu/locker/69/000469/dx/harbor.www/maxus.dieldrin.cr.gif	http://www.rpi.edu/locker/69/000469/dx/harbor.www/maxus.dieldrin.cr.gif
6	Animation of urban sprawl in the San Francisco Bay area http://edcwww2.cr.usgs.gov/umap/umap.html	http://www.ncgia.ucsb.edu/projects/gig/v2/About/abApps.htm	http://www.ncgia.ucsb.edu/projects/gig/v2/About/abApps.htm
7	Software for creating choropleth maps at the state level http://maps.esri.com/ESRI/mapobjects/tmap.htm	http://maps.esri.com/esri/mapobjects/tmap/tmap.htm	Software for creating thematic maps http://factfinder.census.gov/saff/SAFFInfo.jsp?_pageId=thematicmap&_submenuId=maps_0
8	Project Argus (Software for exploring data) http://midas.ac.uk/argus	CommonGIS (Software for exploring data) http://www.ais.fhg.de/SPADE/products/index_eng.html	http://www.ais.fhg.de/SPADE/products/index_eng.html
9	Massachusetts Electronic Atlas http://icg.harvard.edu/~maps/maatlas.htm	The Atlas of Canada http://atlas.gc.ca/site/index.html	http://atlas.nrcan.gc.ca/site/index.html
10	Khoros (Tools for developing your own software) http://www.tnt.uni-hannover.de/soft/imgproc/khoros	GeoVista Studio http://geovistastudio.sourceforge.net/	http://geovistastudio.sourceforge.net/Teaching
11	The Geographer's Craft (Materials for students) http://www.utexas.edu/depts/grg/gcraft/contents.html	http://www.colorado.edu/geography/gcraft/contents.htm	http://www.colorado.edu/geography/gcraft/contents.html
12	Cartography: The Virtual Department (Materials for instructors) http://www.utexas.edu/depts/grg/virtdept/contents.html	http://www.csun.edu/~hfgeg005/cwg/	http://www.colorado.edu/geography/virtdept/contents.html
13		Numerous examples of mapping the Internet and its associated information spaces can be found in Martin Dodge and Rob Kitchin's Atlas of Cyberspace (2001a) and on Martin Dodge's Website http://www.cybergeography.org/atlas	

Figure 22. Change documentation of web resources from *Thematic Cartography and (Geo)Visualization*.

Table 5. Dates of change documentation, coding, and re-coding of web resources.

	Change Documentation	Coding	Re-Coding
Slocum et al.	6/22	7/18	7/29
Kraak & Ormeling	6/23	7/20	7/31

Table 6. Categories of web resources in two textbooks.

Slocum et al.	Kraak & Ormeling
Web maps, geodatabases, and other map-like examples	
Software related to mapping techniques	Initiatives, organisations, and projects
Others (Wikipedia, projects, external lists)	Others (Wikipedia, toponyms)

3.3 Author Interviews

Author interviews are adopted as the first method of complementing the result from the content analysis. The necessity stems from two reasons; First, detailed backgrounds and processes of the book writing can be understood more in depth through interviews. Although authors typically mention their main motivation and objectives in the preface, these descriptions are concise and do not provide specific contexts. Moreover, there may be discrepancies between the written words and their current opinion, which also provide insights regarding the progress of a book. Second, most of the changes between the two editions are unexplained, except for some major structural changes covered in the preface. Other updates, additions, and deletions happening on a subchapter level are very unlikely to be mentioned, making it difficult to infer why the authors

made these changes. And while some updates can be inferred by mobilising general knowledge (e.g., decreasing number of descriptions about televisions between 2000 and 2015 coincides with the decline of television during that period), others cannot.

Considering these two objectives, interview questions are divided into two sections. The first section consists of six questions concerning the general writing process. These are open-ended questions and are asked identically for both author groups. The motivation for updating each edition is included in the questions as well, even though they are covered in prefaces, to capture further background behind the publishing (e.g., the opinion of the publisher and other realistic reasons). Also, considering that most of the authors are university professors, it was inquired what kind of changes they have recognised in the cartographic teaching environment. The second section concerns specific updates of a book that require further elaboration and is prepared separately per book. Since each question pinpoints a specific update, they are formulated in a more closed manner such as ‘Why was this example deleted in the third edition?’.

For the authors of *Thematic Cartography and (Geo)Visualization*, the live interview could not be done because of time zone differences. Therefore, the main author Terry Slocum gathered answers from the co-authors and provided answers to the interview questions in written form. In the case of *Cartography: Visualization of (Geo)Spatial Data*, one of the authors Menno-Jan Kraak was interviewed online. The interview questions are provided in Appendix A and B.

3.4 Text Mining

Text mining is the second method of complementing the result from the content analysis, which provides different perspectives by quantitatively approaching the text data. Unlike numerical data that can directly be processed by computers, text documents are less structured and often contain fuzzy and ambiguous relations. Thus, specific (pre-)processing methods and algorithms are required to extract useful patterns (Hotho et al., 2005). These data processing methods are referred to as text mining; a knowledge-intensive process in which a user interacts with a document collection over time by using a suite of analysis tools. In a manner analogous to data mining, text mining seeks to extract useful information from data sources through the identification and exploration of interesting patterns (Feldman & Sanger, 2006).

The data transformation of hardcopy textbooks was described from 3.1.2 to 3.1.4, which resulted in *.txt* files that can be used as input for text mining. This subsection concerns how the text data are handled to derive information, using the R package ‘tidytext’.

Tidytext is based on a ‘tidy data principle’ suggested by Wickham (2014), where each variable is a column, each observation is a row, and each type of observational unit is a table. Using this, Silge & Robinson (2016) developed the tidytext package which allows the conversion of text to and from tidy formats, and to switch seamlessly between tidy tools and text mining packages. For this project, tidytext is utilised with a data manipulation package ‘dplyr’ to perform two types of rudimentary text mining: word frequency and relationship among words using bigram. The R codes used for the analysis are provided in Appendix C.

3.4.1 Word Frequency

Input. Seven *.txt* files.

Output. Seven lists of top-15 words and digital-related keywords with respective occurrences.

A central question in text mining and natural language processing is how to quantify what a document is about. One measure of how important a word is its term frequency; how frequently a word occurs in a document (Silge & Robinson, 2017). Here, in addition to the most frequently occurring words that represent the character of an edition of a book, their changes along the progress of editions are also mapped.

To achieve this, first, a *.txt* file is opened which contains all the main body text of an edition of a textbook, generated in 3.1.4. This results in a table where each row represents a sentence from a book. Since the analysis is to be applied on a word level, these sentences have to be further broken down. This cutting process—referred to as tokenisation—is performed by the function *unnest_tokens* in *tidytext*, splitting words that are separated by an empty space or a hyphen. This results in a table where each row represents a word.

The next step is to remove stop words. Stop words are words that are not useful for the analysis because their function is mainly grammatical (i.e., articles and conjunctions) or they appear too often in general text documents without adding much meaning (e.g., actually, again, sometimes, et cetera). *Tidytext* contains a dataset *stop_words* that contains a total of 1,149 stop words that are collected from three lexicons. In addition, there are words that are not included in the *stop_words* dataset but appear frequently in both textbooks and should be removed as well. For example, the word ‘figure’ ranks high on word frequency because many sentences are referring to images in the textbook. Including these words in the analysis can be misleading because their usage is associated with the internal structure of the books, rather than their thematic content. Therefore, these textbook-specific stop words are identified (Table 7) and appended to the *stop_words* list, as well as numbers from 0 to 99. After excluding all stop words using the function *anti_join* of *dplyr*, the table contains only relevant words.

Table 7. Textbook-specific stop words and their examples in textbooks.

Stop Word	Example
figure(s)	“This possibility is shown as a dashed line in Figure 1.3.”
e.g	“ e.g. a field or fields of wheat”
chapter(s)	“The dasymetric map will be covered in Chapter 10.”
section(s)	“The basic concept of data standardization was introduced in section 2.4.”
plate(s)	“A map created using their color scheme is shown in Color Plate 9.3.”
based	“contour- based methods”, “computer- based systems”
numbers from 0 to 99	“We will explore the notion of geographic brushing further in Chapter 21 .”

After excluding stop words, a simplification process called stemming is performed to prune conjugators and plurals of words. This step is necessary because verbs and nouns exist in various forms according to the grammatical context in natural languages. Since this hinders the text analysis such as word count, words should be reduced to the common base. For example, a researcher should likely treat two words ‘computer’ and ‘computers’ as one word because they refer to the same concept and their word frequency will be underestimated by the separation.

Stemming can be done both automatically and manually. An example of automatic stemming is the R package ‘SnowballC’ which is based on the stemming algorithm developed by Porter (1980). It treats complex suffixes as compounds made up of simple suffixes, and removes the simple suffixes in several steps, resulting in the most basic forms of words. Although this offers the fastest and simplest solution, automatic stemming is not implemented in this research for the following reasons. First, it wrongly stems some newly coined words that are not in its lexicon such as GIS. Second, it mechanically reduces every conjugation regardless of the difference in meanings. For example, words such as ‘inform’, ‘information’, and ‘informative’ are all reduced to ‘inform’, which is grammatically sensible but inaccurate when considering the definition and the usage of each word. A more problematic case is the ‘projection’, where it is mainly used in the context of map projection and has little semantic connection with its base ‘project’. Although some stemming packages can remove only plurals such as the ‘S-Stemmer’ by Harman (1991), this leaves other cases untreated (e.g., reducing an adverb to an adjective). They also do not offer debugging or allow manual inputs, making it impossible to customise the stemming algorithm.

Alternatively, manual stemming can be done by `dplyr`’s `case_when` function which works similarly to the *if-else-elif* syntax. Using this function, specific conjugation forms can be specified and replaced with a certain word. This manual workflow is not optimal in a case where a large number of stems should be analysed since it is nearly impossible to go through every word in a document and specify its conjugations. On the other hand, if a limited number of stems are the focus—which is the top-15 in this project—this is a safer and more robust approach because it can be controlled which conjugations are to be reduced to which form.

This starts by sorting out 30 candidate words from the simple word count. Starting from the most frequently used word, the word is searched in all editions of a textbook, documenting conjugations that should be reduced to their basic form. For example, in the first edition of Slocum et al.’s book, the word ‘map’ was listed on the top and its conjugated forms included ‘maps’, ‘mapping’, ‘mapmaker(s)’, ‘mapped’, ‘hypermap(s)’, et cetera. Among these, only the word ‘maps’ is selected to be reduced to ‘map’. The words ‘mapping’ and ‘mapped’ are excluded despite their shared base because they are likely to refer to specific mapping methods or the action of making a map, instead of an actual map. After replacing the conjugations, the entire word count is recalculated. This is repetitively done until the 30th-frequently-used word and until the top-15 words stay unchanged. The above process is performed per edition of the two textbooks, resulting in a total of seven lists of top-15 words. This iterative process is schematised in *Figure 23*.

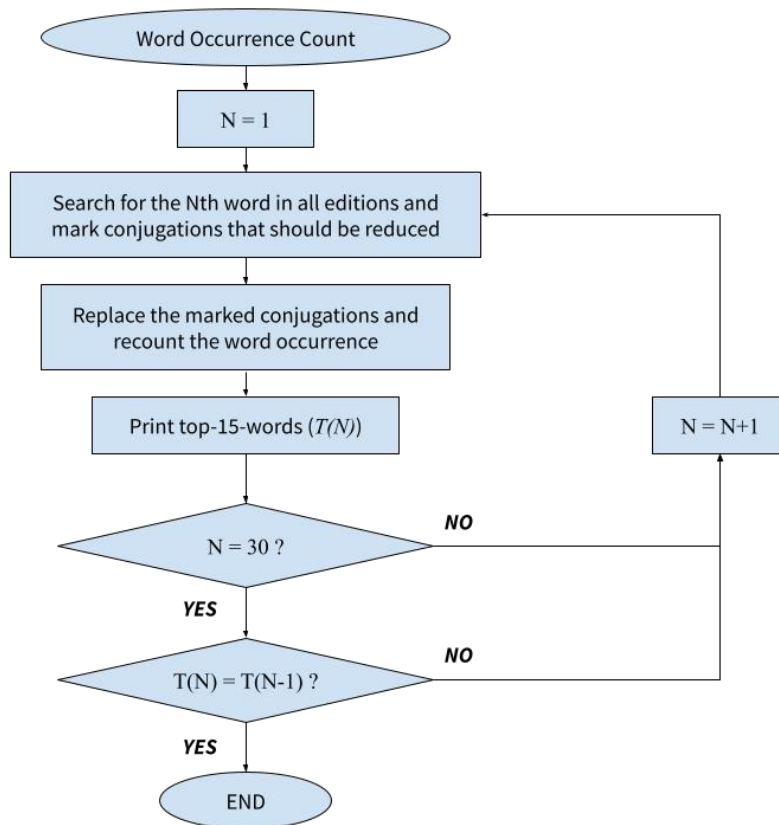


Figure 23. Algorithm of the manual stemming process to calculate top-15 words.

While top-15 words represent the general structure and priorities of a textbook, this research's emphasis is on digital cartography. Therefore, in addition to the top-15 words, the top-15 keywords related to digital cartography are derived separately. To formulate top-15 list of digital keywords, the coded result from the content analysis is utilised. By carefully examining the word count of each edition, top-15 keywords related to digital cartography are documented, based on the content analysis result. When selecting the keywords, ambiguous cases in which the context has to be taken into account are excluded. For example, words such as 'display' and 'monitor' have different meanings when used as a verb or a noun. Since these differences cannot be distinguished from this text analysis method, they are excluded from the keywords. Since this extraction method is prone to error, the process was repeated after 10-14 days of hiatus from the initial keyword selection.

This results in a total of seven lists of top-15 keywords that are relevant to digital cartography. It should be noted that these keywords are edition-specific; Each edition of a book has its own list of keywords based on the order of occurrence. Finally, they are visualised into bar charts using the package 'ggplot2'.

Results from the word frequency are exhibited in 4.1.3.1 and 4.2.3.1.

3.4.2 Bigram Analysis

Input. Seven *.txt* files.

Output. Seven lists of bigrams filtered by digital-related keywords derived from 3.4.1.

While word frequency provides a succinct overview of a document, it contains the least amount of context because the analysis is performed on a word level. Many interesting text analyses are based on the relationships between words, whether examining which words tend to follow others immediately, or that tend to co-occur within the same documents (Silge & Robinson, 2017), which cannot be explored when every word exists independently as a unit.

N-grams can overcome this limitation by considering multiple words instead of one. For example, the word ‘digital cartography’ has more information value than just ‘digital’ or ‘cartography’. This principle of sequencing words is called an n-gram; a bigram is a two-word sequence of words, and a trigram is a three-word sequence. Using n-grams, probabilities can be assigned to sequences of words, which are called language models (Jurafsky et al., 2014), which become a foundation of predictive text input and auto-completion. This research will focus on n-grams’ characteristic of considering multiple words as a unit and examine the relationship among words; which words tend to co-occur in an edition of a book and if their relationships have changed over time. Also, bigram was chosen as the unit because of the relatively small text corpus of the analysed textbooks. While a larger N contains more context and information, it has a trade-off of becoming sparser (*Figure 24*). For example, the chance of ‘digital’ and ‘cartography’ appearing as a pair will be much lower than in their separate occurrences. This is especially critical when the text corpus is small such as in Kraak & Ormeling’s books; A lot of n-grams will end up having extremely low counts. Therefore, bigram is adopted for the analysis of both books.

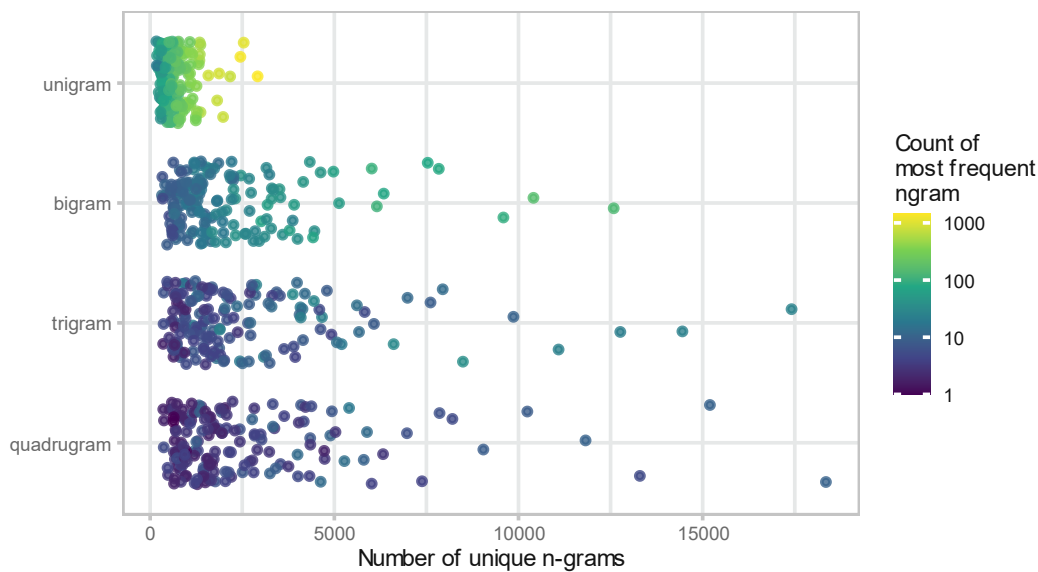


Figure 24. Number of unique n-grams by increasing N.
Each point represents a H.C. Andersen fairy tale. (Hvitfeldt & Silge, 2021)

The analysis is initiated by slicing each edition into bigrams. In tidytext, this is done by giving the argument ($token = 'ngrams', n = 2$) to the `unnest_tokens` function. This results in a table where each row represents a bigram. It should be noted that bigrams are produced with overlaps; For example, the sentence 'Maps are the best' is broken down into 'Maps are', 'are the', and 'the best'.

After this, conjugations and stop words should be removed before the actual bigram count. Since the list of conjugations and stop words are already generated in 3.4.1, they are used for the bigram analysis as well. First, the bigrams are separated by a single space, resulting in a table with two columns where each row of a column is a word instead of a bigram. Then, using the list from above, certain plural forms and conjugations related to digital cartography are replaced with their basic form. From here, rows that contain stop words are removed so that only relatively more meaningful pairs of words are left.

Next, this is further filtered with the list of digital keywords, again already developed in 3.4.1. However, before the filtering, all of the digital keywords across the editions are unioned to form a single list per book. The reason for aggregating them per book is to observe the change over time; For example, if top-15 digital keywords per edition are directly applied to their corresponding edition, keywords that do not overlap will not appear consistently in the result of bigram analysis, making it impossible to observe the change along the editions. Therefore, the digital keywords are unioned across the editions. Then, they are used to filter the table so that at least one of the first or the second words belongs to the digital keyword. In other words, bigrams that do not contain any digital keyword are excluded in this stage. This results in a table of bigrams of which at least one word is a digital keyword and each row's occurrence is counted.

Finally, the table is schematised into a directed network graph using packages 'ggraph' and 'igraph' in R, which contain various algorithms to place nodes and vertices. Therefore, each algorithm was tested with a sample dataset from the first edition of *Thematic Cartography and (Geo)Visualization*. Here, two criteria are set to search for an optimal visualisation; First, there should be no overlapping nodes or vertices, and second, the nodes should be distributed as evenly as possible. The first criterion is to avoid visual congestion, and the second is to ensure all vertices are elongated enough so that the links and arrows are not occluded by labels. As a result, the force-directed algorithm ((h) in Figure 27) by Fruchterman & Reingold (1991) was chosen. Also, bigrams that occurred less than 10 times in an edition are omitted in the visualisation.

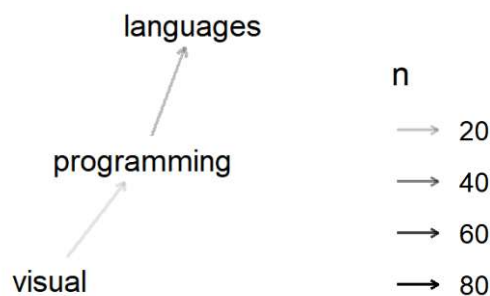


Figure 25. Excerpt from a directed network graph of an edition.



Figure 26. Excerpt from a directed network graph of a book.

In the directed network graph shown in *Figure 25*, there are two variables: direction and opacity. First, an arrow directs toward the succession of words; the word ‘programming’ is likely to come after ‘visual’, and before ‘languages’. Second, the opacity of a vertex represents the occurrence of a bigram. Here, ‘visual programming’ occurred 20~39 times, and ‘programming languages’ occurred 40~59 times. This graph is generated per edition of each textbook. From these, another directed network graph is derived by aggregating the graphs per textbook. In order to preserve the occurrence of bigrams per edition, the vertices are colour-coded (*Figure 26*).

There are two points to be noted when reading and interpreting the directed network graphs. First, any other series consisting of more than two words are not significant. For example, from *Figure 25*, it should not be assumed that ‘visual programming languages’ could be a common trigram; The analysis is only valid on a bigram level. Second, the nodes are placed considering aesthetics. Each algorithm provided by ‘ggraph’ and ‘igraph’ disperses and connects the nodes to best visualise the connectivity, not any other factors that are relevant in this research. For example, the long distance between two words does not imply a lower chance of occurrence or their semantic dissimilarity; The occurrence is represented by the opacity of a vertex, and the semantics of words are not taken into account in this visualisation method.

Results from the bigram analysis are exhibited in 4.1.3.2 and 4.2.3.2.

4 Results

This section concerns the results of analyses applied to two textbooks. Each book forms an independent subsection; Slocum et al.'s book is covered in 4.1 and Kraak & Ormeling's book in 4.2. Both subsections share the same structure and consist of three parts. The first part exhibits categories derived from the content analysis through which the evolution of digital cartography can be observed. The findings are further complemented with author interviews. The second part displays the evolution of web resources, and the third part shows the results from text mining.

It should be noted that a custom in-text citation style will be used throughout the *Results* section. This is because this section involves a large amount of text referenced from the textbooks; While the overall references of this research follow the 7th edition of the American Psychological Association (APA) citation style, this hinders the readability when used frequently as in-text citations. Therefore, a custom format for this section is designed as shown in *Table 8*. This means that references appearing in different citation styles other than the format below are secondary-referenced from the original textbooks. These are included without editing to preserve the original content of the books. However, they are not part of the analysis nor are they included in the bibliography of this research.

In this section, the source of an excerpt from a textbook will be annotated inside square brackets; The first letter(s) of the leading authors' name(s) is/are concatenated with the edition number of the book, followed by the relevant page number (*Table 8*).

Table 8. Custom in-text citation format applied to the *Results* section.

Example	APA 7 th Equivalent	Example	APA 7 th Equivalent
[S1, 198]	(Slocum, 1999, p. 198)	[K01, 25]	(Kraak & Ormeling, 1996, p. 25)
[S2, 11]	(Slocum et al., 2004, p. 11)	[K02, 4~5]	(Kraak & Ormeling, 2003, p. 4~5)
[S3, 8~10]	(Slocum et al., 2014, p. 8~10)	[K03, 167]	(Kraak & Ormeling, 2010, p. 167)
[S4, viii]	(Slocum et al., 2022, p. viii)	[K04, 120]	(Kraak & Ormeling, 2020, p. 120)

Furthermore, this section includes some inconsistencies regarding capitalisation and spelling differences. For example, in both textbooks, the words 'internet' and 'the web' are capitalised. On the other hand, APA discourages the capitalisation of those words (American Psychological Association, 2019). Therefore, they are capitalised for in-text citations and are written in lowercase when written as the main text. Also, Slocum et al.'s books follow American spelling whereas the writing of this project follows British spelling and Kraak & Ormeling's books are a mixture of the two. As a result, the entire document is written following British spelling except for in-text citations and when referring to the name of (sub)chapters of the textbooks.

A <<chapter>> of a book is written in italic, surrounded by a pair of guillemets. A <subchapter> follows the same principle but is surrounded by single guillemets. Headings below subchapters are uniformly referred to as a 'section' and are written in italic with single quotation marks.

4.1 Slocum et al.'s *Thematic Cartography and (Geo)Visualization*

Slocum et al.'s *Thematic Cartography and (Geo)Visualization* was first published in 1999, followed by subsequent editions in 2004 and 2009. Slocum explained the motivation for writing the book that “I saw a need for a textbook that covered recent cutting-edge topics in cartography” because “classic cartography textbooks such as *Elements of Cartography* (Robinson et al., 1995) and *Cartography: Thematic Map Design* (Dent, 1996) did not sufficiently cover the consequences of technological changes in cartography”. Also, even though the fourth edition was released in August 2022 and was not available during the analysis phase, the authors kindly offered the ToC and the preface. Therefore, chapter titles from the fourth edition are included in *Figure 28*.

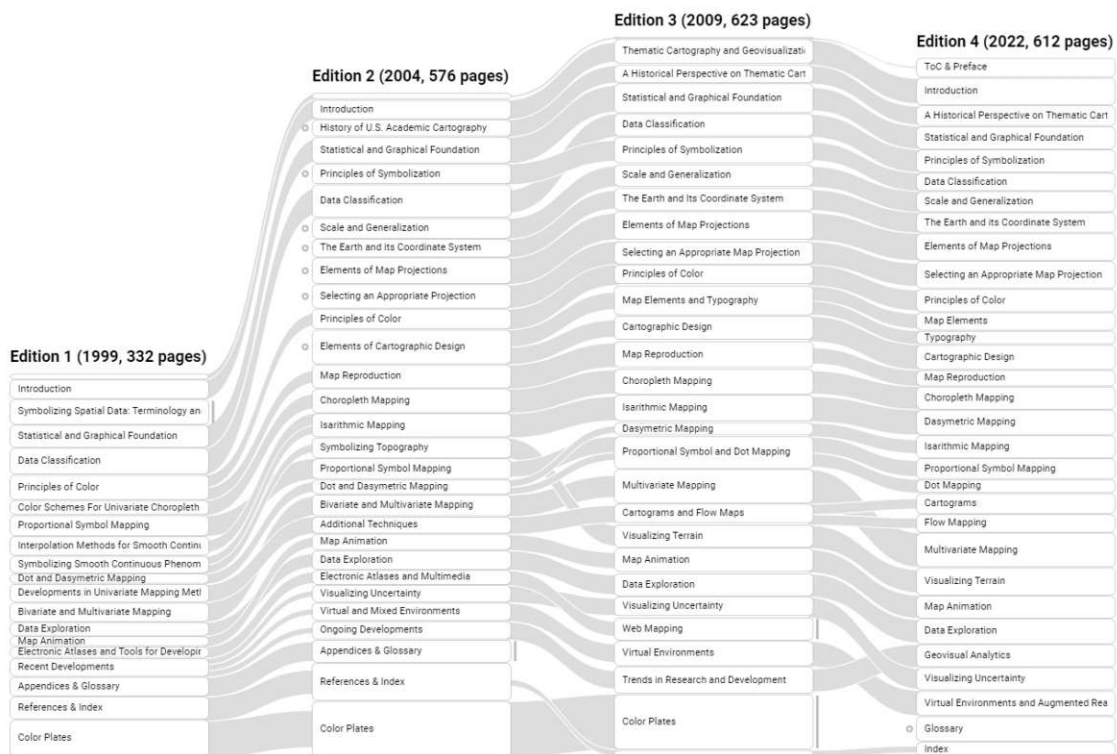


Figure 28. Flow diagram of structural changes of chapters in *Thematic Cartography and (Geo)Visualization*.

While Slocum was the only author of the first edition, three more co-authors—McMaster, Kessler, and Howard—joined in writing since the second edition. Consequently, the largest reform of the book took place in the second edition. The overall volume increased from 332 to 576 pages with the addition of various chapters about map projections and map design. Existing chapters were also largely complemented, or they diverged into multiple chapters; For example, two subchapters under the chapter «*Recent Developments*» now formed independent chapters «*Visualizing Uncertainty*» and «*Virtual and Mixed Environments*». Chapters relevant to digital technologies nearly doubled in size such as «*Map Animation*», «*Data Exploration*», and «*Electronic Atlases and Multimedia*». In the preface of the second edition, the authors attributed this

change to “the growing importance of geographic visualization” [S2, viii]. Restructuring of chapters is also a noticeable change. For example, in the first edition, two chapters were based on the characteristic of data instead of their mapping techniques: «*Interpolation Methods for Smooth Continuous Phenomena*» and «*Symbolizing Smooth Continuous Phenomena*». This was revised in the second edition by redistributing their content to «*Isarithmic Mapping*» and «*Symbolizing Topography*». Also, a significant part of the chapter «*Principles of Color*» that dealt with printing and other dissemination technologies was relocated to a new chapter «*Map Reproduction*».

Since the second edition, the overall structure has been maintained until the fourth edition, with relatively small fluctuations. Regarding the third edition, Slocum explained in the interview that “it was revised to provide greater clarity on a variety of topics and also to discuss the most recent developments in cartography”. A critical change that reflects this statement is the advent of the chapter «*Web Mapping*». The authors introduced this chapter that it “focuses on the World Wide Web and how this technology has fundamentally changed the way in which mapping activities take place” [S3, 499]. Although the chapter «*Web Mapping*» was almost newly written, it is connected with the chapter «*Electronic Atlases and Multimedia*» from the second edition since it inherited some electronic atlases. This is also supported by the authors’ description that they “consider some online atlases (such as The National Atlas of the United States) within the larger realm of Web mapping” [S3, 10]. This lineage will again be demonstrated in 4.1.1.8.

On the other hand, chapters «*Web Mapping*» and «*Trends and Research and Development*» were deleted in the fourth edition. Instead, the new chapter «*Geovisual Analytics*» was added. In the preface, the authors justified this update as below.

A dedicated chapter on Web Mapping seemed unnecessary given that we now make heavy use of web links throughout the textbook. Some of the material from Trends in Research and Development has been moved into other chapters, and we now cover recent developments within chapters where those developments have occurred. We have added a new chapter on Geovisual Analytics, reflecting the growth in this area that we noted above [S4, xxii].

In addition to the introduction of «*Geovisual Analytics*», cartograms and flow maps were now assigned with an independent chapter as well as map elements and typography.

The results of the analysis are explained in three subsections; 4.1.1 shows the result of documented changes along the editions, 4.1.2 analyses the changes in web resources, and 4.1.3 summarises the information derived from the text mining.

4.1.1 Categories from Content Analysis

Here, the results from the content analysis of changes in the first three editions of *Thematic Cartography and (Geo)Visualization* are exhibited. This section is written by interweaving the documented changes from the books and complementing some parts with author interviews. Initially, the documented changes were coded into a total of 33 codes. Then, in order to show the changes in an organic composition instead of fragments, multiple codes are grouped into one category. Each subsection from 4.1.1.1 to 4.1.1.8 is representative of a category.

Every subsection starts with an overview where the category and the involved codes are explained. Afterwards, the documented changes are exhibited in two parts: changes between the first and the second edition, and changes between the second and the third edition. Finally, a summary is provided to review the written changes and to extract meanings from them.

4.1.1.1 Map Production Methods and Colour

Overview. Map production methods in Slocum et al.'s book can be divided into three types: printed, computer-based, and internet- and web-based. While it could be contested that printed maps are not strictly digital, it should be noted that printing hardware has undergone changes in accordance with the development of computers and that a lot of early digital maps had a printed version as their precursor. As a result, technologies related to printing are also an inevitable component in understanding the evolution of digital maps and are therefore included in the analysis. Also, the representation of colours is a topic that has gone through major changes along with the development of hardware and software involved in map production. Therefore, map production and colour are bound to one category. Codes found in this category are *colour*, *hardware*, *display*, *printing*, *software*, and *web*.

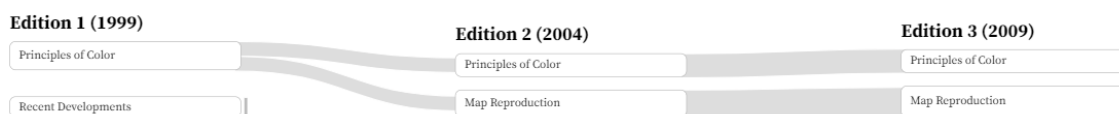


Figure 29. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* that are involved in the formation of the category *Map Production Methods and Colour*.

Changes between the first and the second edition. Most of the updates observed in the second edition concern printing methods and associated hardware; A lot of them are deletions of the text describing technologies that are deprecated or have become too common to be covered in the textbook. For example, it was described in the first edition that “Both printers and plotters are used to create hard-copy maps” [S1, 88], and they mentioned the ‘Hewlett Packard series’ as an example that simulates traditional pen-and-ink cartography. However, they also predicted that “It appears that this technology [plotter] will soon disappear” [S1, 89]. This proved to be true and the entire part was reduced to a single footnote that “Certain devices, such as pen plotters, do not require that data be converted into a raster image for printing, but these devices are less widely used today” [S2, 231].

Also, it was written in the first edition that “Four types of color printers are commonly available in the marketplace today: ink-jet, laser, thermal-wax transfer, and dye-sublimation” [S1, 94]. Here, the four printer types were treated equally under the section ‘Printers’. When it comes to the second edition, they were reformatted into the subchapter ‘Low-volume Print Reproduction’, where laser and ink-jet were each assigned with their own section but thermal-wax and dye-sublimation were joined under ‘Additional Printing Methods’. They provided context concerning

this decision; “Laser printers have proven to be practical and reliable for low-volume reproduction” [S2, 237], and “The quality of ink-jet printing has advanced dramatically over the past decade; equally dramatic has been the drop in prices for these devices” [S2, 238].

Slocum also included an anecdote about colour printing in the first edition of the book; “When trying to create maps using Brewer’s (1989) method (...), the colors produced by an inexpensive ink-jet printer were quite different from those appearing in her published work. But when the maps were printed on a more expensive dye-sublimation printer, color differences were negligible” [S1, 111]. This part—which demonstrated the difficulty of balancing the cost and the quality of colour-printed maps—was deleted in the second edition.

Contrary to printing methods that were downsized, computer-based dissemination was largely complemented in the second edition. Slocum defined the reproduction of maps in digital form in comparison with printed maps; “It consists primarily of digital file duplication, as performed by a computer’s operating system. Reproduced digital maps can be stored on media such as the CD, DVD, and Zip disk” [S2, 245], and mentioned the local area network (LAN) as a way of distributing electronic maps. They also introduced various file formats in which maps can be exported such as EPS, PDF, JPEG, GIF, TIFF, and BMP.

Furthermore, the new subchapter *Internet and Web-based Dissemination* was added in the second edition. Slocum assessed that the internet plays an important role in map dissemination because “It is capable of distributing static and interactive maps to large numbers of users almost instantaneously” [S2, 245]. They also distinguished internet- and web-based methods; Internet-based dissemination is exchanging files through an online server such as File Transfer Protocol (FTP), while the web is a subset of the internet but with more flexibility and interactivity. They introduced key concepts and file formats including HTML, Flash, XML, and SVG and covered Java and JavaScript as a way of expanding the functionalities of the web.

In the first edition, eased possibilities of producing colour maps were selected as one of the consequences of technological changes in cartography. The democratisation of colour maps manifested itself through both media: hardcopy [printed] and softcopy [digital] maps. Here, Slocum wrote that “This proliferation has been driven largely by increasing capabilities (and decreasing costs) of graphic display systems and color printers, and the ease of transferring information digitally via the Internet” [S1, 10]. And they suggested the readers think about how affordable hardware has become.

Consider the color capability that can now be achieved in our own homes: for under \$1,500, personal computers are available that will display 256 colors at one time out of a palette of 16.8 million, and color printers of good graphic quality are now available for under \$500 [S1, 10].

However, in the second edition, the marvel at the new possibilities of representing colours was replaced with new excitement over the internet and web (*Table 9*). Also, whereas the first edition included recent research on colours concerning the optimal application of colour schemes in the chapter *Recent Developments*, this was deleted in the second edition and they mentioned other novel mapping techniques and related digital technologies.

Table 9. Comparison of descriptions regarding the sixth consequence of technological changes in *Thematic Cartography and (Geo)Visualization*.

[S1, 10]	[S2, 10]	[S3, 10] <i>Unchanged.</i>
A sixth consequence of technological change is the proliferation of color maps.	A sixth consequence of technological change is the ability to access maps and related information via the World Wide Web (WWW, or simply the Web), which is part of the larger Internet.	A sixth consequence of technological change is the ability to access maps and related information via the World Wide Web (WWW, or simply the Web), which is part of the larger Internet.

In terms of hardware that is responsible for representing colours, the most evident updates are found in describing the transition from CRT (cathode ray tube) to LCD (liquid crystal display) in the chapter «*Principles of Color*». In the first edition, although Slocum acknowledged that CRT is likely to be replaced in near future because of its disadvantages, detailed information about LCDs was intentionally left out. They annotated two reasons behind this decision; “First, CRTs are still the dominant technology for personal computers, (...) and second, it is unclear which of these [flat-panel] displays will ultimately dominate the marketplace” [S1, 93]. This entire part was deleted in the second edition and they wrote that LCDs are starting to be adopted in desktop computers—which CRTs used to dominate—not only in laptop computers and computer projection devices. From here, LCD was given its own section under the subchapter «*Hardware Considerations in Producing Soft-copy Color Maps*».

When it comes to software, changes regarding colours can be found in parts that handle colour management systems (CMS). CMS are software that “provide consistency and predictability of color by identifying differences in gamuts and correcting the variations in color introduced by each device” [S2, 237] to ensure that the colours on a graphic display match with those produced by a printing device. They were first covered in the chapter «*Principles of Color*», but from the second edition, the content was relocated to a new chapter «*Map Reproduction*» which presents topics related to both printed and digital maps.

In the first edition, it was mentioned that CMS are either bound to specific software applications or a computer operating system. In the second edition, they revised this part that “working at the operating system level is the trend today” [S2, 237] and deleted the earlier examples that worked only with specific software. The continuities of CMS examples are shown in *Table 10*.

In the second edition, the formation of the International Color Consortium (ICC) was introduced. Their achievements are summarised as “developing a standard for a vendor-neutral, cross-platform color profile called ICC color profile, and establishing architectures for the Color Management Module (CMM) that performs color corrections as well as the color management solution that facilitates the color management process at the operating system level” [S2, 237]. Examples of ICC colour management solutions are the two cases of CMS that have made it to the second edition: Apple’s ColorSync and Microsoft’s ICM2.

Furthermore, the first edition annotated that CMS's quality, ease of use, and cost vary since they are a recent phenomenon. Referencing the reviews by Sugihara (1995) and Hilliard (1995) who found out that the systems bundled with specific applications are the most sophisticated, Slocum suggested that "reviews of software by cartographers will be necessary to determine which is most appropriate for mapping purposes" [S1, 95]. This entire paragraph was deleted in the second edition since the CMS that are bound to specific software applications were no longer covered and new standards were established by the ICC.

Changes between the second and the third edition. In comparison to the second edition, fewer changes are observed in the third edition in terms of map production methods and colour. There was some restructuring at the subchapter level and deletions of obsolete technologies. For example, the subchapter that used to cover the low-volume printing disappeared. Instead, the authors provided an in-depth explanation of offset lithography—a single dominant method for high-volume printing. In the interview, Slocum explained this deletion of the subchapter; "Since the technology of such printers is constantly changing, we felt that it was unimportant to focus on particular characteristics of printers".

A remarkable change is the downsizing of descriptions regarding television, which used to be attentively documented until the second edition. Television was a highly relevant technology before microcomputers because "it was the first medium where maps were disseminated widely in non-paper form" [S2, 245]. While they mentioned in the second edition that the relatively low resolution of television was partially resolved in the second edition with the advent of digital television, the entire part was discontinued in the third edition. This is not an individual case; A lot of descriptions about evolving hardware and display technologies of the time (e.g., lowered costs, set-top boxes) have been either deleted or reduced to a few sentences in the third edition.

In terms of display technologies, the authors became certain that LCDs are now the standard display technology. For example, the sentence in the second edition "Today, CRTs are the dominant technology for the production of soft-copy maps, but LCDs are becoming quite common, and might eventually displace CRTs as the dominant technology" [S2, 197] was replaced with "Traditionally, CRTs were the dominant technology for graphics displays, but LCDs have recently become more common" [S3, 208]. Below is a diagnosis by the authors concerning the rapid development in display technologies from the chapter <<*Virtual and Mixed Environments*>>, which was also deleted in the third edition.

With the exception of the desktop display, high-quality versions of these displays currently are beyond the reach of the typical geographer. Not only does the hardware for the display tend to be expensive, but high-end computers (e.g., Silicon Graphics) generally are required to power the display and specialized software is necessary. Although this is true today, keep in mind that technological advances are continually being made and the cost of computer-based technology is dropping, so we anticipate that many of these displays will be more generally available within a few years [S2, 437].

Summary. In Slocum et al.'s books, the evolution in map production methods and colour can be observed along with the life cycle of digital technologies. A lot of descriptions about printing technologies (e.g., ink-jet, laser, thermal-wax transfer, and dye-sublimation) gradually

decreased as well as television which used to be the most dominant dissemination technology. On the other hand, the computer-based dissemination was largely complemented in the second edition, followed by the addition of the internet and web-based dissemination in the third edition. The advent of the web also replaced the excitement about the proliferation of colour maps. In terms of colour displays, CRT lost its dominance to LCD over time, and the price issue of high-end displays raised by authors was also resolved. Furthermore, the evolution of CMS gives an overview of how software-hardware dependency has gradually been alleviated along with the international effort (i.e., the formation of the ICC) to standardise the associated software.

Table 10. Continuities of CMS examples in *Thematic Cartography and (Geo)Visualization*.

1st ed. (1999)	2nd ed. (2004)	3rd ed. (2014)
Kodak Precision CMS (in Adobe System Page-Maker 6.0)		
EfiColor XTension (in Quark-Xpress 3.2)		
PANTONE Personal Color Calibrator (for Macintosh and Windows 95)		
Apple's ColorSync (for Macintosh)		
Microsoft's ICM (for Windows)	ICM2	

- CMS that work in specific software applications
- CMS that work in specific operating systems

4.1.1.2 GIS and Map Design

Overview. Slocum et al. have emphasised the rapid development of GIS and their potential since the second edition. The increased relevance of GIS can be witnessed in two blocks: early chapters where they introduce components and the history of cartography, and later chapters which deal with map design. And since some map design procedures have shown meaningful evolution when explained in the context of GIS, the map design is categorised together with GIS. Codes found in this category are *software*, *GIS*, *design*, *education*, and *remote sensing*.

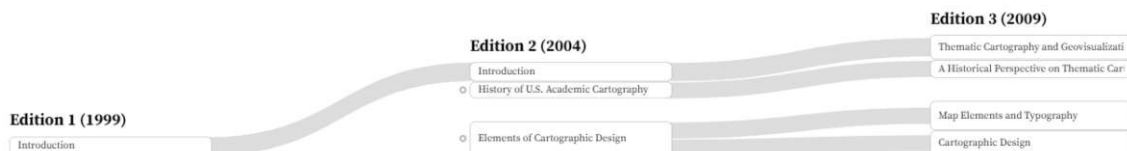


Figure 30. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* that are involved in the formation of the category *GIS and Map Design*.

Changes between the first and the second edition. In the chapter <<Introduction>> of the second edition, the authors listed the recent development in three realms that cartographers should be aware of. These are GIS, remote sensing, and quantitative methods, each of which is elaborated on under the new subchapter <<Related Techniques>. Here, GIS are mentioned as a more sophisticated and advanced form of software compared to earlier ones that primarily focused on thematic mapping. They wrote that “Today, GIS software typically has considerable thematic mapping capability, in addition to its inherent spatial analysis capabilities, and its spatial analysis capabilities also enable GIS to handle more sophisticated mapping problems than traditional thematic mapping software” [S2, 13].

In the new chapter <<History of U.S. Academic Cartography>>, the authors stated that GIS are a key factor that brought a paradigm shift in the academia of cartography. Here, they defined four major periods of the US academic cartography: The Incipient Period, Post-World War II Graduate Education Centers of Excellence, Diffusion of Cartographic Programmes in Geography Departments, and The Transition Period. They identified the last period as the past 10 years [from 2004] where “the intellectual landscape of cartography has changed significantly” [S2, 29]. Since GIS permeated into cartography and their applications became ubiquitous in society, the distinction of the field cartography was weakened and diluted, which the authors described as “an erosion of the professional base of cartography” [S2, 29]. With a focus on university teaching environments, they diagnosed that “cartography has become increasingly integrated within GIS curricula, which resulted in fewer academic positions in cartography, fewer students educated as thoroughly in thematic cartography, and growth in geographic visualization” [S2, 18].

The evolution of GIS is also observed in chapters that are related to map design. In the second edition, improvement in map design abilities of GIS was pointed out. For example, in the chapter <<Elements of Cartographic Design>>, they mentioned that “graphic design applications such as Macromedia FreeHand and Adobe Illustrator typically provide the greatest control over graphics and type, but recent advances in the design capabilities of GISs have narrowed the gap, allowing you to produce high-quality maps in a user-friendly GIS environment” [S2, 200], hinting that the typical division of ‘map production with GIS and finishing it up with graphic design software’ could be overcome in near future.

Changes between the second and the third edition. In the chapter <<A Historic Perspective on Thematic Cartography>> of the third edition, the ‘Transition Period’ was renamed to ‘Integration with GIScience’. While the overall tone concerning the integration of cartography and GIS remained unchanged, the authors have come to a more optimistic conclusion. For example, in the second edition, they set a rivalry between cartography and geovisualisation that the integration of cartography and GIS resulted in “fewer academic positions in cartography, fewer students educated as thoroughly in thematic cartography, and growth in geographic visualization” [S2, 31]. In the third edition, this was replaced with the following sentence: “Since the 1980s, cartography has become more integrated with GIScience, though recently there has been a resurgence in the focus on cartography, geovisualization, and representation” [S3, 23].

Aside from the educational context, the history of GIS was complemented with the introduction of ‘critical GIS’. The authors defined critical GIS as “an examination of the social and ethical implications of the broader field of GIScience” [S3, 17] and showed two studies as an example. These are; a feminist visualisation by Mei-Po Kwan (2002) who mapped the space-time paths of Black women in Portland, and an alternative dasymetric map suggested by Crampton (2004) which considers ancillary information such as land use and land cover.

Regarding the application of GIS and map design software, the changes are observed in two chapters «*Map Elements and Typography*» and «*Cartographic Design*», which ramified from «*Elements of Cartographic Design*» of the second edition. In the chapter «*Map Elements and Typography*», three software applications were added: Label-EZ, Maplex, and TypeBrewer. Label-EZ and Maplex are labelling software designed for optimal placement of map labels and TypeBrewer is for selecting appropriate typefaces on a map. In describing the labelling software, their interoperability with GIS was emphasised. For example, it was illustrated that users can import a series of map layers derived from GIS, place labels in Label-EZ, and transfer the layers back to the GIS application. Similarly, they mentioned that “Maplex is designed for use primarily with ESRI GIS applications”, and that “there is also an extension for ArcGIS that focuses exclusively on automated type placement” [S3, 235].

In the chapter «*Cartographic Design*», they introduced a new subchapter «*Case Study: Real Estate Site Suitability Map*». It consists of detailed instructions that readers can follow step-by-step using ESRI ArcMap and Adobe Illustrator to create a real estate site suitability map. The authors described the goal of this subchapter as to “create an efficient, attractive map that represents the relative suitabilities of residential building parcels (thematic symbols), together with appropriate base information” [S3, 249].

Summary. The increased relevance of GIS has been reflected in Slocum et al.’s book since the second edition. This can be seen from two aspects: the historical/educational viewpoint and the map design aspect. Whereas the blurred border of traditional cartography due to the rapid development of GIS was largely seen as a crisis in the second edition, the authors started to affirm the integration in the third edition. Also, the inclusion of the concept of ‘critical GIS’ mirrors the increased complexity and ubiquitousness of GIS, reminding readers of the need for reflection and responsibility of cartographers in the digital era. From a map design point of view, while GIS are established as a standard tool for generating maps, the advantages of using map design software along with GIS have been steadily emphasised. Their increased functionalities and interoperability resulted in the new subchapter in the third edition where readers are invited to create a map using ArcMap and Adobe Illustrator.

4.1.1.3 Generalisation

Overview. The first edition did not have a chapter dedicated to generalisation, and only one generalisation technique from Langford and Unwin (1994) was mentioned in the chapter «*Dot and Dasymetric Mapping*». Then, the second edition introduced the chapter «*Scale and Generalization*» which is divided into five subchapters. Among these, «*Models of Generalization*» and

«*An Example of Generalization*» went through major updates in the third edition. Since the former subchapter—introducing conceptual models behind the cartographic generalisation process—is largely theoretical, this subsection will focus on the latter which is more associated with digital technologies. Codes found in this category are *generalisation*, *software*, and *web*.



Figure 31. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* that are involved in the formation of the category *Generalisation*.

Changes between the first and the second edition. There is no corresponding chapter in the first edition. Also, no significant changes are detected regarding the generalisation technique from Langford and Unwin (1994) in the chapter «*Dot and Dasymetric Mapping*».

Changes between the second and the third edition. In the subchapter «*An Example of Generalization*», the second edition used raw TIGER vector data for the Tampa-St. Petersburg area of Florida. On this dataset, Slocum et al. applied the Visvalingam algorithm which “uses an areal tolerance to select critical points and is considered to be robust in maintaining the original character of the line” [S2, 118]. After this, they wrote that “It is still difficult to solve generalization problems with off-the-shelf software due to the limited capability of the algorithms and complexity of the databases” [S2, 118] and that the National Historical Geographic Information System (NHGIS) project at the University of Minnesota is underway to provide multiple scale versions of census data to enable users to select the scale most appropriate for their use.

This project was completed by the time when the third edition was published, and the NHGIS website replaced the example from the second edition. NHGIS produced a multi-scale database that utilises different levels of generalisation. According to Slocum et al., there are three levels: “a scale of 1:150,000, which is the original scale of the TIGER files and the digital line graphs (DLGs), and two additional databases, one at 1:500,000 and one at 1:1,000,000” [S3, 111]. They also briefly introduced the processing procedure of the NHGIS generalisation system consisting of two parts. The first part “eliminates small areas (e.g., islands, small parts of multipart features, and slivers caused by historical boundary changes)” [S3, 121], and the second part generalises boundaries using the line simplification algorithms.

Moreover, the third edition added a new subchapter «*MapShaper: A Free Web-based Generalization Service*», a website developed by Harrower and Bloch. The authors introduced five characteristics of MapShaper; “(1) Users can automatically construct topology; (2) Users can specify critical points; (3) Users can easily change the level of generalization; (4) Users can see the results of generalization immediately; and (5) Users can see the effects of various simplification algorithms with a mere click of the mouse” [S3, 122]. Here, they juxtaposed two images to compare the effects of Douglas-Peucker and modified Visvalingam-Whyatt simplification algorithms. Aside from encouraging the readers to experiment with the website by themselves, Slocum et al.

also asked them to contemplate the purpose of the generalisation to evaluate and the effects of simplification, instead of solely relying on the automated process.

Students of modern cartography should realize that much of what we know today about the generalization process is the result of what cartographers over the centuries learned through manual cartography, where maps were made with pen and ink. When changing scale, cartographers would carefully study a feature at the original scale, and then decide how the feature needed to be modified for the new scale. Did it need to be moved (displaced) to avoid colliding with another feature? Did it need to be simplified and/or smoothed? A series of techniques were developed over hundreds of years to help cartographers generalize maps and map features [S3, 112].

Summary. Through the changes between the second and the third edition of Slocum et al.'s book, it can be learned that the automation of generalisation used to be a difficult task. This was due to the limited capability of the algorithms and the complexity of the databases. These issues were largely resolved in the time of the third edition, which is demonstrated by examples of the National Historical Geographic Information System (NHGIS) project and a free web-based generalisation service 'MapShaper'. Furthermore, despite the convenience of automated processes, the authors emphasised that it is important to understand the underlying principles and utilise appropriate generalisation methods.

4.1.1.4 Multimedia

Overview. The topic of multimedia has frequently changed its main residence. In the first edition, it was introduced as a tool for conveying information in electronic atlases and was also covered under the chapter <<Recent Developments>>. In the second edition, it obtained its own subchapter under the chapter <<Electronic Atlases and Multimedia>> and appeared in the new chapter <<Virtual and Mixed Environments>> as well. In the third edition, it was covered in three chapters: <<Web Mapping>>, <<Virtual Environments>>, and <<Trends in Research and Development>>. Codes found in this category are *multimedia*, *software*, *accessibility*, *web*, *atlas*, and *user*.



Figure 32. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* that are involved in the formation of the category *Multimedia*.

Changes between the first and the second edition. In the first edition, the concept of multimedia equated to the 'combination of sounds with visuals' and was covered in the subchapter <<Using Sound to Represent Spatial Data>> in the chapter <<Recent Developments>>. Here, four examples were introduced where sounds are incorporated to complement animated maps or to alternatively represent map information. While these examples barely changed in the second edition, the authors added a paragraph to justify how using sounds can be helpful for the audience with

and without disabilities. They identified four reasons; “First, sound can be useful when the work environment requires that vision be used for other things besides map interpretation; Second, sound can assist in interpreting a map when visual interpretation is not clear; Third, sound has the potential for use in multivariate mapping, as different characteristics of sound (e.g., loudness and pitch) can be assigned to different attributes; Finally, sound might be useful as a means of orienting the user and in 3D displays where visual information might be occluded” [S2, 453].

Also, the second edition started to incorporate the touch (haptic) along with the visuals and sounds. Here, ‘Haptic Soundscapes’ was introduced as an example where the researchers experimented with different vibrations using a force-feedback mouse, combined with sound variables such as pitch and volume to show the population of a county.

In addition, the second edition coined the term ‘multimedia systems’ and assigned them with a separate subchapter in *«Electronic Atlases and Multimedia»*. They were defined as “Mapping systems that place much greater emphasis on multimedia aspects [than the collection of maps stored in the atlas]” [S2, 416]. Despite the presence of its subchapter, multimedia systems were not thoroughly explained, and the authors displayed sporadic examples that suggest the potential of using multimedia in neighbouring fields such as urban planning and natural resources.

Changes between the second and the third edition. The third edition further extended the definition of multimedia. From here, the authors started using the term ‘other senses’ instead of ‘sounds’ and covered additional studies that experimented with sound, touch, and even smell. Even though they did not come up with any actual maps where users can use the sense of smell, they indicated that the future is near, that “hardware and software have been developed for generating smells”, and that “it will be interesting to see what role smell plays in cartography” [S3, 549]. As a result, real-world examples of multimedia maps still revolved around sound and haptics. For example, they updated the description of Haptic Soundscapes with additional feedback such as “One of the difficulties of interpreting Golledge et al.’s work is that many of us do not have access to haptic interface devices; we hope that this research will lead to an increased proliferation of the hardware necessary for haptic interfaces” [S3, 549].

Concerning the use of sounds, they added two software applications: ‘iSonic’ and the ‘Cybercartographic Atlas of Antarctica’. iSonic is a software application where geographical names are read out loud and their associated values are represented in violin pitches. The authors admired the developers’ effort of grasping the need of users; Here, sounds play from left to right and from top to bottom instead of in alphabetical order so that listeners can better understand the spatial distribution. In the Cybercartographic Atlas of Antarctica, real sounds from natural languages are used instead of abstract sounds that were widely applied to represent the thematic values. The authors described its key function; “When a user places the mouse over an uncontested territory, the user hears a news report only from the country claiming this region, whereas over a contested territory, the user hears multiple news reports simultaneously” [S3, 547].

As the semantics of multimedia have broadened by including various senses, the authors started to contemplate the term ‘geovisualisation’ as well. They introduced the proposal by Fraser Taylor—the developer of the Cybercartographic Atlas of Antarctica—that the term ‘cyber-

cartography’ would be a more appropriate term to embrace the multisensory, interactive, and web-based characteristics of maps that are on the rise. While they stated that this term is not widely adopted by cartographers, they agreed with Taylor’s point that “It seems clear that ultimately we will need a term that is more inclusive than geovisualization” [S3, 14].

Also, the entire subchapter <Multimedia Systems>—introduced in the second edition under <<Electronic Atlases and Multimedia>>—was deleted in the third edition. To the question regarding the reason for deletion, Slocum answered that “We didn’t see the term ‘multimedia systems’ used heavily in the literature, and so chose not to focus on it in the third edition”.

Summary. In Slocum et al.’s book, the meaning of multimedia has been broadened along with the editions; from sound and visuals to other senses, including touch and smell. Cartographers experimented mainly with sounds and haptics to represent the thematic values in combination with visuals, which was expected to benefit readers with and without disabilities. Also, in order to embrace the new trend of multimedia functions used in maps, some have suggested alternative terms that could replace ‘geovisualisation’.

4.1.1.5 Thematic and Topographic Visualisation

Overview. Changes in thematic and topographic visualisation are observed throughout various chapters; Updates in thematic visualisation are represented by new visualisation and quantitative methods, and updates in topographic visualisation concern 3D visualisation and modelling. Although not all examples covered in Slocum et al.’s book are inherently digital, they are still mostly accompanied by software that can assist in the visualisation stage. Therefore, this subsection focuses on the changes in visualisation techniques and examples that are associated with computer-based systems. Codes found in this category are *visualisation*, *3D*, *statistics*, *software*, *web*, and *hardware*.

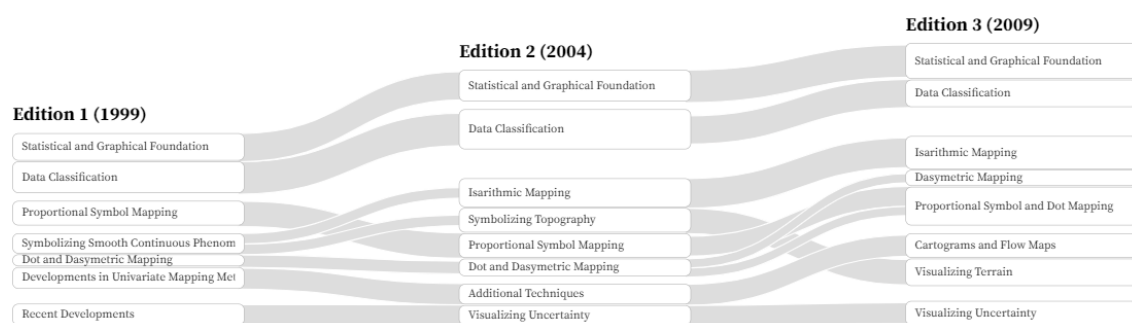


Figure 33. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* that are involved in the formation of the category *Thematic and Topographic Visualisation*.

Changes between the first and the second edition. In the chapter <<Statistical and Graphical Foundation>> of the second edition, the authors added four points to consider when using correlation and regression methods. Among these, the third point is relevant to digital cartography, that specific regression techniques are developed to handle spatially autocorrelated data.

Spatial autocorrelation—the tendency of thematic values of a location in geospatial data to be similar or dissimilar to that of neighbouring locations—is a key concept in chapters that handle statistics and visualisation. Here, they mentioned a software package ‘SpaceStat’ which is designed to deal with geospatial data and spatial statistics.

As more refined visualisation techniques and statistically complex methods were developed, realising them in computer-based systems became computationally more expensive, which was a point that had to be made at the time of the first edition. For example, after covering triangulation, inverse distance, and kriging as typical interpolation methods, they described the possible computational stress that can be caused by geospatial analyses.

Using a 50 megahertz 486-based microcomputer, Declercq computed grid point estimates for a 121 * 361 grid in less than 5 minutes using either linear triangulation or inverse distance, but almost an hour was required for some kriging computations. A smoothed triangulation approach (using a quadratic-spline interpolation) was clearly the slowest, requiring almost 12 hours [S1, 147].

However, at the same time, the authors were aware of the rapid growth of hardware technologies. They annotated that “microcomputers are becoming substantially faster, and that for large data sets, one can simply let the computer run while doing something else” [S1, 147]. In the second edition, this sentence was replaced with “The usefulness of such studies [by Declercq], however, is problematic, as processor speeds, storage capacity, and the amount of RAM increases exponentially each year” [S2, 283].

Regarding topographic visualisation, there was a major restructuring in the second edition; In the first edition, isarithmic mapping and terrain visualisation was covered together in the chapter <<Symbolizing Smooth Continuous Phenomena>>. In the second edition, this was separated into <<Isarithmic Mapping>> and <<Symbolizing Topography>>. While each chapter inherited the content from the first edition, they were also largely complemented with additional subchapters.

In the chapter <<Symbolizing Topography>>, a software package ‘LandSerf’ was added, which lets users utilise more parameters other than aspect and slope when visualising terrains. Here, the authors experimented with an example of the Columbia River area in Washington State and shared the intermediate results in the colour plate. Also, three methods were added in the new subchapter <<Oblique Views: Block Diagrams, Panoramas, and Draped Images>>. Block diagrams and panoramas are techniques that have existed long before the emergence of computers, therefore, their descriptions have little connection to GIS or software. On the other hand, ‘draped images’ are relevant to digital technologies, which are maps that are made by draping remotely sensed images on a 3D map of elevation. Here, the authors wrote that “The popularity of this approach is indicated by the fact that several hundred packages have been developed for 3D visualization of the terrain” [S2, 307].

Furthermore, the authors went beyond the concept of 3D visualisation in computer software and proposed the possibility of generating actual 3D terrain models.

Although physical models have been around since the time of Alexander the Great, they have been used infrequently by cartographers because of the difficulty of obtaining terrain information, the cost and time of creating them manually, and the difficulty of disposing of them. Today, such limitations no

longer exist, as digital data are readily available, several technologies have been developed for converting the digital data into a solid model, and disposal is less of an issue given the range of available technologies (Caldwell 2001) [S2, 307].

Here, two examples were introduced: Solid Terrain Modeling (STM) created for National Geographic's Explorer's Hall in Washington DC, and the physical model of Death Valley. STM is generated by cutting the polyurethane foam based on DEM data, and then printing satellite images on it. In the case of Death Valley, the models are created using the San Diego Supercomputer Center's Tele-Manufacturing Facility, by laminating individual pieces of paper.

Changes between the second and the third edition. Whereas there were no significant changes in topographic visualisation in the third edition, thematic mapping was substantially updated, which involved multiple chapters: <<Data Classification>>, <<Isarithmic Mapping>>, <<Dot and Dasymetric Mapping>>, <<Proportional Symbol Mapping>>, and <<Cartograms and Flow Maps>>.

In the chapter <<Data Classification>>, they introduced a new subchapter <<Using Multiple Criteria to Determine Class Intervals>>. This subchapter deals with the advanced question of 'how can we combine multiple criteria to create a choropleth map that is visually effective but also statistically correct?'. Here, they introduced 'ChoroWare', a software application that produces multiple choropleth maps so that users can examine the trade-offs between the criteria and find a set of class intervals that they find the most appropriate.

In the chapter <<Isarithmic Mapping>>, the entire paragraph explaining Declercq's experiment—where the computation took almost 12 hours with the smoothed triangulation approach—was deleted. Now, the only mention about the execution time was that smoothed triangulation methods can take longer than inverse-distance or kriging.

In the chapter <<Dot and Dasymetric Mapping>>, they introduced 'LandScan'—a global database that provides population estimates for grid cells of approximately 1 square kilometre, by utilising various information such as roads, slopes, populated places, and night-time lights. Here, the authors explained various data sources of ancillary layers and how they can function as an indicator of population. LandScan was first released in 2001, which suggests that it could have been included in the earlier edition. Therefore, it was inquired during the interview why it was only added to the third edition. To this question, Slocum answered that "We chose to create a chapter focusing on just dasymetric mapping for the third edition, and thus we were looking for material that could help fill out that chapter".

In the chapter <<Proportional Symbol Mapping>>, they updated the part where they explained why circles are the most preferred form of symbol. In the first and the second edition, they listed five reasons; "(1) Circles are visually stable; (2) Users prefer circles over other geometric symbols; (3) Circles conserve map space; (4) When constructing circles, it is easy to determine appropriate relations between data and circle size; and (5) Circles are easy to construct" [S1, 121]; [S2, 313]. Then, they annotated that the latter two have largely disappeared with the rise of digital mapping because "formulas for calculating the sizes of geometric symbols can now be implemented in spreadsheets, and software can be used to automatically construct geometric symbols" [S1, 121];

[S2, 313]. In the third edition, the latter two reasons were deleted as well as the annotation, leaving only three reasons why circles are the most preferred.

The chapter <<*Cartograms and Flow Maps*>> was complemented with additional studies and software applications. For example, they introduced a novel approach in cartograms by Gastner and Newman (2004, 2005) that retains correct topological relationships between enumeration units and can be executed quickly. After demonstrating this approach, the authors introduced ‘The Worldmapper’—a website with a variety of cartograms produced by using Gastner and Newman’s method.

In the case of flow maps, they introduced a new software application by Phan et al. (2005) which can handle typical problems in flow map design (e.g., merging arrows and shifting the base map to make room for the flow lines). Similarly, the authors again attached a website where readers can see the maps produced by using their algorithms. Furthermore, they provided a recent update concerning the flow map software developed by Tobler (1987). Considered one of the earliest software displaying migration flows, it depicts “one-way migration from a particular U.S. state to others by arrows of varying width” [S2, 364]. In the third edition, a new paragraph was added to the original description that Jones (2003) redeveloped this software so that it works in a Windows-based environment.

Summary. As novel statistic techniques emerged, hardware capacities were improved and various software applications were released correspondingly. In topographic mapping, changes are mainly found in the second edition, which concerns how to display 3D landscapes and even generate a 3D physical model. In thematic mapping, changes took place in various topics and techniques. Software applications such as SpaceStat and ChoroWare are a reminder that there have been programmes that were explicitly developed for cartographers to meet specific needs, despite the omnipotence of GIS nowadays. A sophisticated geodatabase such as LandScan shows how technologies associated with cartography (e.g., remote sensing, information technology, web) can assist in refining and improving quantitative mapping methods. Furthermore, changes observed in cartograms and flow maps show a progressivist timeline of a technology where a scientific method experiences a huge leap; overcoming earlier deficits, its results are distributed on the internet, and further computer-based improvements are suggested.

4.1.1.6 *Virtual Environments*

Overview. In the first edition of Slocum et al.’s book, virtual environments (VE) were covered in the subchapter <<*Virtual Reality and Virtual Realism*> in <<*Recent Developments*>>. Then, it was promoted to a separate chapter in the second edition. Although VE are not yet incorporated in map production, their technology is inherently digital and shows the evolution in various neighbouring fields of cartography. Thus, they formed an independent category. Codes found in this category are *virtual environments, web, hardware, 3D, display, software, and user*.



Figure 34. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* that are involved in the formation of the category *Virtual Environments*.

Changes between the first and the second edition. In the first edition, VE were covered in the subchapter *Virtual Reality and Virtual Realism* under the chapter *Recent Developments*. While a mixture of terms floated around to refer to this up-and-coming technology (e.g., virtual world, virtual realism, and virtual simulation), they introduced an example and related software. These were: Virtual Los Angeles, CLRview Software, and Emaps Package. Only Virtual Los Angeles was continued until the third edition and the rest disappeared after the first edition.

Virtual Los Angeles is an example of virtual cities which is characterised by a highly detailed visual representation of buildings. In the first edition—when the project was still in the early development phase, its blueprint was briefly mentioned and that it will be developed using the proprietary ‘uSim’ software. In the second edition, this was broadened to a variety of software applications: AutoCAD, MultiGen/Creator, IRIS Performer, OpenGL, and ArcView.

CLRview and Emaps are software applications developed in the early 1990s to create and edit realistic 3D features, which is an essential prerequisite in virtual reality. Typical shortcomings of early software were mentioned in the first edition; “CLRview currently requires a sophisticated hardware environment (Silicon Graphics), and Emaps required the use of additional software running on a sophisticated graphics workstation” [S1, 252]. Moreover, the current shift in the development environment of GIS was mentioned that Emaps is not supported anymore but most of its functionalities have become available on ArcView and Java.

The second edition—in addition to giving VE a separate chapter—openly stated their increased relevance. In the chapter *Introduction*, they listed virtual environments as well as augmented reality as one of the consequences of technological changes in cartography (Table 11).

Table 11. Comparison of descriptions regarding the fifth consequence of technological changes in *Thematic Cartography and (Geo)Visualization*.

[S1, 9]	[S2, 1 & 9]	[S3, 1 & 10] <i>Unchanged.</i>
A fifth consequence of technological change is the ability to access maps and related information via the Internet (or information superhighway).	A fifth consequence of technological change is the ability to create realistic representations of the environment (virtual environments or virtual reality) and the related notion of augmented reality (enhancing our view of the real world through computer-based information).	A fifth consequence of technological change is the ability to create realistic representations of the environment (virtual environments or virtual reality) and the related notion of augmented reality (enhancing our view of the real world through computer-based information).

While the chapter «*Virtual and Mixed Environments*» of the second edition had some connections to the first edition, almost all of its contents were newly written. In the first subchapter «*Defining Virtual and Mixed Environments*», the authors included a wide range of discussions about various terms surrounding VE along with some earliest research. Here, they concluded that the term ‘virtual environment’ is more appropriate than ‘virtual reality’. Specifically, they defined VE as “A 3D computer-based simulation of a real or imagined environment that users are able to navigate through and interact with” [S2, 435].

Although it makes sense to simulate the tangible real world, there are instances in which we might need to display something intangible. (...) As a result, researchers have chosen to use the term virtual environment to encompass both tangible and intangible simulations [S2, 435].

The second edition assorted the VE examples into five sections: ‘*Virtual Cities*’, ‘*The Virtual Field Course*’, ‘*Water-Resources-Related Applications*’, ‘*Digital Earth*’, and ‘*Virtual Forests*’. ‘*Virtual Cities*’ acquired Virtual Los Angeles from the first edition and included three more examples. While each of the examples was created by using a unique combination of software applications, they shared common features such as realistic views of cityscapes, user interaction capabilities, and various viewpoint settings. In the remaining four categories, all examples and descriptions were newly written in the second edition. For example, ‘*The Virtual Field Course*’ described a project in the UK which sought to emulate a field environment for geography students. The authors introduced two major software components: ‘traVelleR’ and ‘panoraMap’. While their way of contributing to the VE was similar to the examples in ‘*Virtual Cities*’—realistic landscapes, free navigation, and interactive panorama views—their interoperability with the web and other GIS software was highlighted.

As with traVelleR, the real power of panoraMap is the ability to link other kinds of information with these panoramic images and to explore this information in a spatial context. (...) Both traVelleR and panoraMap have been designed to take advantage of the VFC Hub mechanism, which connects software to the Internet and means that data stored on remote servers can be incorporated into a session [S2, 441].

Since ‘*Digital Earth*’ and ‘*Virtual Forests*’ went through major updates in the third edition, their content will be covered later with the changes in the third edition.

In addition to VE, the notion of mixed environments (ME) was introduced in the second edition, which is a combination of the real world and a virtual representation such as augmented reality; Thus, on the spectrum of reality and virtuality, ME are closer to the real world than VE. Here, two examples are shown: ‘Project Battuta’ and Hedley et al.’s work. Project Battuta was initiated to develop a new approach for collecting and using geospatial data. Their two key components are “a wearable computer and an associated augmented reality display package that enables users to overlay geospatial data and images on the landscape as one travels through the environment” [S2, 445]. Here, the authors emphasised the importance of hardware and the new interface for augmented reality. Furthermore, Hedley et al. conducted user studies on mixed environments. Their main finding was that “augmented reality interfaces allow people to

complete basic spatial problem-solving activities more quickly and more accurately than when using desktop interfaces for the same 3D geographic visualizations” [S2, 446].

Changes between the second and the third edition. The overall structure remained unchanged in the third edition and most of the updates were about the hardware and VE examples. In the subchapter *Technologies for Creating Virtual Environments*, they mentioned five basic forms of display technologies: desktop, wall-size, head-mounted, room-format, and drafting table-format. Here, they added a paragraph about realising stereoscopic vision on a desktop display by presenting separate images to the left and the right eyes. Also, a specialised wall-sized display called ‘GeoWall’ was introduced, of which potential the authors were excited about. They wrote that “the GeoWall is becoming popular in geology; Johnson et al. indicate that “more than one-third of undergraduate non-major Earth science students in the U.S. already use a GeoWall in their coursework”” [S3, 523].

Among the five categories of VE examples, *Digital Earth* and *Virtual Forests* went through significant updates in the third edition. In *Digital Earth*, they added ‘Google Earth’ which replaced the former two major examples from the second edition—‘Digital Earth’ by the US government and ‘Digital Earth Workbench’ by NASA. Google Earth fits the closest to Slocum et al.’s definition of VE, not only because of its high resolution and level of detail but also because it can create an immersive experience when coupled with other software and hardware. To demonstrate how users can browse through 3D space with stereoscopic vision in Google Earth, the authors introduced the specialised software ‘TriDef Visualizer’ and the hardware ‘GeoWall’. Another example added in the third edition is ‘Science on a Sphere’. Developed by the National Oceanic and Atmospheric Administration (NOAA) of the US, it provides “a room-sized global display that permits the animation of Earth processes such as storms, ocean temperature, and climate changes” [S3, 530].

Also, *Virtual Forests* was almost newly written in the third edition. The software application from Buckley et al. was replaced with Dunbar et al.’s. The authors explained that its strength comes from ‘Visual Nature Studio’, one of the most flexible software packages for visualizing forested landscapes. Furthermore, they demonstrated the complexity of constructing a virtual forest by the fact that “Dunbar et al. used not only Visual Nature Studio, but also four other types of software: remote sensing image analysis, GIS, image editing, and video editing” [S3, 531].

Some minor updates happened in *Virtual Cities*; In the example of Virtual Los Angeles, a link to a website was added which shows various VE projects initiated by the University of California, Los Angeles (UCLA). Although Slocum said during the interview it is unsure if the project was successfully ported to the Windows platform as described in the second edition (Table 12), they confirmed that VE projects are still ongoing at UCLA to assist in urban design and planning. The authors also illustrated the recent efforts at University College London (UCL) to make VE products more accessible to the public using the web; “In another effort known as GeoVUE: Geographic Virtual Urban Environment, CASA researchers have undertaken the ambitious plan of allowing users to develop their own VEs using online tools” [S3, 528].

Table 12. Comparison of descriptions regarding Virtual Los Angeles in *Thematic Cartography and (Geo)Visualization*.

[S1, 251]	[S2, 440]	[S3, 527]
A review of the software [uSim] and associated hardware by Liggett and Jepson (1995) suggests that it will be some time before the lay mapmaker has access to such capability.	Presently, Virtual Los Angeles is available only to those at UCLA who have access to sophisticated Silicon Graphics equipment, but experiments are underway to port the system to a Windows platform.	The Web site for Virtual Los Angeles contains numerous examples of projects that have been undertaken and sample videos.

Summary. Since the second edition, VE have been anticipated as a technique that can create realistic representations of the environment. Some notable examples of VE include Virtual Los Angeles and Google Earth. The observed changes in VE can be seen from two aspects: technological developments and increased accessibility. A lot of examples newly introduced in the third edition show how specialised hardware and software have been developed and refined to construct the immersive environment such as GeoWall. On the other hand, the introduction of Google Earth and efforts from institutions such as UCLA and UCL show how VE have become accessible to the public by using online tools.

Table 13. Continuities of virtual environment examples in *Thematic Cartography and (Geo)Visualization*.

1st ed. (1999)	2nd ed. (2004)	3rd ed. (2014)
Virtual Los Angeles	CommunityViz	
	Virtual Cities Project at UCL	
	UrbanModeller	
CLRview Software		
Emaps Package		
	traVelleR	
	panorama	
	Digital Earth	
	Digital Earth Workbench	
	Keyhole	
		Google Earth
		Science on a Sphere
	Terra Vision	
	ArcGlobe	

Chesapeake Bay VE	
Virtual Puget Sound	
Digital River Basin	
Buckley et al.'s Work	
	Dunbar et al.'s Work
Project Battuta	
Hedley et al.'s Research	

■ Virtual Cities	■ Water-resources-related Applications
■ The Virtual Field Course	■ Virtual Forests
■ Digital Earth and Digital Globes	■ Unclassified

4.1.1.7 Data Exploration and Animation

Overview. Slocum et al. approach animation in two ways; first, as a tool for data exploration, and second, as a stand-alone component. The former is shown in the chapter <<Data Exploration>> and the latter belongs to the chapter <<Map Animation>>. Although these chapters have remained separate throughout all editions, animation exists as the uppermost term that encompasses the key changes and software examples in both chapters. Therefore, data exploration and animation are bound into one category. Codes found in this category are *animation*, *ESDA*, *software*, *visualisation*, *web*, *user*, *statistics*, and *3D*.

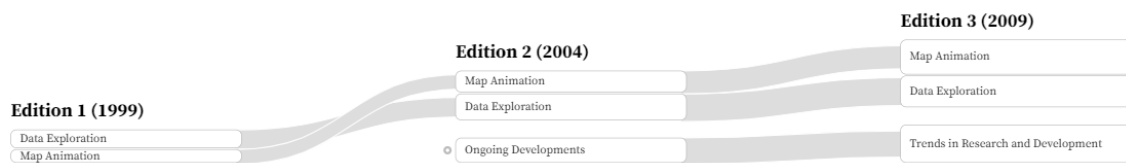


Figure 35. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* that are involved in the formation of the category *Data Exploration and Animation*.

Changes between the first and the second edition. In the first edition, data exploration is defined as “revealing unknowns via high human-map interaction in a private setting” [S1, 210]. Since early data exploration software developed in the 1990s started to contain animation functions, animation was highly anticipated as a novel tool to assist in data exploration. In their first edition, eight software applications were introduced which were described as types of data exploration software that “include a significant animation component” [S1, 210]. Among these, three were classified as deprecated in the second edition, and were labelled as “no longer available or has definite limitations when compared to today’s software” [S2, 389]. Also, the second edition divided the remaining software into two: software developed primarily with geographers in mind, and software not explicitly developed for geographers but still useful for geographers.

As visualised in Table 00, a total of 14 software applications appeared at least once throughout the editions. Some are preserved in all editions and the others have been either newly added or

deleted. Since no significant changes were found in the description of five programmes that are present in all editions, the remaining nine will be described.

ArcView was introduced in the first edition as “popular GIS software that integrates analysis and display capabilities” [S1, 219]. They saw ArcView as a strong data exploration tool because it enables users to view data from a variety of perspectives, and lets users experiment with different symbol schemes. In the second edition, while the description of ArcView remained constant, a paragraph introducing ArcGIS was appended. It was rated as a more sophisticated programme that eventually will replace ArcView 3.x. However, the authors also predicted that the two programmes will coexist, especially because “ArcView 3.x has a large pool of users and ArcGIS requires a more sophisticated platform to run on” [S2, 401].

Transform and Vis5D were both documented until the second edition. Transform was introduced as software that can visualise data associated with both graphs and maps with several unusual features of the time. For example, a ‘Fiddle Tool’ was covered which permits shifting or compressing colours within a scheme. Vis5D was introduced as software that “can handle five-dimensional data such as meteorological data containing three spatial dimensions, a time dimension, and a dimension for the attributes” [S1, 215]. Along with functions such as creating iso-surface, visualising volume, and real-time animation, Slocum pointed out the computational expensiveness of Vis5D in the first edition. On the other hand, this part was partly revised in the second edition with an additional note that “this limitation is now changing” [S2, 403] (Table 14).

Table 14. Comparison of descriptions regarding Vis5D in *Thematic Cartography and (Geo)Visualization*.

[S1, 217]	[S2, 403]	[S3]
Such capability does not come without a price, as one limitation of Vis5D is that it is most effective only on high-end workstations. (I used a Silicon Graphics Iris Onyx/2 with two 150MHz processors, 128MB RAM, and a VTX graphics subsystem.)	Such capability does not come without a price, as one limitation of Vis5D is that it is most effective on high-end workstations. This limitation is now changing, however, as PC-based systems (with the appropriate graphics cards) are now approaching the power of expensive workstations.	<i>Not found in this edition because Vis5D was discontinued.</i>

Two data exploration software applications were added in the second edition: CommonGIS and HealthVisPCP. CommonGIS was introduced as software under development of which goal is to visualise attributes associated with enumeration units of point locations via the web. Its main strength was described as “the ability to automatically select an appropriate method of symbolization, given a set of attribute data that a user wishes to map” [S2, 398]. On the other

hand, HealthVisPCP focuses on visualising the parallel coordinate plot (PCP), a chart that can compactly show correlations of multiple attributes.

Another change in the second edition is the emergence of the term ESDA. According to the authors, this is a concept borrowed from exploratory data analysis (EDA) whose purpose is “to develop new questions and hypotheses via statistical analysis in interactive graphics systems” [S2, 45]. With the semantic expansion as well as the introduction of new software such as CommonGIS, ArcGIS, and HealthVisPCP, the authors started to acknowledge that animation is not the only main functionality that data exploration software can offer (Table 15). As a result, whereas the first edition focused on documenting the animation capabilities of the software examples, the evaluation criteria of data exploration software were diversified in the second edition. The authors added two subchapters *Goals of Data Exploration* and *Methods of Data Exploration* under which ten typical methods of data exploration are summarised. They described these as “a broad range of methods that they would hope to find in data exploration software” [S2, 390] and these are applied to evaluate the functions of CommonGIS and HealthVisPCP.

Table 15. Comparison of descriptions regarding the data exploration software in *Thematic Cartography and (Geo)Visualization*.

[S1, 220]	[S2, 404]	[S3, 477]
In this chapter, we have seen that animation frequently is an important element of data exploration.	In this chapter, we have focused on data exploration software that emphasizes the mapped display. (...) In contrast, ESDA methods emphasize numerical summaries of the data and attempt to more fully integrate them with the map display.	In this chapter, we have looked at a variety of software that has been developed for exploring spatial data. (...) Since a full appreciation of data exploration requires that you interact with the data, we encourage you to experiment with several of the software packages that we have described.

Despite its close link with data exploration software, animation was already given an independent chapter *Map Animation* since the first edition. Animations introduced here resemble a short video clip and are defined as ‘stand-alone’, meaning that they are developed outside data exploration software or electronic atlases.

In the second edition of the chapter *Map Animation*, they added the new subchapter *Early Developments*. Here, they traced back to the 1950s when Norman Thrower (1959; 1961) was enamoured by the potential of animated cartography, who wrote that “By the use of animated cartography we are able to create the impression of continuous change and thereby approach the ideal in historical geography, where phenomena appear “as dynamic rather than static entities”” [S2, 376]. Furthermore, they mentioned Campbell and Stephen Egbert (1990) who pointed out reasons for the stagnation of the development in animated cartography. These were “the lack of

funds in geography departments, the lack of a “push toward the leading edges of technology” in these departments, a negative view of academics who develop and market software, and the difficulty of disseminating animations in standard print outlets” [S2, 376]. However, the authors concluded that this criticism remains in the era before the 1990s and that many of these limitations appear to have been overcome.

Another major update in the second edition was the addition of user studies which are discussed under the subchapter *«Does Animation Work?»*. Here, the authors questioned; “Although those developing animations are often enamored of their capability, we (and others) often find them difficult to understand. Is it possible that the capability of animation has been overhyped?” [S2, 386]. After introducing several contradicting study results, they concluded that “the effectiveness of animation will be a function of numerous factors, including the type of symbology, how the animation is viewed, whether the user has expertise in the domain being mapped, and experience with animation” [S2, 387].

Concerning the examples of stand-alone animation, the first edition included seven examples and the second edition added five more. The second edition also classified them into two categories: those having little or no interaction and those characterised by interaction. Except for five additions, their descriptions have not gone through significant changes. Also, each animation example is highly specific, not showing much connection to other digital technologies. Therefore, the examples are not covered in detail. The continuities of stand-alone animation examples are shown in *Table 17*.

The influence of animation and ESDA reached farther than the boundary of mapping techniques. In the chapter *«Ongoing Developments»* of the second edition, two streams that share the lineage with data exploration were added as new subchapters: *«Information Visualization»* and *«Spatial Data Mining»*. Information visualisation concerns the representation of nonnumerical or abstract information by using spatial metaphor—a process described as ‘spatialization’. One of the examples the authors introduced was the work by Skupin (2002) where they visualised the relationships among more than 2,000 abstracts submitted to the annual meeting of the Association of American Geographers. In another subchapter *«Spatial Data Mining»*, the authors emphasised the necessity of extracting knowledge from very large datasets. Here, spatial data mining was defined as “exploratory approaches for finding interesting patterns in databases and ultimately uncovering knowledge about those databases” [S2, 461]. After illustrating two examples, they reaffirmed the relevance of cartography in spatial data mining, by referencing Adrienko et al. that spatial data mining should be utilised with interactive visualisation techniques.

Changes between the second and the third edition. In the chapter *«Data Exploration»* of the third edition, four data exploration software examples were deleted. ArcView and ArcGIS were replaced by GeoDa which was introduced as “the latest incarnation of tools for exploratory spatial data analysis” [S3, 473]. Also, Transform and Vis5D were replaced by Integrated Data Viewer (IDV), which inherited the majority of Vis5D’s functions with novel visualisation techniques such as displaying the Earth as it is seen from space. Along with IDV, ESTAT was also newly introduced. It was described as an interactive geovisualisation environment that offers “a

scatterplot, parallel coordinate plot, bivariate choropleth map, and time-series graph” [S3, 475]. The continuities of data exploration software examples are shown in *Table 16*.

Also, the authors further elaborated on the cognitive and perceptive issues of animation. Here, they introduced the notion of ‘change blindness’ that “Information in an animation can be missed, especially when the animation is complex (e.g., when changes are happening at five or six different places on the map)” [S3, 454]. Also, they referenced the concerns proposed by Harrower and Sheesley (2005) about the ‘fly-by’ animation technique. These were “oblique perspective, focus on realism and thus information overload, visual occlusion, and user disorientation” [S3, 442]. Furthermore, they showed the possibility of improving the usability of animation using 3D space in the new subchapter *Using 3-D Space to Display Temporal Data*. For example, after explaining the software developed by Richard Brownrigg (2005), they positively commented that “It is apparent that information can be gleaned from 3D views that is difficult or impossible to obtain from an animation” [S3, 453].

In the chapter *Ongoing Developments* of the third edition, another subchapter *Visual Analytics* was added. Visual analytics is defined as “the science of analytical reasoning facilitated by interactive visual interfaces that should detect the expected and discover the unexpected” [S3, 557]. Here, the authors explained this new field through activities of the U.S. National Visualization and Analytics Center (NVAC) which played a key role in utilising the synthesis of visualisation and analytics in the US. Furthermore, they noted the potential of visual analytics by mentioning Turdukulov et al. (2007) and Weaver et al.’s (2007) research.

Traditionally, researchers have tried to use animation for this purpose, but these small features [rip channels] are difficult to detect, due to change blindness. As an alternative, Turdukulov et al. developed a sophisticated analytical approach for detecting rip channels, as well as novel methods for visualizing rip channels’ change over time [S3, 558].

Weaver et al. describe the development of an interactive visual tool for exploring the visitation patterns of guests at Pennsylvanian hotels in the late 1800s. Their paper provides an interesting example of interdisciplinary work, in this case conducted by visual analytics experts and historical geographers [S3, 558].

Summary. From Slocum et al.’s books, it can be found that animation was initially spotlighted in cartography as a tool to assist in data exploration. Thus, it was introduced as the foremost feature in a lot of early data exploration software. However, the perspective of data exploration was diversified since the second edition with the emergence of the ESDA concept, and other functionalities other than animation became integrated with the data exploration. Despite their close ties with data exploration, animations also existed as stand-alone applications. While they have long existed as a highly anticipated visualisation technique, subsequent research and user studies were followed to evaluate the effectiveness and cognitive limitations of animation. Furthermore, the influence of animation and ESDA is observed in new neighbouring fields of cartography that are added along the edition such as information visualisation, spatial data mining, and visual analytics.

Table 16. Continuities of data exploration software examples in *Thematic Cartography and (Geo)Visualization*.

1st ed. (1999)	2nd ed. (2004)	3rd ed. (2014)
Moellering's Software		
ExploreMap		
Aspens		
MapTime		
Project Argus		
ArcView		
	ArcGIS	
	CommonGIS	
		GeoDa
Transform		
Vis5D		
	HealthVisPCP	
		ESTAT
		IDV

- Either no longer available or has limitations but provides a historical perspective
- Developed primarily with geographers in mind
- Not explicitly developed for geographers but useful for geographers

Table 17. Continuities of stand-alone animation examples in *Thematic Cartography and (Geo)Visualization*.

1st ed. (1999)	2nd ed. (2004)	3rd ed. (2014)
Thunderstorm Development		
Portrayal of the Ozone Hole		
Animation of Climate Change		
Peterson's Nontemporal Methods		
Animations for Course & Lectures		
	Animation of Urban Growth	
	Animation of Map Design Param.	
		Examination of Rate of Change
		Animated Atlas of Air Traffic
Dorling's Work		
Monmonier's Graphic Scripts		
	Harrower's Work	
	Andrienko et al.'s Work	
	Deep Thunder	

- Animations having little or no interaction
- Animations characterised by substantial interaction

4.1.1.8 Electronic Atlases and Web Maps

Overview. In Slocum et al.'s book, electronic atlases existed as an independent chapter in the first and the second edition. This changed in the third edition; the chapter was discontinued and only a portion of electronic atlas examples was transferred to <<Web Mapping>>. Although most of the chapter <<Web Mapping>> was newly written, it is linked with the chapter <<Electronic Atlases and Multimedia>> for the following reasons; First, the advent and success of web influenced electronic atlases and a lot of them were relocated to the web from CD-ROM. Second, web map examples in the third edition share a lot of functionalities of electronic atlases. Therefore, two chapters concerning electronic atlases and web maps are linked and formed into a single category. Codes found in this category are *web*, *atlas*, *multimedia*, *user*, *ESDA*, and *software*.



Figure 36. Flow diagram of chapters from *Thematic Cartography and (Geo)Visualization* that are involved in the formation of the category *Electronic Atlases and Web Maps*.

Changes between the first and the second edition. In the first edition, an electronic atlas was defined as “a collection of maps (and database) that is available in a digital environment. [S1, 233]. The earliest examples date back to the mid-1980s when they were stored in physical external storage and interaction capabilities were very limited, just like paper atlases but in a digital form. With the steep growth of digital technologies, however, electronic atlases have evolved into more than a collection of maps and could “take advantage of the capabilities that computer-based systems provide, permitting data exploration, animation, and multimedia” [S1, 231].

The first edition contained eight electronic atlas examples, all of which made it to the second edition. Also, except for ‘The Domesday Project’ which was a system designed for specialised BBC microcomputers in UK schools, they wrote that the other examples were “all available at a modest cost (or free via the Internet)” [S1, 233].

‘Digital Atlas of California’ was introduced as an example which can be accessed for free by anyone with an internet connection. In the second edition, this was incorporated into a larger project ‘The Electronic Map Library’ which consists of a series of atlases for cities and states of the US, along with a set of panoramic maps. Both in the first and the second edition, the same limitation was listed; “No interaction with the maps is possible since the maps are in GIF format, which is useful for browsing and downloading, but not for providing information about the underlying data” [S1, 234]; [S2, 408].

For ‘The Interactive Atlas of Georgia’, the authors focused on its interactive tools and compared them with those from the data exploration software ‘ExploreMap’. They also pointed out its limitations that “it does not contain text, photographs, or graphs; nor does it contain animation or multimedia capability” [S1, 234], which they attributed to the limited funding available

for state atlases. In the second edition, they added a paragraph introducing the website ‘Georgia GIS Data Clearinghouse’ which contains state-produced maps, including these electronic atlases.

Next to the state of Georgia, the second edition added the ‘Atlas of Oregon’ as an example of state atlases. It was published in both printed and digital versions; The former inherited the earlier printed edition, and the latter had the same set of maps as the printed ones but with interactive and animated capabilities. The authors evaluated the maps focusing on their effectiveness and user interfaces. For example, a ‘magnify tool’ was introduced as an unusual tool with which users can enlarge the area within a square, while the rest of the map remains at the original scale. Although nowadays this is seemingly unintuitive since universal zooming and panning are the most common features in digital environments, they noted that “We found this approach an attractive option for examining detail while still keeping the overall map in mind” [S2, 414].

‘Microsoft Encarta Virtual Globe’ was introduced as a recent example of commercial atlases in the first edition. It offered various map styles and advanced interactive capabilities for the time, such as rotating the globe by clicking and dragging the mouse. This was continued in the second edition with the updated version called ‘Encarta Reference Library 2003’ which was further complemented with data exploration techniques. For example, they added a feature called ‘Map Treks’ which provides guided tours of various geographic features and concepts around the world. Also, it contained a set of ‘Virtual Tours’ that include 360° panoramic photographs of a variety of places around the world, which was evaluated that “Such realism moves the library into the realm of virtual environments, as the intention is to give you the feeling of actually being in a place” [S2, 415]. Also, the authors mentioned the advantage of the offline distribution, that the maps are more responsive compared to the web-based.

‘TIGER Mapping Service’ was designed for mapping 1990 US census data, introduced in the first edition as “an intriguing example of the capability of digital mapping, but also illustrates the difficulty of defining what constitutes an electronic atlas” [S1, 236]. This is because it does not showcase completed maps but rather lets users build their own maps using the underlying dataset. As a result, they concluded that “the term “map server” is potentially more appropriate as such systems “serve” maps to users [S1, 237]. The major update in the second edition was that a new service ‘American Fact Finder’—designed for mapping 2000 US census data—was launched, and it was bound together with TIGER Mapping Service under the section ‘*US Census Bureau Online Mapping*’. Here, the authors compared two services from a thematic mapping perspective and pointed out that both versions only offer choropleth visualisation, which is particularly problematic considering that census data are usually provided in absolute numbers.

The other examples from the first edition: ‘The Domesday Project’, ‘World Atlas’, ‘Death Valley: An Animated Atlas’, and ‘Grolier’s Multimedia Encyclopedia’ were mostly unchanged in the second edition. The Domesday Project is one of the earliest electronic atlases, developed in 1986 in the UK to provide a contemporary snapshot of the nation during the 1980s. It was delivered in two video disks containing approximately 250 megabytes of text, maps, and photos. World Atlas was one of the first widely distributed commercial electronic atlases and contained sound, video, and pictures. Incorporating videos was a hurdle because of their storage requirements, and the

authors wrote that “Video clips were shown only for each major city in the world, and only a limited number were included for each city; furthermore, the clips were small (taking up approximately one-tenth of the screen) and lasted only a few seconds” [S2, 414]. On the other hand, *Death Valley: An Animated Atlas* contained plenty of animations with spoken text and was introduced as an example of where animated maps can be used to provide a series of narratives of a certain region. Similarly, *Grolier’s Multimedia Encyclopedia* used the combination of multimedia and animated maps to deliver information about major historical events.

The chapter *«Electronic Atlases and Multimedia»* was enriched in the second edition with five additional electronic atlas examples: the Atlas of Oregon, American Fact Finder, the National Atlas of the US, Atlas of Switzerland Interactive, and David Rumsey Map Collection. The former two examples were described above with the corresponding examples in the first edition: Interactive Atlas of Georgia and TIGER Mapping Service. Thus, the latter three will be covered below.

‘The National Atlas of the US’ contains five major categories of mapping capability: Interactive Maps, Multimedia Maps, Printed Maps, Printable Maps, and Map Layers Data Warehouse. It enables users to map various types of thematic categories and utilises animation and interactivity to examine a variety of topics. Moreover, users can order large format hard copy maps, print maps that are suitable for home printers, and download actual data that can be imported into other programmes, such as ArcView and ArcGIS. The authors commented that “[It] illustrates the considerable flexibility that electronic atlases can provide, especially when they are implemented in a web format” [S2, 412].

‘Atlas of Switzerland Interactive’ is a CD-ROM-based atlas with thematic maps for Europe, with a series of topographic maps for Switzerland. The authors praised its capability to examine data through time and at different spatial levels and to create alternative views of topographic maps.

‘The David Rumsey Map Collection’ is an online set of more than 8,800 maps from David Rumsey’s personal collection of 18th- and 19th-century maps of North and South America. In addition to its function as an online database of historical maps, the authors introduced its ongoing project; “Rumsey has begun to link some of the maps with modern geographic data, such as satellite images, and a variety of vector and raster data available from the US Geological Survey” [S2, 408]. Using Rumsey’s custom GIS browser, users were allowed to drape modern images and maps over historic maps, or even create 3D images by overlaying historical images on DEMs.

Aside from updates in examples, there was another major change in the second edition; the subchapter *«Tools for Developing Your Own Software»* was deleted. In the first edition, Slocum mentioned the necessity of developing one’s own data exploration software because “software packages frequently have weaknesses and limitations that prevent a user from accomplishing a desired task” [S1, 237]. They provided an overview of popular programming languages of the time (e.g., Visual Basic, Delphi, Tcl/Tk, Java, et cetera), authoring software such as Macromedia’s Director that is used for developing educational presentations, and visualisation software such as IBM Visualization Data Explorer. In the preface of the second edition, the authors clarified that they decided to not include this subchapter anymore because “this is of less interest for the introductory student, and it is difficult to ensure that the material is up-to-date” [S2, ix].

Changes between the second and the third edition. The chapter «*Electronic Atlases and Multimedia*» was replaced by the new chapter «*Web Mapping*». Here, the authors stated that “It is our intent to survey those websites that provide maps and mapmaking capability directly. We do not include websites that, for example, offer software that you must purchase as a CD (or download directly) and install on a local computer via the web” [S3, 503]. As a result, the majority of electronic atlases from the second edition were omitted since they were distributed by CD-ROMs and were not transferred to the internet. Only three of them survived; American Fact Finder, the National Atlas of the US, and the David Rumsey Map Collection. The continuities of the electronic atlas and web map examples are shown in *Table 24*.

This chapter included a total of 13 web map examples, differentiated by five criteria that Slo-cum et al. developed. These are: (1) Pre-existing maps versus custom-tailored maps; (2) Low interaction versus high interaction; (3) Distributed data versus user-contributed data; (4) Updated data versus non-updated data; and (5) Animated maps versus static maps.

The first criterion concerns the users’ right to participate in the cartographic process. Maps that allow users to execute cartography (e.g., specify the kinds of data shown on the map, choose the scale, modify the coordinate systems, experiment with map designs) are classified as *custom-tailored maps*. ‘American Fact Finder’ from the second edition was introduced here as an example and the description was updated from comparing its functions with TIGER Mapping Service to explaining what users can do during the mapmaking. On the other hand, the maps that can only be viewed or downloaded from the website are classified as *pre-existing maps* with the example of ‘National Park Service’. Here, it was described how users can view, print, and download maps from the US National Park system in PDF or Adobe Illustrator (.ai) format.

The second criterion is the level of interactivity which is further broken down into three: *low*, *medium*, and *high*. The authors elaborated on the levels as: “websites that allow no interactivity; websites that permit limited user interactivity, such as zooming in and out, panning, and re-centring the map; and websites that allow to specify a coordinate system, add or remove various data layers, query the data attributes, measure distances, and so forth” [S3, 504]. ‘The National Atlas of the US’ from the second edition was introduced as *medium interaction*. And ‘TypeBrewer’ belonged to *low interaction*, of which function is to let users test and adjust type attributes in a basic mapping environment. In contrast, ‘Google Earth’ and ‘World Wind’ were covered as websites with *high interactivity* where users can zoom and fly around the virtual globe and view geospatial data. The authors illustrated their common functionalities such as displaying satellite imagery, zooming into a specific location, and the presence of an extensive user forum.

The third criterion is about the centrality of data—distributed by an authority or created by user contribution. However, these two were not incompatible and the authors included examples that share both ends. For example, ‘The National Map’ was mentioned as a combination of *federal-distributed* and *user-contributed data* where users can download data distributed by the US Geological Survey and share their map results with the geospatial community. For websites that focus on the *distribution of centralised data without sharing*, West Virginia’s Department of the Environment was introduced. Afterwards, they covered the ‘New York City Subway Smell Map’

which relies on voluntary data collection; Users can report odours they have sensed in a subway station as well as their personal experiences while exposed to the smell.

The fourth criterion is whether the map is regularly updated. Among websites where the geospatial data is *not updated*, ‘Perry-Castañeda Map Library’ was introduced. The authors described that users can browse the extensive collection of maps and download them in .jpg or .pdf format. For the *updated data*, two websites are introduced: ‘AirportMonitor’ and ‘iVidlog’. AirportMonitor is a website that “shows the current status of aircraft arrivals and departures at selected airports across the United States” [S3, 511]. iVidlog pursues a novel approach to real-time data, by showing users virtual driving videos of the road network in Pennsylvania.

The final criterion is animation. Here, the authors reminded readers of the significance of animation; “Animation as a way to visualize data has been a focal point of cartographic activity for the past few decades” [S3, 512] and contrasted them with static maps. ‘The David Rumsey Collection’ from the second edition was introduced as an example of *static maps*, with almost no changes in the description. For *animated maps*, they showed the map that shows the number of US and other coalition fatalities in Iraq. Furthermore, they provided detailed information about Adobe Flash as a development tool for authoring multimedia on the internet.

Aside from web map examples, semantic inventions and expansions related to web mapping are observed in the chapter <<Trends in Research and Development>> of the third edition. For example, they introduced a new subchapter <Mobile Mapping and Location-based Services (LBS)> where the emergence of LBS was attributed to “having access to the web and determining the geographic location of the device (e.g., via the global positioning system)” [S3, 558]. The authors illustrated the changes that the unique mobile environment would bring to map design, such as the small screen size and the need for rapid response. Also, foreseeing the further distribution of smartphones and LBS technologies, they introduced two novel terms associated with mobile mapping; The first is ‘TeleCartography’ which refers to “the distribution of cartographic presentation forms via wireless data transfer interfaces and mobile devices”. The second is ‘ubiquitous cartography’ which is defined as “the ability for users to create and use maps in any place and at any time to resolve geospatial problems” [S3, 558]. Here, the authors predicted that mobile mapping and LBS could incur another crucial shift in cartography as the web did.

Assuming that many of these subscribers will be utilizing smartphones and related location-based services, the impact on cartography could be comparable to the impact the Web has had. Certainly, this new technology has the potential to dramatically change the extent to and the manner in which people utilize maps [S3, 558].

Summary. Along the editions of Slocum et al.’s book, the transitional period is revealed where a lot of electronic atlases in CD-ROMs were transferred to the internet, and then eventually replaced by web maps. In all editions, the description of their capabilities and limitations provide a snapshot of the time concerning which functionalities were valued and anticipated. In the third edition, these priorities have become more pronounced through the five criteria of web maps. Also, the influence of the web has propagated further and resulted in novel ways of producing and consuming geodata such as location-based services (LBS) and mobile mapping.

Table 18. Continuities of electronic atlas and web map examples in *Thematic Cartography and (Geo)Visualization*.

1st ed. (1999)	2nd ed. (2004)	3rd ed. (2014)
The Domesday Project		
Digital Atlas of California	The Electronic Map Library	
The Interactive Atlas of Georgia	Atlases for Georgia	
	Atlas of Oregon	
World Atlas		
Encarta Virtual Globe	Encarta Reference Library 2003	
Death Valley: An Animated Atlas		
Grolier's Multimedia Encyclopedia		
TIGER Mapping Service	TIGER Map Server	
	American Fact Finder	Custom-tailored map
		National Park Service Pre-existing map
		Type Brewer Low interaction
	The National Atlas of the US	Medium interaction
		Google Earth and World Wind High interaction
	Atlas of Switzerland Interactive	
		West Virginia Distributed
		The National Map Distributed & User-contributed
		New York City Subway Smell Map User-contributed
		AirportMonitor Updated map
		iVidlog Updated map
		Perry-Castañeda Map Library Non-updated map
		Iraq Coalition Fatalities Animated map
	The David Rumsey Map Collection	Static map

- Pre-existing vs. custom-tailored maps
- Updated vs. non-updated data
- Low interaction vs. high interaction
- Animated vs. static map
- Distributed vs. user-contributed data
- Unclassified

4.1.2 Evolution of Web Resources

From the first to the third edition, the absolute occurrence of hyperlinks has increased from 26 to 161. Since this is overrepresented by the increased volume of the book—the first edition was 332 pages whereas the third edition is 623 pages—the absolute count is normalised by dividing it by the total number of words of each edition. As a result, even after considering the volume of each edition, the relative occurrence of hyperlinks doubled in the second edition and more than tripled in the third edition. This change is largely due to new chapters related to digital technologies that reflect the trend of more maps being published on the internet. For example, the chapter «*Virtual and Mixed Environments*» in the second edition included 18 hyperlinks which increased to 24 in the third. The chapter «*Web Mapping*» alone includes 29 hyperlinks.

Meanwhile, in the second edition, Slocum et al. started providing hyperlinks for software applications and maps that were continued from the first edition. These examples include CommonGIS, IBM's Visualization Data Explorer, and MigMap. Even when the description remained unchanged—which is why the latter two are not covered in the above content analysis—their websites were added in the second edition. Also, in the chapter «*Map Animation*» where half of the examples in the third edition existed since the first edition, the number of hyperlinks steadily increased from one, 10, and eventually to 14. This accounts for the relatively steep increase of the number in the category 'Software applications related to mapping techniques', which is not only due to the newly introduced maps and software packages but also because the existing content was complemented with web resources.

Table 19. Absolute and relative occurrence of hyperlinks in *Thematic Cartography and (Geo)Visualization*.

	1st ed. (1999)	2nd ed. (2004)	3rd ed. (2014)
Occurrence of Hyperlinks	26	95	161
Total Word Count	119,357	218,264	230,403
Relative Occurrence of Hyperlinks	21.8	43.5	69.9

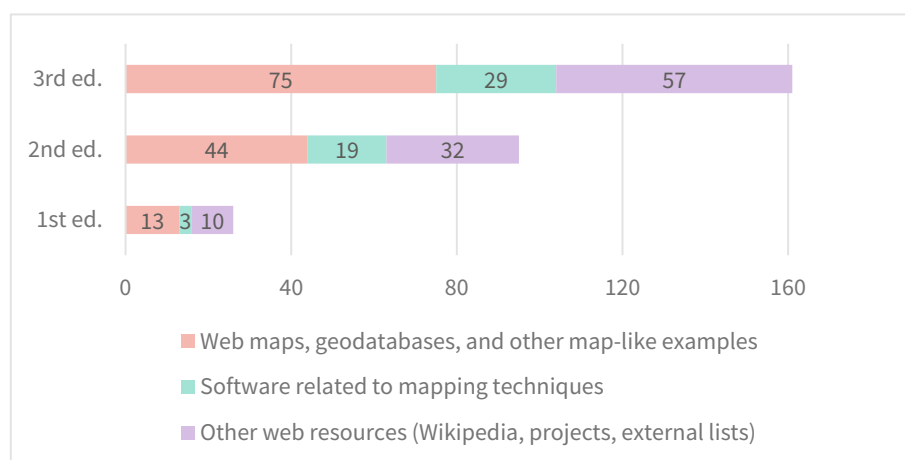


Figure 37. Classification of hyperlinks in *Thematic Cartography and (Geo)Visualization*.

Aside from hyperlinks that are related to maps and software, Slocum et al. started introducing more external references along with editions. This includes web sources such as Wikipedia and websites of projects that provide further information about the topic, and external lists containing a vast number of examples that the book could not all incorporate. For example, the second edition and the third edition introduced websites that list map servers or publications for reference purposes. Furthermore, there are a few websites that are less related to mapping techniques or software, which also increased in numbers. For example, in the second edition, when they introduced a wristband with a GPS function which can pose privacy issues, the website of the manufacturer was attached. Although these web resources are mainly auxiliary, it shows the change in textbook writing and how it has become more common and convenient to outsource the ‘further reading’ by referencing external web resources.

4.1.3 Derived Information from Text Mining

Here, results from two text mining methods are shown: word frequency and relationships among words using bigram. First, word frequencies are calculated to form a list of top-15 words and top-15 digital keywords. Top-15 words provide a general overview of what a book is about, and the top-15 digital keywords focus on topics that are more relevant to the overarching theme of digital cartography. Second, bigrams of the digital keywords are computed and visualised to investigate the relationships among these keywords and their neighbours.

4.1.3.1 Top-15 Words and Top-15 Digital Keywords

After removing stop words and conjugations, word frequencies are calculated per edition and ordered by occurrence (*Table 20*). Words that belong to digital keywords—extracted using the content analysis results—are highlighted in yellow.

In all editions, *map*, *data*, *color*, and *method* ranked high on the list. The substantial restructuring between the first and the second edition is noticeable; Not only have overall occurrences of words grown significantly because of the increase of the volume (*Table 19*) but also the rank of words changed. For example, *projection* and *printing* were newly included in the second edition due to new chapters dedicated to map projection and map production. Also, *software* and *scheme* from the first edition were replaced by new words such as *user*. Also, some general words such as *approach* and *show* were demoted; Instead, *line* turned out to be an important word, which is repeatedly mentioned in multiple chapters such as «*Principles of Symbolization*» and «*Scale and Generalization*» that were newly introduced in the second edition. Furthermore, a lot of occurrences of the word *variable* were overtaken by *attribute* in the second edition in the context of mapping, which accounts for the demise of *variable* and the advent of *attribute*.

A familiar thematic map is the choropleth map, in which enumeration units (or data collection units) are shaded to represent different magnitudes of a **variable** [S1, 2].

A common thematic map is the choropleth map, in which enumeration units (or data collection units such as states) are shaded to represent different magnitudes of an **attribute** [S2, 2].

Since the book structure has been stabilised since the second edition, the ranks did not fluctuate in the third edition in comparison to the second. *User* climbed up to the fifth, however, *printing* was not included anymore. This concurs with the results from the content analysis as well as the interview that the printing technologies became less relevant whereas user—which often appears in the context of web maps to explain their interactive functionalities—have become more important.

When the scope is narrowed to digital keywords (Figure 38), *user* was the winner since the second edition. In comparison to the first edition, its occurrence increased more than three times—from 173 to eventually 593. *Animation* was the most frequently mentioned in all three editions among the terms related to mapping techniques, followed by (*data*) *exploration*. The substantial change in the first and the second edition is again visible with the ranks of digital keywords; New words such as *generalization*, *virtual*, *device*, and *web* appeared whereas older ones such as *electronic*, *sound*, and *GIS* were ranked lower. This could provide insights into which topics gained more relevance during the restructuring while others experienced relative loss. For example, the occurrence of *electronic* almost did not change despite the increase in the volume of the second edition. This indicates that the quantitative coverage of topics such as ‘electronic atlas’ and ‘electronic version of maps’ was constant. On the other hand, *interactive* increased almost three times, which suggests that web-based maps that offer interactive functionalities were largely increased in the second edition.

In the third edition, while *user* remained at the very top, its occurrence also increased from 435 to 593. This is a notable increase considering that the entire word count of the third edition only increased 5.6% compared to the second edition (Table 19). Also, the occurrence of *web* rose from 90 to 264, ranking as the fifth most frequently mentioned digital keyword. *Electronic* and *exploration* were removed from the list and became surpassed by new terms such as *uncertainty* and *http*. ‘Uncertainty’ is a word repeatedly mentioned in the chapter «*Visualizing Uncertainty*» which was largely complemented in the third edition, and ‘http’ is representative of hyperlinks which also sharply increased. Whereas *digital* increased from 57 to 137 in the second edition, the number barely changed in the third edition, resulting in a lower rank.

While the word count provides a succinct overview, it should be noted that some words are underrepresented, and therefore, it cannot be used as a direct proxy for the importance of topics. This is because the usage of words is a personal decision, and the occurrence also depends on it as well. For example, the word *uncertainty* equals ‘data quality’ in the third edition of Slocum et al.’s book, without many competitor words or conjugated forms. Also, it is repetitively mentioned inside a large chapter, which accounts for its high rank. On the other hand, (*data*) *exploration* is a topic that is covered in multiple chapters, and the activity is named differently with conjugated forms of ‘exploration’. Unfortunately, among these, only ‘exploratory’ was counted with the original word ‘exploration’ because other forms such as ‘explore’ and ‘exploring’ were too often used outside the context of data exploration. And since it is impossible to consider the context of each use case in a simple word count, the word ‘exploration’ was counted less than its actual usage, resulting in what counters the perceived importance of the topic.

Table 20. Top-15 words of *Thematic Cartography and (Geo)Visualization*.

1st ed. (1999)		2nd ed. (2004)		3rd ed. (2014)	
Word	Occurrence	Word	Occurrence	Word	Occurrence
Map	1616	Map	2874	Map	3195
Data	1221	Data	1593	Data	1713
Color	807	Color	1064	Color	983
Symbol	450	Method	618	Method	629
Method	426	Projection	583	User	593
Variable	414	Symbol	557	Symbol	571
Value	342	Line	494	Projection	569
Class	331	Attribute	483	Line	528
Animation	317	Value	462	Information	497
Software	290	User	435	Attribute	491
Unit	278	Class	432	Class	486
Approach	242	Information	424	Value	459
Information	237	Animation	412	Animation	433
Show	226	Printing	392	Approach	420
Scheme	224	Unit	389	Unit	401

To summarise, the word count in Slocum et al.'s book well represents the major restructuring between the two consequent editions, especially when filtered by digital keywords. By observing which keywords are contained in the top-15 list, it can be inferred which topics have gained relevance. These were *generalization*, *virtual*, *device*, and *web* in the second edition, and *uncertainty* and *http* in the third edition. Furthermore, words such as *web* and *user* showed a disproportionately large increase both in the second and the third edition, which means their importance has become phenomenally larger. However, words such as *exploration* and *uncertainty* suggest that the interpretation of results requires precaution. This is because some words are accurately counted because of their homogenous use cases, while some are counted less than their use since they are referred to in various forms and used in different contexts. As a result, additional information is required to view the changes from another perspective and to upscale the interpretation of the word count to the chapter or topic level.

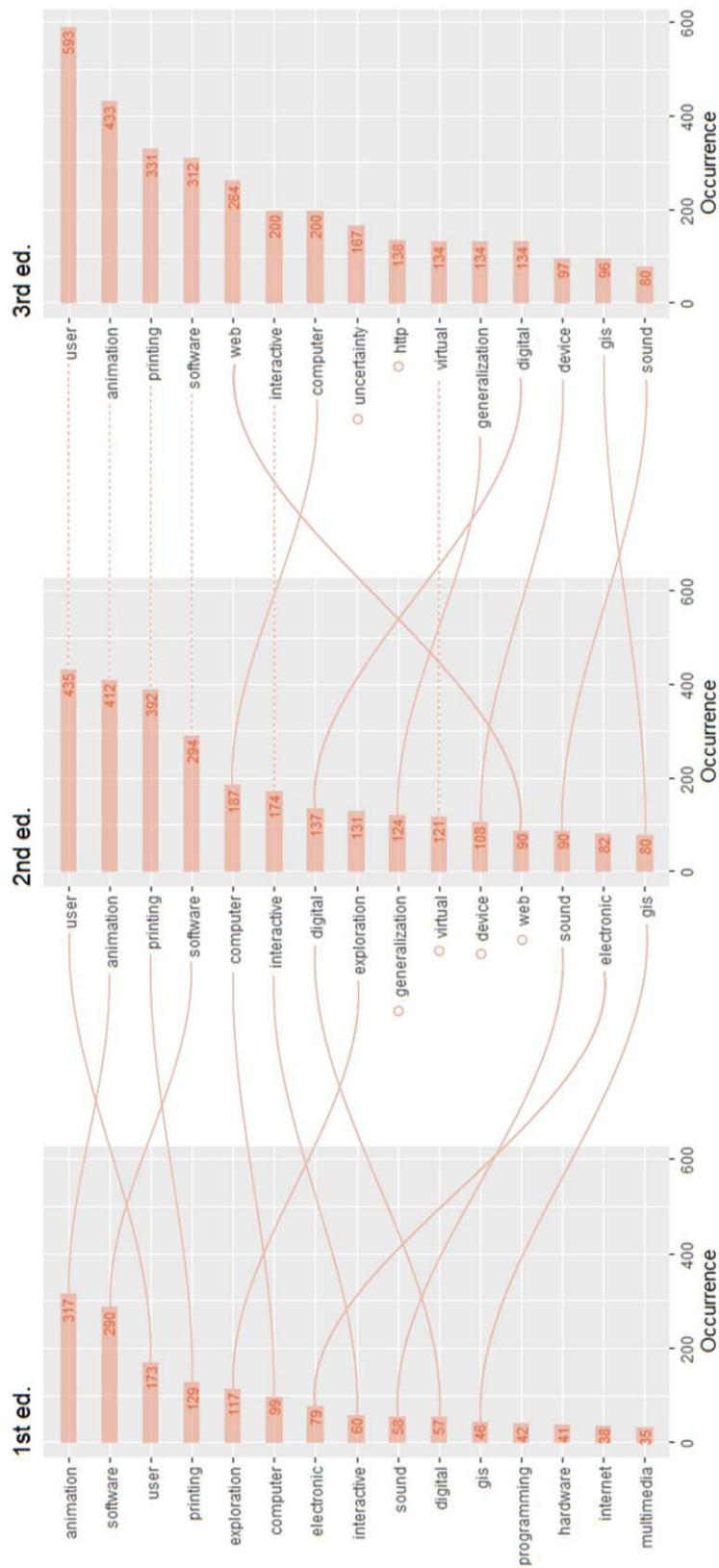


Figure 38. Flow diagram of bar charts showing changes in the rank of digital keywords from *Thematic Cartography and (Geo)Visualization*. Keywords are connected to the next edition with a flow line. Dotted lines are drawn for keywords whose rank did not change between the editions. Keywords that are newly included in the top-15 are marked with a dot on the left side of the label. Keywords without any lines indicate that they were not included in the top-15 in the following edition.

4.1.3.2 Relationships among Digital Keywords and Their Neighbours

Bigrams are used in order to consider the context of words, looking at the pair of words that co-occur in an edition. After calculating the occurrence of bigrams, they are filtered by digital keywords. Next, they are visualised into a directed network graph showing the frequency of bigrams that occurred at least 10 times in the edition (*Figure 50*). Finally, the graphs are aggregated per book so that the changes along the editions are better pronounced.

The strongest link is *data exploration* which occurred more than 75 times in all three editions. Considering that the word ‘exploration’ in the context of data exploration is undercounted because of its conjugated forms such as ‘explore’ and ‘exploring’, the actual coverage of the topic of data exploration in the book is likely to be even larger. *Electronic atlas* was a common bigram until the second edition, then disappeared in the third, which is due to the restructuring where the entire chapter was deleted and replaced by <<*Web Mapping*>>. As a result, bigrams such as *web site(s)* and *(world) wide web* emerged. Also, as more hyperlinks for web resources were included in the book, *http* coming after *web site* became a common expression in the third edition. Also, it can be seen how *virtual reality* became less preferred since the second edition as the term *virtual environment* became normalised. Similarly, the mention of *programming languages* became less common in the second edition, largely due to the deletion of the subchapter <<*Tools for Developing Your Own Software*>>.

In all three editions, *software* was usually preceded by *(data) exploration*, *GIS*, *mapping*, and followed by *package*. Bigrams *visualization software* and *authoring software* were only common in the first edition; The former used to refer to software such as Adobe Illustrator and was replaced by the term ‘graphic(al) software/package’ in the second edition. The latter was introduced as “a flexible set of graphical tools and multimedia/hypermedia capability that enable cartographic applications to be developed more easily than do programming languages” [S1, 232], covered in the deleted subchapter <<*Tools for Developing Your Own Software*>>. On the other hand, *software application* and *labeling software* emerged as common terms in the second and the third edition, due to new examples that were added.

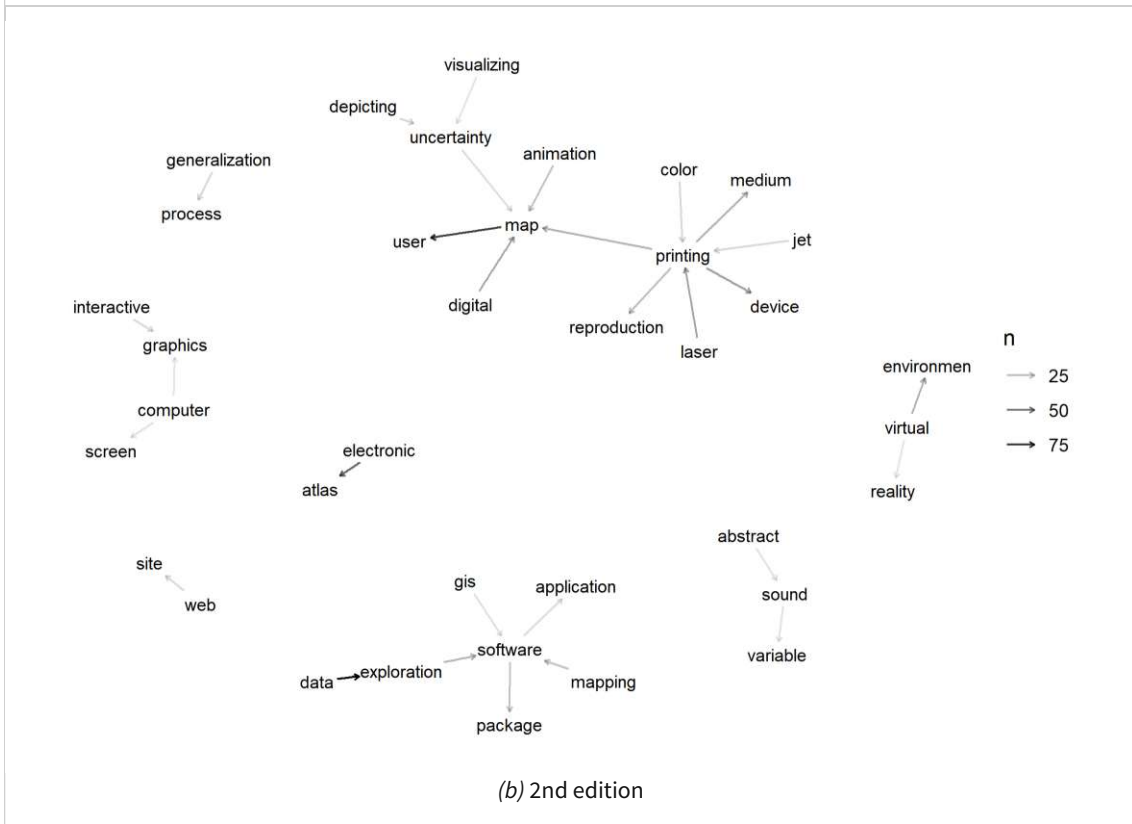
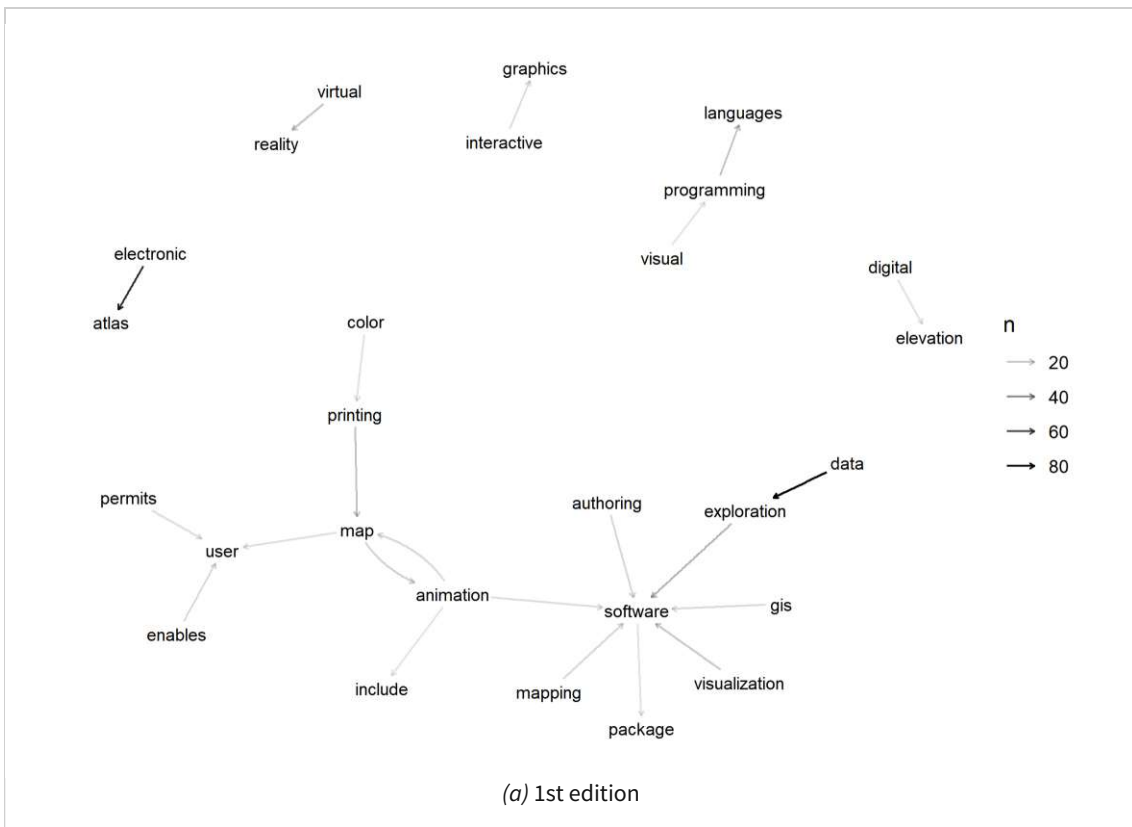
The bigrams of the word *printing* are only pronounced since the second edition, probably due to the smaller volume of the first edition. Specific printing methods such as *(ink-)jet printing* and *laser printing* occurred less in the third edition, which aligns with the observation that a lot of low-volume printing methods were removed from the third edition because the authors “felt that it was unimportant to focus on particular characteristics of printers”.

The word *map* was usually preceded by *printing (printed)*, *animation (animated)*, *digital*, and *uncertainty*. Among these, bigrams *digital map* and *uncertainty map* occurred only since the second edition; *Digital map* was rarely used in the first edition (only four occurrences in total) because almost all digital map examples were electronic atlases and the map production mainly focused on printing methods. And the bigram *uncertainty map* only appeared since the second edition due to the new chapter <<*Visualizing Uncertainty*>> as well as the term ‘uncertainty’ starting to be preferred over ‘data quality’. As already observed in the word count, the term *user* increased abruptly in the second edition, resulting in the growth of the bigram *map user*. Also, the

bigram *computer screen* became common since the second edition, and *wearable computer* since the third edition due to an increased number of descriptions regarding recent developments in virtual environments and related hardware.

To summarise, the bigram analysis provides another aspect from which the results of the word frequency can be viewed. By considering the context of digital keywords, some subchapter level changes such as the deletion of the subchapter *Tools for Developing Your Own Software* are better pronounced which were less visible when looking at individual words as a unit. Also, the new combination of words that have become common such as *wearable computer* and the changes in terms such as *virtual reality* versus *virtual environment* could be identified. Furthermore, the general observation concurs with the findings from content analysis and the word count; The relevance of the web and its interactive functionalities have increased. However, it should be noted that in Slocum et al.'s book, the first edition has a smaller text corpus compared to the latter two. This means that counting the absolute number of bigrams could be an unfair comparison for the first edition, suggesting another approach that is more robust to the size of the text corpus would be more appropriate.

4 Results



4.2 Kraak & Ormeling’s *Cartography: Visualization of (Geo)Spatial Data*

Kraak & Ormeling’s *Visualization of (Geo)Spatial Data* was first published in 1996, followed by subsequent editions in 2003, 2010, and 2020. The textbook traces back to its original Dutch version *Kartografie*. In the interview, Kraak said that “there was no Dutch cartography textbook during the time” and that their initial motivation was to write a textbook for students learning cartography in the Netherlands. And based on the second edition of *Kartografie* (1990), Kraak & Ormeling published the first English version to reach out to a broader audience. The Dutch version no longer continued after its final update in 1999; Kraak mentioned that “the student population in the Netherlands was not high and it was cumbersome to continuously update the book in two languages”. As a result, the English version became the standard in 2003 as shown in *Figure 41*.

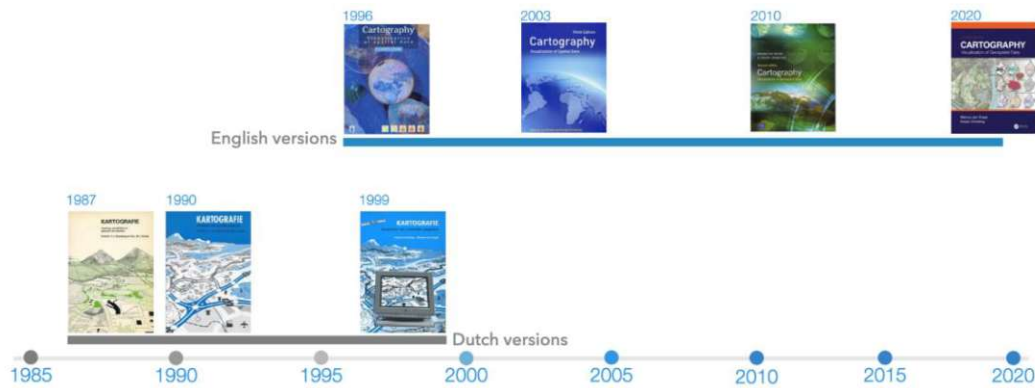


Figure 42. Timeline of *Cartography: Visualization of (Geo)Spatial Data* (Source: Kraak, Personal Communication).

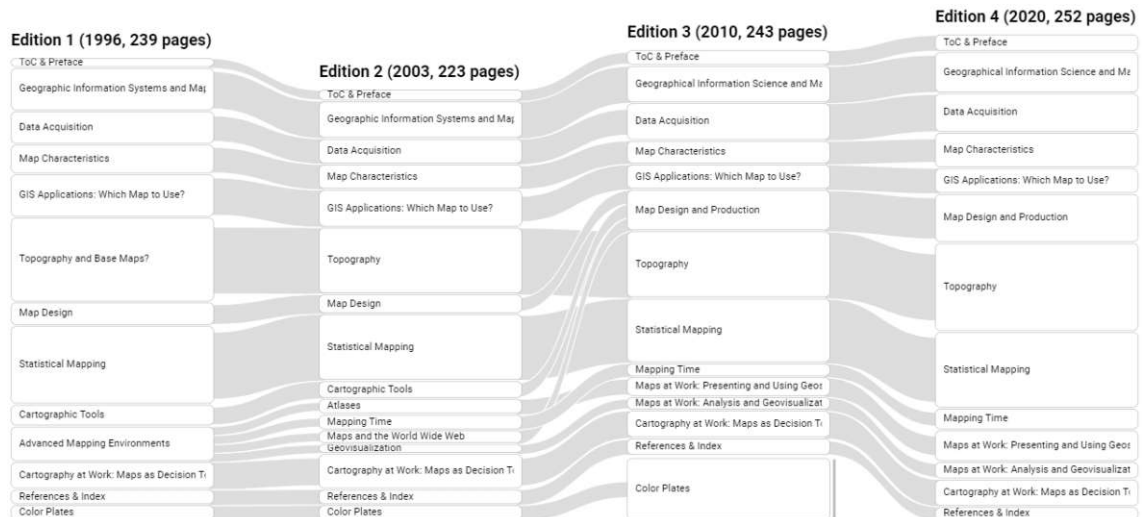


Figure 41. Flow diagram of structural changes of chapters in *Cartography: Visualization of (Geo)Spatial Data*.

In the preface of the first edition, the authors emphasised the changing environment from analogue to digital. They wrote that “We are currently experiencing an information revolution

which provides us all with computing power that we have never had at our disposal before” and that it “allows us to analyse large spatial data files at will” [KO1, vii]. Along with the changes that GIS would bring to cartography, they also predicted that the web “becomes a noticeboard and source for maps” [KO1, vii]. Also, they stated that the objective of the book is to provide “sufficient relevant knowledge of cartography for GIS users for the production and use of effective visualizations of spatial information” [KO1, vii].

In the second edition, they started to place greater emphasis on the web. The emergence of the web was described as “another revolution” that “boosted the possibilities for interaction and for querying the databases behind the maps” [KO2, viii]. Demonstrating how accessible databases have become via the web, they positioned the map as an important interface of this new data distribution environment. Furthermore, they wrote that “If map-making with GIS mainly involved geo-professionals, the World Wide Web potentially allows everyone having access to this new medium to create maps” [KO2, viii]. Consequently, the objective of the book was updated as below, including new terms such as geovisualisation as well as the web.

So the focus of this second edition has changed into how to provide sufficient relevant knowledge of cartography/geovisualization concepts and techniques to those accessing the World Wide Web for the production and use of effective visualizations of geospatial information [KO2, viii].

This is also reflected by the introduction of the new chapter «*Maps and the World Wide Web*», which is largely derived from the chapter «*Advanced Mapping Environments*». Whereas novel technologies such as electronic atlases, multimedia, and animation were covered under this chapter in the first edition, it diverged into four independent chapters in the second edition, including «*Maps and the World Wide Web*». The authors described that these chapters would belong to maps for exploration, in contrast to the earlier ones that are maps for communication.

The third edition—where it was described that “the position of the web has strengthened and stabilized itself” [KO3, viii]—went through some restructurings. The largest change was the disintegration of the chapters «*Maps and the World Wide Web*» and «*Cartographic Tools*» which were merged with map design principles and now formed a single chapter «*Map Design and Production*». Also, they labelled the advent of online mapping systems (e.g., Google Earth, Google Maps, Microsoft Virtual Earth) as the largest impact, saying that “It is their worldwide detailed coverage and the ability for users to add their own data that form the success factors of these products and that addicted so many people to maps” [KO3, viii].

The overall structure of the fourth edition was unchanged since the third edition and all changes took place in subchapter levels. The authors did not mention GIS anymore, except for the single sentence that contrasts GIS and the web. Maintaining a heavier emphasis on the web, their viewpoint from the second edition was maintained that maps have an important function as a tool for exploring databases which have become available via the web. However, the authors started to emphasise the map design aspect of the book. Here, they newly identified two objectives, which are “to provide an overview of the role that maps will play both today and in the near future in the world of geospatial data handling” and “to teach map design” [KO4, ix]. The evolution of the objectives of Kraak & Ormeling’s book is shown in *Table 21*.

Table 21. Comparison of objectives of the book from prefaces of *Cartography: Visualization of (Geo)Spatial Data*.

[KO1, vii]	[KO2, viii] & [KO3, viii]	[KO4, ix]
(...) to provide sufficient relevant knowledge of cartography for GIS users for the production and use of effective visualizations of spatial information.	(...) to provide sufficient relevant knowledge of cartography/ geovisualization concepts and techniques to those accessing the World Wide Web for the production and use of effective visualizations of geospatial information.	(...) to provide an overview of the role that maps will play both today and in the near future in the world of geospatial data handling. (...) teaching map design.

The results of the analysis are explained in three subsections; 4.2.1 shows the result of documented changes along the editions, 4.2.2 analyses the changes in web resources, and 4.2.3 summarises the information derived from the text mining.

4.2.1 Categories from Content Analysis

Here, the results from the content analysis of *Visualization of (Geo)Spatial Data* are exhibited. This section is written by interweaving the documented changes from the books and complementing some parts with author interviews. Initially, the documented changes were coded into a total of 22 codes. Then, in order to show the changes in an organic composition instead of fragments, multiple codes are grouped into one category. Each subsection from 4.2.1.1 to 4.2.1.6 is representative of a category.

Every subsection starts with an overview where the category and the involved codes are explained. Afterwards, the documented changes are exhibited in three parts: changes between the first and the second edition, the second and the third, and the third and the fourth. Finally, a summary is provided to review the written changes and to extract meanings from them.

4.2.1.1 GDI and Geodatabases

Overview. Many examples introduced in Kraak & Ormeling's books are online databases that distribute geodata. Thus, the term geodatabase is used to embrace various services such as national geoinformation systems, international mapping projects, websites for toponyms, et cetera. Another frequently used term is geodata infrastructure (GDI) and spatial data infrastructure (SDI). Since they are used interchangeably and the authors did not distinguish those two, the term GDI will be uniformly used. Codes found in this category are *GDI*, *web*, *GIS*, *database*, and *international*. Since almost all chapters of the book are involved in this category, a diagram showing relevant chapters is omitted.

Changes between the first and the second edition. The second edition is where the concept of geospatial data infrastructures (GDI) first appears. Diagnosing that GDIs gave maps an

additional role in geospatial analysis, the authors wrote that “These new infrastructures for accessing geospatial data are being developed all over the world in order to allow access to the geospatial datafiles created and maintained in order to monitor the population, resources and environment” [KO2, 1]. With the advent of GDI, the title of the subchapter *<The Relation Between GIS and Cartography>* was replaced with *<The Relation Between GDI and Cartography>*. In contrast to the first edition where they justified how important maps are in GIS environments, the focus moved on to how beneficial utilising GDIs could be in geospatial analyses.

The authors emphasized the enhanced interoperability and compatibility that GDIs have incurred. This is enabled by two components: clearinghouses and geospatial warehouses. Clearinghouses deal with metadata to make sure that different datasets are compatible with each other and that there are no duplicates. Geospatial warehouses are repositories that store actual datasets as well as metadata where users can access and download the data. Here, electronic navigational charts were provided as an example of geospatial warehouses.

Also, they pointed out the internet’s potential for “querying the databases behind the maps” [KO2, 3]. For example, they prospected that the web can greatly enhance geospatial analysis using GDIs. Mentioning “the possibilities of applying GIS processing on the web to all the data that have currently been made available there through the GDIs” [KO2, 15], they evaluated the benefit of data sharing and geospatial analysis that can be performed on the web.

Updates are also observed in other sources of geodata in the private sector. For example, in the subchapter *<Working with Digital Data>*, contrary to the first edition which did not include any examples, database management systems that offer both vector and raster data were introduced. They are ESRI’s Spatial Data Engine, Oracle’s Geospatial, IBM’s DB2 Spatial Extender, and Informix Spatial Database. Also, in the first edition, ArcData by ESRI was introduced as an online geodatabase along with a CD-ROM-based German topographic map Arc-Deutschland’ 500. In the second edition, they were replaced by Geography Network, a new online platform launched by ESRI where users can download and share geodata.

A new subchapter *<Management and Documentation of Spatial Information>* was introduced where the authors explain what kind of information should be annotated in the map margin and the metadata form. The process was described both in an analogue setting (e.g., index cards) and in a digital setting. As a guideline to document digital geodata, they mentioned the International Standard for Bibliographic Description for Cartographic Material (ISBD-CM) rules. Furthermore, they mentioned ongoing projects that seek to promote the exchange and integration of geodata such as Global Spatial Data Infrastructure (GSDI) and Oddens’ Bookmarks. GSDI is an initiative to construct a global clearinghouse and Oddens’ Bookmarks is a portal that provides 19,000 links related to maps and cartography.

In the subchapter *<Topographic Data: Mapping and Charting Organizations>* of the first edition, Kraak & Ormeling covered topographic mapping activities taking place in different countries: the US, the UK, the Netherlands, and Germany. In the second edition, the former two were complemented and updated whereas the latter two were replaced by a European and a global project. The first edition showed an ongoing project in the Netherlands to construct a 1:10,000

topographic database with the DLM called TOP10vector. It contains “basic topography such as infrastructure, land use categories, and buildings” and “topography, administrative boundaries and road centre lines” [KO1, 110~112], which are planned to be regularly updated. Similarly, the Amtliches Topographisches Kartographisches Informationssystem (ATKIS) from Germany was covered as an effort to build a centralised national database in three scale levels. They described ATKIS as “the first topographic information system that introduced the terms ‘digital landscape model’ (DLM) and ‘digital cartographic model’ (DCM)” [KO1, 113]. In the second edition, they were both replaced with Global Map and the Seamless Administrative Boundaries Europe (SABE).

Global Map is a project stemmed from the UN Environment & Development Conference in Rio in 1992, to “provide easy and open access to globally consistent geographical data for research and decision-making related to global change monitoring and environmental analysis” [KO2, 93]. It aims to produce 1 km-resolution raster data and 1:1,000,000 scale of vector data consisting of eight map layers: boundaries, drainage, transportation, elevation, population centres, land cover, land use, and vegetation. Here, mentioning that there are only a few countries that released their data, the authors demonstrated the difficulty of constructing a global geodatabase, which is “due to many arguments related to national security, political sensitivity, technology, economics and standards” [KO2, 93].

Created in 1991 for census purposes, SABE is a database of European administrative boundaries at various levels. The second edition showed the 1997 version of this database available on a scale of 1:100,000 containing “the geometry and semantics (names, codes, etc.) of all administrative hierarchies of 27 European countries” [KO2, 94]. They also introduced some of the GIS operations applied in the map construction such as line generalisation and edge matching for international boundaries and coastlines. Regarding the replacement of the TOP10vector and ATKIS by Global Map and SABE, Kraak explained in the interview that it was a combination of reasons; The first edition focused on European countries because they were transferred from the Dutch version *Kartografie* which was intended for Dutch students, whereas they decided to incorporate more international mapping databases since the second edition. Also, Kraak said in the interview that “TOP10vector is not used anymore” and “ATKIS was equivalent for this development during the time, but it also became outdated”.

Unlike the Netherlands and Germany, the examples from the US and the UK nominally survived until the second edition. However, their contents were largely reformed; For example, when they covered the United States Geological Survey (USGS) in the first edition, the focus was on their map series distributed as TIGER files and Digital Line Graphs (DLGs). TIGER is a format “produced in order to collect and visualize 1990 census data” and DLGs are “created by manually or automatically digitizing existing map sheets” [KO2, 113]. And in the last chapter where ongoing mapping projects are showcased, the authors mentioned that the USGS is developing the 1:24,000 topographical database of the entire country, which is “supposed to be completed around the year 2000” [KO1, 213]. This was actualised by the time of the second edition and the authors replaced the entire part describing the map series of the USGS with two sentences; “The major map series (over 55,000 sheets) is published at scale 1:24,000. However, these maps are on

average over 20 years old” [KO2, 94]. Instead, the section now introduced the recent blueprint of this plan ‘The National Map’. They wrote that it will consist of “high-resolution digital orthorectified imagery, high-resolution surface elevation data, vector feature data for hydrography, transportation, structures and boundaries as well as geographical names and land cover data” [KO2, 95]. The authors also annotated that this is rather an ambitious project and that the USGS is currently running several pilot projects.

The corresponding organisation in the UK is the Ordnance Survey (OS), covered both in the first and the second edition. The first edition introduced the modern history of topographic mapping by the OS; Whereas it started by digitising existing maps as drawing files, the growth in GIS demanded the maps be topologically structured, resulting in the necessity to build its own database. Here, the authors described the dilemma of the organisation during the transition period of analogue and digital.

If the Ordnance Survey intended to sell these data, they should meet the needs of the customer. However, an increase in digital sales would probably decrease the demand for paper maps. At that time, it was decided that 1995 would be a critical year since potential customers had indicated that this would be a deadline; If the data were not available by then, these customers would start their own mapping activities [KO1, 116].

As a result, while constructing a digital system that would “provide an infrastructure to collect, store, manipulate and supply structured topographic data” [KO2, 116], the OS also launched a project called Superplan which lets people purchase an on-demand paper map.

In the second edition, all of these descriptions were deleted except for the short recapitulation of the history of the OS’s map digitisation. The focus was newly set to its new database ‘the OS Mastermap’, a large-scale digital map covering the entire country with 400 million features that are equipped with a unique identifier number. Here, the authors focused on the advantages of its database hosted via the web with various information, “ranging from extensive background information on the organization, including its history, to copyright rules” [KO2, 97]. They showcased the functionalities that the OS Mastermap offers, differentiated by user groups such as business users, children, university students, researchers, et cetera.

Changes between the second and the third edition. In the third edition, the definition of GDI was provided for the first time as below.

A GDI can be defined as (Groot and McLaughlin, 2000) ‘A set of institutional, technical and economical arrangements, to enhance the availability (access and use) for correct, up-to-date, fit-for-purpose and integrated geo-information, timely and at an affordable price, with the goals to support decision-making processes related to countries' sustainable development’ [KO3, 17].

The authors started to prefer the term spatial data infrastructure (SDI) to GDI, which resulted in renaming the subchapter to *<The Spatial Data Infrastructure and Maps>*, even though both terms were still used interchangeably. While most of the content remained unchanged, they introduced INSPIRE—a GDI initiative by the European Union—and the Open Geospatial Consortium (OGC). OGC develops protocols which “describe interfaces and encodings that allow so-called geoservices to operate” [KO3, 17], ensuring that geodata is provided in a standardised manner.

INSPIRE follows the standards from OGC along with specific guidelines such as data security, data accessibility, reusability, et cetera.

In terms of geodatabases, the authors indicated that the Digital Chart of the World (DCW) is the precursor of Global Map. Produced in 1990 by the US military mapping agency, DCW is a global scale map at the scale of 1:1,000,000 containing hydrography, territorial boundaries, coastlines, and administrative names. Here, the authors wrote that “DCW is exchanged for Global Map, to make available topographic files on a country basis with a similar resolution as DCW but more recent data and a more extensive thematic coverage” [KO3, 23]. They also added that now, more countries are participating in Global Map. Whereas it was pointed out in the second edition that only a few countries had released their data due to various reasons, now they wrote that “Up till 2009 over 60 countries had made their data available” [KO3, 113]. Global Map was also newly added to the part where small-scale maps are introduced, next to the International Map of the World and Karta Mira that were continued since the first edition. On the other hand, the example of Oddens’ Bookmarks—a portal that provides 19,000 links related to maps and cartography—was removed. Regarding this change, Kraak said in the interview that “Oddens was a person that was in charge of the map library in Utrecht University, and it was impossible to maintain [the website with numerous links that are constantly updated]”.

Another noticeable update is the addition of the subchapter *Geographical Names*. Regarding this update, Kraak explained that it was a combination of changes in the field and personal interest; First, they became more important because a lot of online databases containing geographical names have become available. Second, Ormeling is an expert in geographical names. In this chapter, they covered the issue of dealing with names of places that should be standardised to ease the communication and the mapping process. They introduced some organisations and projects that are in charge of the standardisation such as the United Nations Group of Experts on Geographical Names (UNGEGN) and EuroGeoNames.

Changes between the third and the fourth edition. The fourth edition did not experience many updates in terms of GDI. One change is the advent of the word ‘geodata infrastructure portal’. Whereas the internet was mentioned as the medium of acquiring geospatial data in the second and the third edition, the fourth edition started to refer to geodata infrastructure portals.

Table 22. Comparison of descriptions regarding the acquisition of geodata in *Cartography: Visualization of (Geo)Spatial Data*.

[KO2, 25]	[KO3, 22]	[KO4, 25]
(...) GPS data that are produced only once and are therefore much more vulnerable. Nowadays, files with geospatial data can also be acquired through the Internet or the World Wide Web.	GPS data can be distributed on DVD, but are mainly available via download through the Geodata infrastructure. Nowadays, files with geospatial data can also be acquired through the Internet or the World Wide Web.	(...) GPS data that are distributed mainly through the geodata infrastructure portals.

In terms of geodatabases, ESRI's Geography Network—introduced in the second edition replacing ArcData and maintained until the third edition—was updated with their new system 'Living Atlas'. Whereas the main function of accessing and sharing geodata remained the same, the availability and the variety of data were significantly improved. For example, the earlier version mentioned that “mainly North American organizations offer their data for free” [KO2, 91]. On the other hand, Living Atlas includes “generic global datasets for base maps with and without hill shading, boundaries, satellite images from Sentinel-2 and others, as well as aerial photographs from all kinds of regions” [KO4, 128].

Additional geodatabases generated on a global level were added such as OpenStreetMap (OSM) and Natural Earth. OSM is frequently mentioned throughout the edition as an example of free, global-scale, and collaborative web map services. Here, they focused on how geodata is crowdsourced “via digitization of satellite imagery or real photographs, by adding GPS track collected by cycling and walking the terrain, and updating attribute information” [KO4, 135]. OSM replaced Global Map—introduced in the second edition as a global topographic database and maintained until the third edition, because the project ended in 2018 after being transferred to the UN. The authors wrote that “DCW type of data is often replaced by OSM” but also pointed out the shortcoming of OSM data that “height information is lacking in OSM data” [KO4, 220].

Natural Earth is a free website that “provides free vector and raster map data at 1:10, 1:50 and 1:110 m scales in ESRI shapefile and TIFF format” [KO4, 84]. The other geodatabases added in the fourth edition include the Data Distribution Centre (DCC) of the Intergovernmental Panel on Climate Change (IPCC) and the Corine Land Cover from the European Union.

They also updated the guidelines of metadata. Whereas ISBD was still maintained as an international standard, its link to cartography was deleted. Instead, they added a new standard: Descriptive Cataloguing of Rare Materials – Cartographic (DCRM-C). Similarly, a follow-up was provided to the GSDI initiative, that there is now “a group of experts that on behalf of the United Nations support the Global Spatial Data Infrastructure” [KO4, 220]. In the part where they introduced data formats that follow ISO standards and the OpenGIS consortium, GeoJSON was newly added next to Geographic Markup Language (GML).

Furthermore, a recent feature provided by the Ordnance Survey was added to their description. While the majority of the description about the OS MasterMap—its topographic database website launched in 2001—remained unchanged, the fourth edition added its new functionality 'OS Open Zoomstack'. It provides a single map of the UK that can be customised and downloaded based on different user needs. Furthermore, they mentioned that its classic paper maps are “still popular and are available as an app on mobile devices with interactive functionality” [KO4, 139].

In terms of toponyms, they added a new website based on the data of the National Geospatial Intelligence Agency (NGA) of the US called GeoNames. Here, they showed the functionalities of the website where users can search for toponyms worldwide and retrieve information about a place such as coordinates, height above sea level, et cetera. They also updated the paragraph concerning the EuroGeoNames project that its service is no longer continued, however, the

regional dictionary can be downloaded since April 2019. The continuities of geodatabase examples are shown in *Table 23*.

Summary. The rapid growth of the internet has sparked the development of geodatabases to distribute geodata. They are a central concept in Kraak & Ormeling's books in understanding the evolution of digital cartography since the authors focus on sources and the use of geodata throughout the book. Changes observed regarding the USGS and the Ordnance Survey show how the national organisations have adapted to the changing environment and needs. Also, descriptions in each edition about international projects such as SABE and Global Map provide a snapshot of how global databases are designed and implemented. The third edition can be seen as a watershed in terms of contemporary geodata management; It defined GDI as well as introduced OGC standards and the European Union initiative INSPIRE. Along with international guidelines about geodata management, the fourth edition added more examples of private and open-source geodatabases such as ESRI's Geography Network, OSM, and Natural Earth. As a result, it can be observed with the progress of edition how online sources of geodata have been diversified and democratised, supported by initiatives and policies that foster the standardisation and the exchange of geodata.

Table 23. Continuities of geodatabase examples in *Visualization of (Geo)Spatial Data*.

1st ed. (1996)	2nd ed. (2003)	3rd ed. (2010)	4th ed. (2020)
NL TOP10vector			
DE ATKIS			
USGS TIGER & DLG	USGS The National Map		
UK Ordnance Survey (OS)	UK OS Mastermap		
DCW	Global Map		
			OpenStreetMap
			Natural Earth
	SABE		EuroBoundaryMap
	GSDI		GGIM
ESRI ArcData	ESRI Geography Network		ESRI Living Atlas
ESRI Arc-Deutschland' 500			
	Odden's Bookmarks		
		UNGEKN	
		EuroGeoNames	
			US NGA GeoNames

- National topographic databases
- International mapping projects and geo-portals
- Toponym databases

4.2.1.2 Electronic Atlases

Overview. In all four editions, electronic atlases were covered in two subchapters: *Electronic Atlases* and *Working with (Web-Based) Electronic Atlases*. Since the latter subchapter has been barely updated throughout the editions, it is excluded from the analysis. Also, except for the first edition, the other three editions had a subchapter *Paper Atlas* dedicated to printed atlases, however, its content also remained unchanged and had little connection to digital technologies. As a result, this subsection will only concern changes in the subchapter *Electronic Atlases*. Codes found in this category are *atlas*, *multimedia*, *animation*, *web*, and *database*.



Figure 43. Flow diagram of chapters from *Cartography: Visualization of (Geo)Spatial Data* that are involved in the formation of the category *Electronic Atlases*.

Changes between the first and the second edition. In the first edition, electronic atlases were introduced under the chapter *Advanced Mapping Environments* along with animation, multimedia, and ESDA. The authors defined an electronic atlas as “computerized GIS, related to a certain area or theme in connection with a given purpose, with an additional narrative faculty in which maps play a dominant role” [KO1, 185]. They classified them into three types: view-only, interactive, and analytical. View-only electronic atlases are digitised versions of paper atlases without interactive functionalities and two examples are introduced; the Atlas of Arkansas and the Atlas of Florida. On the other hand, the interactive ones let users manipulate the datasets, with the example of electronic atlases of France produced by RECLUS. Finally, analytical electronic atlases allow the most interaction, as well as manipulation and “much of the GIS functionality would be available here” [KO1, 185], such as querying.

While most of the descriptions remained unchanged, the second edition updated the examples; The Atlas of Arkansas, the Atlas of Florida, and the electronic atlases of France were deleted. Instead, the National Atlas of Germany, the Atlas Mittleres Ruhrgebiet, and the Atlas of Switzerland Interactive were newly added. Also, whereas they introduced Software Toolwork’s World Atlas as an example of forging multimedia functionalities in the first edition, the second edition replaced this with Microsoft Encarta World Atlas. The continuities of electronic atlas examples are shown in *Table 24*.

Moreover, the second edition added two new sections—*Atlas Information Systems* and *Web Atlases*—under the subchapter *Electronic Atlases*. Atlas information systems refer to a set of various datasets that are connected with each other, and web atlases are the ones that are accessible via the web. The authors compared the pros and cons of three common media: paper, CD-ROM, and the web. Noting that the web atlases can overcome shortcomings of atlases disseminated by paper and CD-ROM, two of their pros were demonstrated; the automatic update through linked data servers, and their ability to geographically order data. However, the authors were also aware of the technical limitations of the web atlases of the time.

(...) and here is the strength of the WWW: it can be updated continuously, although it would not always be easy to access if too many people decided to visit the atlas site simultaneously. (...) At the same time the resolution of the information in the website would not be as fine as that of the book or the CD: it would take too much time to load the updated images otherwise [KO2, 158].

Changes between the second and the third edition. The section ‘*Atlas Information Systems*’ was removed and the word CD-ROM was replaced with DVD. All examples concerning electronic atlases remained the same, except for the addition of Microsoft Visual Earth which was described as “intentional combinations of specially processed spatial datasets, together with the software to produce maps from them” [KO3, 163].

Changes between the third and the fourth edition. The fourth edition revised the types of electronic atlases; The third type ‘analytic’ was replaced with ‘interactive and integrated into a spatial data infrastructure’. Whereas the ‘analytic’ was merely an extension of ‘interactive’, the new type accommodates maps that are linked to external data servers. They also introduced another criterion of classification: ‘stand-alone electronic atlases’ versus ‘web-based’, resulting in a total of six possible combinations. Except for the reclassification, all of the examples have been covered at least once in previous editions and their description did not change significantly when compared to the third edition.

Since web atlases have now become the counterpart to stand-alone electronic atlases, the updates in the fourth edition were mainly about the description of web atlases. The authors maintained their stance about the limitation of web atlases that “as it would take more time to download information as compared to when taken from a DVD, the maps need to be less complex in order to minimize their transfer time” [KO4, 198]. After introducing the National Atlas of the United States as an example of web atlases connected to the spatial data infrastructure, the authors demonstrated the potential of this linkage by listing possible applications. They also noted that there are ongoing projects in Spain and Ukraine to realise this as of 2020.

(...) the atlas soil map links to the national soil mapping programme and detailed soil maps of the area centred on; climate maps link to the actual weather situation, a higher education participation map links to the relevant statistics for the clicked-on municipality; highway system maps link to the actual traffic density map as monitored by the relevant sensor system [KO4, 198].

Summary. In Kraak & Ormeling's books, electronic atlases have been largely covered in the context of multimedia and interactive functionalities. In the third and the fourth edition, it can be observed how their capabilities have been furthered by the internet. The third edition noted the possibility of hosting atlases online, which was confirmed by the fourth edition that additionally classified the examples as electronic versus web-based. Also, the introduction of the new category ‘interactive and integrated to the spatial data infrastructure’ shows the trajectory of digital technologies regarding map use; Whereas the earlier editions focused on functionalities that computer-based systems can offer, the development of geodatabases has opened new possibilities of accessing and manipulating the original data using atlases. As a result, the changes observed in electronic and web atlases provide historical insights through the intersection of map products, computers, and the web.

Table 24. Continuities of electronic atlas examples in *Cartography: Visualization of (Geo)Spatial Data*.

1st ed. (1996)	2nd ed. (2003)	3rd ed. (2010)	4th ed. (2020)
Atlas of Arkansas			Atlas of Arkansas
Atlas of Florida			
Electronic Atlases of FR			
	National Atlas of DE	National Atlas of DE	National Atlas of DE
	Atlas Mittleres Ruhrgebiet	Atlas Mittleres Ruhrgebiet	
	Atlas of Switzerland	Atlas of Switzerland 2	Atlas of Switzerland 3
		Atlas of Public Health NL	Atlas of Public Health NL
			National Atlas of the US

	Stand-Alone Electronic Atlases	Web Atlases
View-only	■	■
Interactive	■	-
Interactive & Integrated	-	■

4.2.1.3 Animation, Multimedia, and Web Mapping

Overview. In the first edition, animation and multimedia were covered under the chapter «*Advanced Mapping Environments*», along with electronic atlases and ESDA. In the second edition, they were relocated to two separate chapters: «*Mapping Time*» and «*Maps and the World Wide Web*». Whereas the chapter «*Mapping Time*» remained stable until the fourth edition, the other chapter was discontinued in the third edition and its content was relocated to «*Map Design and Production*». Codes found in this category are *animation*, *multimedia*, *GIS*, *user*, and *web*.



Figure 44. Flow diagram of chapters from *Cartography: Visualization of (Geo)Spatial Data* that are involved in the formation of the category *Animation, Multimedia, and Web Mapping*.

Changes between the first and the second edition. In the first edition, Kraak & Ormeling emphasised the necessity of animated maps that “the GIS environment need to deal with processes as a whole, and no longer with single time slices” [K01, 191]. After introducing three different methods to generate animations, they classified the animation into two: temporal and non-temporal. Here, except for the fact that animated maps are generated by computer-based systems, there are few links to digital technologies. Multimedia, on the other hand, has more frequent mentions of necessary digital technologies. The authors covered individual multimedia

components—sound, text, (video) images, animations, and hypermaps—focusing on each component's relation with maps and how they are realised in computer-based systems. Also, they envisioned how the use of multimedia can add value to maps.

Since multimedia equipment can produce music of CD quality one can easily imagine what a multimedia composers' map of Europe would look like. Pointing to Beethoven's place of birth would activate his ninth symphony, show his picture and music sheets as well as a video of a landscape that visualizes the atmosphere of the music [KO1, 189].

Although the overall description of animations and multimedia were mainly unchanged in the second edition, the influence of the web has become more pronounced. For example, it was added that the web has eased the distribution of animation because “media players integrated in the web can be used to run the animations” [KO2, 167]. Similarly, it was added that “the WWW is an ideal platform to combine different multimedia elements with maps” [KO2, 171]. Also, hypermaps—defined as georeferenced multimedia—were unlisted from the components of multimedia and were incorporated into the section ‘*Animations*’. Here, the authors emphasised how the web has allowed the user “to navigate the data, both through geospace and through different topics” [KO2, 172]. In addition, they added that special plug-ins (e.g., Macromedia’s Shockwave) and programmes created by Java or Javascript are required for advanced interfaces of animation.

The increased relevance of the web is further confirmed by the introduction of the new chapter <<*Maps and the World Wide Web*>>. While the second subchapter <<*Web Maps and Multimedia*> inherited most of the content from the first edition’s <<*Maps and Multimedia Systems*>, the first subchapter <<*Web Map Design*> was newly written. The authors emphasised that “these [map design] basics do not change because of the new web environment” [KO2, 168], however, they also provided additional possibilities and constraints in web mapping. Their focus was on interactivity and multimedia, such as showing additional information via clicking or hovering, filtering certain regions on a map, and using hyperlinks. The authors also wrote that the web has enabled the use of new visual variables: blur, focus, and transparency. Furthermore, they introduced the programming language VRML, which can “display three-dimensional maps that can be looked at from any viewpoint” and “allows viewers to manipulate the model in three-dimensional space” [KO2, 170].

Changes between the second and the third edition. The subchapter <<*Web Maps and Multimedia*> of the third edition newly included the cognitive limitation of animations; that users might fail to recognise changes when too many changes are displayed simultaneously. Also, whereas Shockwave, Java, and Javascript were mentioned in the second edition as additional tools required for advanced interface options, they were replaced by Adobe’s Flash and Scalable Vector Graphics (SVG). Also, the authors started incorporating external websites that can handle animations and multimedia. For example, they added a sentence that “cartographic animations are even found on a website like YouTube” [KO3, 157] as well as mentioning Flickr, a website where geotagged photos are uploaded and shared. On the other hand, the other subchapter <<*Web Map Design*> remained unchanged.

Changes between the third and the fourth edition. The fourth edition described how more widely animations have become available using GIS software and the web. Here, the interactive animation that visualises the history of topographic mapping by the National Mapping Agency of the Netherlands was introduced as an example. Flickr from the third edition was replaced with Encarta Interactive World Atlas Online as an example representing how multimedia can be applied in cartography. On the other hand, some parts were deleted; The paragraph describing the multimedia map about Beethoven as well as hypermaps were removed. Furthermore, they deleted the sentence describing additional tools that are required for advanced interfaces in animation and replaced it with the possibility of displaying interactive animations on a browser using HTML5 and D3.

Table 25. Comparison of descriptions regarding the realisation of animation in *Cartography: Visualization of (Geo)Spatial Data*.

[KO2, 167]	[KO3, 157]	[KO4, 191]
For advanced interface options, special plug-ins such as Macro-media's Shockwave or interactive programs created in Java or Javascript are still needed.	Cartographic animations are even found on a website like YouTube. For advanced interface options, special plug-ins such as Flash or Scalable Vector Graphics (SVG) are still needed.	Cartographic animations can be distributed via websites such as YouTube or Vimeo. Interactive animations can also be displayed in a browser, using, for instance, a combination of HTML5 and D3.

In contrast, the subchapter *Web Map Design* experienced only two changes; First, they mentioned Leaflet as “one of the most popular open-source Javascript libraries for interactive maps” [KO4, 94]. Second, the description of the programming language VRML was deleted.

Furthermore, the new subchapter *Story Map* was added, specifically concerning the product StoryMaps by ESRI. Often referred to as visual storytelling, story maps combine maps, multimedia, and other visualisation techniques to create a narrative. Here, they briefly introduced major functionalities of StoryMaps such as map tour, swipe, series, et cetera. In a similar context, they also added a paragraph about dashboards which are defined as “a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance” [KO4, 217]. Focusing on how dashboards can assist communication through various visualisation techniques and interactivity, they showed an example created by Tableau software.

Summary. Descriptions of animation and multimedia in Kraak & Ormeling's books focus on classifying types and identifying each component. Since the second edition, they have acknowledged the potential of the web regarding the distribution of animation and conveying multimedia content. This is reflected by the introduction of the new subchapter *Web Map Design* where they demonstrated novel data exploration techniques that web maps can enable

such as hovering, filtering, and the usage of hyperlinks. In this regard, the authors mentioned plug-ins for advanced interface options in realising animation, which has changed along with the editions. They started with Shockwave, Java, and Javascript, and then moved on to Flash and SVG. Eventually, they were replaced by HTML5 and D3 (Javascript), programming languages that enable interactivity on the web. They also started incorporating social media and web-based platforms that contain animations and multimedia such as Youtube and ESRI StoryMaps. As a result, the changes observed in animation and multimedia are intertwined with the development of web mapping and interactive functionalities that the internet has made possible.

4.2.1.4 Map Production

Overview. Map production processes in Kraak & Ormeling’s books are mostly covered using GIS and graphic software. In the first and the second edition, they walk readers through the process in the chapter «*Cartographic Tools*». Since the third edition, the contents were merged with parts dealing with map design and formed a new chapter «*Map Design and Production*» with additional subchapters about web mapping. Although most of the contents concerning web mapping were covered above under 4.2.1.3, some are explained in the context of map production, and therefore, will be mentioned here. Codes found in this category are *design, GIS, software, web, file format, printing, and hardware*.



Figure 45. Flow diagram of chapters from *Cartography: Visualization of (Geo)Spatial Data* that are involved in the formation of the category *Map Production*.

Changes between the first and the second edition. Chapter «*Cartographic Tools*» of the first edition starts by listing five requirements a GIS software application should have. It should be able to “handle topographic and thematic data”, “create a legend, north arrow, scale bar, and map title”, “change the coordinate system, generalise topography, or classify attribute data”, and “send maps to raster or vector plotters”, and “provide a graphical user interface” [KO1, 172~173]. Here, the authors mentioned the current development of interactive user interfaces in GIS software, referring to them as ‘data-viewers’. After explaining each requirement, they covered the general process of map-making using GIS and desktop publishing (DTP) software. DTP software is equivalent to today’s graphic design programmes that “allow the user to drag the text to the correct position with the mouse, and even scale it interactively” [KO1, 174]. Finally, they covered five steps of map production—from the map design to the final printing.

While keeping the overall structure intact, the second edition inserted a new subchapter in the front concerning the management of geodata. Here, they introduced the standards and initiatives such as ISBD-CM, GSDI, and Oddens’ Bookmarks, all of which are covered under 4.2.1.1 in the context of the geodatabase.

Other updates in the second edition are mainly deletions which resulted in the overall decrease in the volume of the chapter. For example, two paragraphs explaining the ongoing development of interactive user interfaces were deleted. Also, in the following subchapter <Desktop Mapping>, they deleted the paragraph pointing out the poor interoperability between GIS and DTP software applications; that “it is not possible to produce a map with GIS software, transfer it to a DTP environment, enhance its graphics, return it to the GIS environment” [KO1, 175]. Whereas ‘Cart/o/Info’ and ‘Cart/o/Graphix’ were shown in the first edition as an example of possibly overcoming this interoperability issue, they were replaced by Microsoft’s MapPoint and the collaboration of MapPublisher, ArcView, and Freehand.

In terms of printing, they maintained the observation that “current printing technologies can print only one colour at a time for each printing plate” [KO2, 145]. However, a sentence was added that “today digital presses exist that do not require printing plates” [KO2, 146]. Consequently, descriptions of publishing packages that help separate colours of the map into each colour of CMYK were deleted, such as MapPublisher and MapFinisher. Similarly, other descriptions associated with printing plates or plotters were either deleted or reduced in size. Below are two deleted examples concerning plotters.

It may vary from 1,000 ECU [electronic control unit] for a simple black and white A4 laser printer to more than 50,000 ECU for a colour AO electrostatic plotter [KO1, 179].

It is not only Postscript that is used to activate printing and plotting devices. Another ad hoc standard is HPGL, the Hewlett Packard Graphic Language. Most laser printers can handle this language but, more importantly, all pen plotters are activated by it. Since pen plotters are vector-oriented, HPGL in its most simple form contains pen-down and pen-up commands (...) [KO1, 180].

Instead, they added a description of the PDF format which enables “to visualize the maps designed on any output medium” [KO2, 145]. The term computer graphics metafile (CGM)—a file format for storing maps—was also replaced by bitmap and PDF.

Changes between the second and the third edition. In the third edition, the chapter was merged with map design principles, resulting in a new chapter <<Map Design and Production>>. From here, the term DTP was replaced with ‘graphical software packages’.

Also, the part describing the interoperability of GIS and graphic software packages was largely complemented, with three ways of producing a digital map. The first is to work with a special GIS plug-in for graphics that allows the highest level of integration such as MapPublisher and Adobe Illustrator. In this case, the changes made in the graphic software are directly reflected in the original GIS database. The second option is to create the map using GIS, export it as pdf, tiff, or jpeg files, and edit it in the graphic software. Finally, they introduced the possibility of using a web browser to create maps. Here, they covered various formats and standards such as Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS), Scalable Vector Graphics (SVG), and Geographic Markup Language (GML).

Furthermore, manual digitizing started to be referred to as an old technology which is being replaced by scanners. After describing the working principle of a digitising tablet, they added

that this hardware is hardly used anymore. Some paragraphs that were used to delineate the process of digitising a complex analogue map were also removed in the third edition.

For ‘professional’ digitizing of large series of maps, configurations are used which are ergonomically adapted to the working attitudes, allowing for a minimum of tiresome bending over, while the digitized lines and points will be visualized immediately onscreen, so that the operator can check on-line whether he or she is following the marked information properly. (...) [KO1, 28]

Part of the preparation for digitizing of complex analogue maps also exist in highlighting the lines that have to be digitized with a marker, and of entering manually the codes that will be applied to the lines and to the areas they separate on the sheets beforehand. The direction in which the lines will be digitized will be determined and indicated as well (...) [KO1, 32]

On the other hand, the improved storage option of computers was mentioned, which can “store the original data and not the derived or aggregated data” [KO2, 35]. The authors demonstrated the usefulness of pertaining to the original dataset during the map production.

Changes between the third and the fourth edition. The extensive part describing the map production process using a graphic package with the example of the US population density map was deleted. Instead, a lot of content relevant to web mapping started permeating into non-web-related parts of the chapter. For instance, in the subchapter *Requirements for the Cartographic Component of GIS Packages*, the authors mentioned standards and formats relevant to web mapping such as HTML5, CSS3, SVG, and WebGL. Moreover, they demonstrated the advantages of using web-based mapping tools as below.

(...) with some scripting knowledge tools such as Mapbox and Kepler.gl, let you create advanced interactive cartographic visualizations, as well as by using libraries like OpenLayers. With even more scripting skills, the options using the data-driven visualization library D3.js are nearly unlimited [KO4, 84].

The explanation about using a web browser to create maps was also complemented with additional background information (e.g., World Wide Web Consortium (W3C) and OGC). Here, GeoJSON was added next to GML, both as examples of the WFS standard. Also, they mentioned the possibility of storing data on a cloud, which is introduced as “storing and accessing programs and data on the Internet instead of in one’s computer” [KO4, 37].

Summary. The map production process in Kraak & Ormeling's books is described using GIS, graphic software, and the web. Whereas printing hardware (e.g., plotters, printing plates, and digitisers) has been largely reduced, GIS and graphic software became dominant mapping tools. From the first to the third edition, it can be observed how they have been increasingly integrated and compatible with each other. Also, a web browser is introduced as another method of producing digital maps using WMS, WFS, and WCS. This content was further complemented in the fourth edition with web-based libraries and mapping tools such as Mapbox and D3. Also, the web has become not only a key platform to display maps but also an important tool to produce maps, blurring the boundary between web mapping and non-web digital mapping. As a result, the above changes show how digital mapping has been eased with the integration of GIS and graphic software, and how the dissemination of digital files replaced printed maps.

4.2.1.5 Remote Sensing Technologies and Applications

Overview. Kraak & Ormeling classify the data acquisition methods into nine categories. Among these, the changes have been observed dominantly in remote sensing and GPS data. While they are mainly covered in chapters <<Data Acquisition>> and <<Map Characteristics>>, there are also episodic mentions of them in other chapters, and these are incorporated here as well. Codes found in this category are *remote sensing*, *database*, *GIS*, *international*, and *web*.

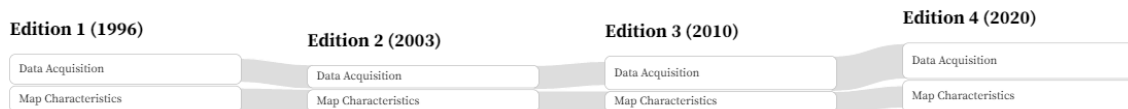


Figure 46. Flow diagram of chapters from *Cartography: Visualization of (Geo)Spatial Data* that are involved in the formation of the category *Remote Sensing Technologies and Applications*.

Changes between the first and the second edition. The second edition introduced new technology in photogrammetric surveys: laser altimetry. It constructs terrain models with an aircraft equipped with GPS receivers and a laser range finder which “allows for distance measurements up to 1 cm theoretical precision” [KO2, 20]. Also, in the chapter <<Map Characteristics>> where they introduce maps that are used for maintenance of ground objects, they showed how remote sensing can be applied to monitoring sea-dykes in the Netherlands.

After every major storm the dykes would be photographed again from the air, and these images would be photogrammetrically processed and turned into new maps, to be compared with the old ones in order to ascertain whether the coastlines had been damaged by sand or dunes being swept away. In those cases where this was indeed the case, the sand would be replenished as soon as possible [KO2, 44~45].

In the subchapter <<Terrain Visualization>> where they show how topographic maps can be used along with digital terrain models (DTM), a picture showing a topographic map draped over a DTM was added. Along with the picture, they annotated that “These days GIS software with digital terrain modelling capacity has draping functions available” to “create realistic three-dimensional views” [KO2, 88].

Changes between the second and the third edition. Noting that new satellites were launched with higher resolution and an increased number of bands, the authors introduced some examples. These are SPOT, Ikonos, Quickbird, GeoEye-1, Cartosat-3, and Aster. Without covering their detailed specifications, they listed each satellite’s resolution for panchromatic and multispectral applications. Also, later in the subchapter <<Geospatial Analysis in Geography: Use of Small-Scale Maps>>, they mentioned how images from the SPOT satellite are used in environmental management. With an example of the rural development plan in Guinea, they wrote that it can “identify the current agricultural land use as well as constraints on and potential of the remaining land” [KO3, 56].

Furthermore, they started to show the applications of remote sensing technologies and GPS data in Google Earth and Google Maps. Describing them as “a true revolution” where anyone can “display satellite data and maps for free on a level of detail that was not heard of before” [KO3,

18], the authors showed their possible use cases. For example, they wrote about how users can track their trajectories using GPS on Google Maps, and superimpose a map on the satellite imagery using Google Earth to perform geometric transformations and corrections.

Changes between the third and the fourth edition. The laser altimetry—introduced in the second edition as a new method of photogrammetrical surveys—was renamed LiDAR and was described as the general laser scanning method. Also, contrary to the second edition where they wrote that the sea-dykes in the Netherlands can be monitored using airborne photogrammetry, now it was replaced by LiDAR technology while the example remained the same. Furthermore, they included the recent digital terrain model of the Netherlands generated by laser scanning.

In terms of satellite technologies, three recent global positioning systems (GPS) were added: Galileo by the European Union, GLONASS by Russia, and BeiDou by China. Also, the case of the rural development plan in Guinea was replaced with Bhutan, with Sentinel-2 instead of SPOT, and the paragraph describing the resolution of various satellite programmes was deleted. This was replaced with the sentence “Today, the resolution goes beyond the one-metre resolution” [KO4, 27]. Here, they only briefly mentioned the Copernicus Programme of the European Union which offers satellite images for free and outsourced further information about other satellites to the ITC’s satellite sensor database, followed by a hyperlink. Furthermore, in the midst of constantly developing technologies related to data acquisition, they wrote that cartography can play an important role as below.

The ever more detailed satellite imagery available, the increasing number of sensor networks and new techniques for analysing textual sources with spatial references like geoparsing all lead to highly varied ‘big data’, (...) In order to make sense of them and derive meaning or trends, cartography, with its capacity of generalizing data in order to fit their purpose, plays an important role [KO4, 4].

Summary. The consistent development of remote sensing technologies is a noticeable change in Kraak & Ormeling's books. First, there are novel techniques such as laser altimetry that are introduced and later established as a standard technique. Second, the resolution and the number of bands of satellites have increased over time. For example, in the fourth edition, due to the large number of satellite products available and their frequent update, the authors outsourced further information about satellite sensors to the external website. In terms of the application, Google Maps and Google Earth are mentioned as tools where anyone can collect GPS data and use remotely sensed data for their analysis. As a result, they show how far remote sensing technologies have evolved and how relevant they have become in cartography with web-based tools that enable easier access and use of location-based data and satellite imagery.

4.2.1.6 Semantics and Policies

Overview. Kraak & Ormeling’s books possess a generalist approach to cartography and frequently mention definitions, international initiatives, and legal issues in cartography. Although they do not include actual maps, a lot of them are relevant to digital technologies, therefore, are covered in this subsection. Codes found in this category are *definition*, *copyright*,

international, web, and VGI. Since the codes involved in this category are covered throughout the book without being bound to several chapters, a diagram showing relevant chapters is omitted.

Changes between the first and the second edition. In the first edition, the authors provided the definition of cartography by Taylor (1991) as “the organisation, presentation, communication and utilisation of geoinformation in graphic, digital or tactile form. It can include all stages from data preparation to end users in the creation of maps and related spatial information products” [KO1, 43]. In the second edition, they offered their own definition next to Taylor’s. They defined cartography as “making accessible spatial data, emphasizing its visualization and enabling interaction with it, aimed at dealing with geospatial issues” [KO2, 35].

The second edition introduced a new subchapter *Copyright and Liability*. Here, the authors explained how copyrights can be a complex issue in maps because they involve various datasets from different sources. Not only does each of these actors have its own copyright policy, but also the resulting map as a cartographer’s creation could be subject to copyright. After introducing the concept of copyright, they cover related topics such as the copyright of geodata, exceptions, international differences, and the liability of distributing wrong data.

Changes between the second and the third edition. In the first chapter, of the third edition, the authors diagnosed that “the rise of Internet brought the next revolution of mapping” and “the IT related developments have resulted in a convergence of different disciplines working with geoinformation” [KO3, 1]. Since here, they started using the term ‘GIScience’ which includes various disciplines that deal with geoinformation. While the definition of cartography remained unchanged, they added the definition of a map provided by the International Cartographic Association (ICA), which is “A symbolised representation of a geographical reality, representing selected features and characteristics, resulting from the creative effort of its author’s execution of choices, and designed for use when spatial relationships are of primary relevance” [KO3, 41].

Also, a new subchapter *Map Use and Usability* is introduced. Here, the authors emphasised that usability has to be considered in a digital mapping environment and defined it as “the capability of the map to be attractive, understood, and used under specified conditions” [KO3, 187]. While noting that usability has moved beyond measuring the efficiency of user interfaces, they introduced several ways to evaluate the usability of maps. They showed how different user groups (e.g., tourists, children, mobile device users) can have different use purposes and approaches in interacting with maps and claim that the map design should always be user-centred.

Furthermore, they added a new subchapter *Geovisual Analytics* under the chapter *Maps at Work: Analysis and Geovisualization*. Pointing out that the traditional geovisualisation methods are not sufficient for handling emergent global issues (e.g., climate change, health, energy), they justified the need for combining visualisation with analytics. After introducing the basic concept of visual analytics provided by the National Visualization and Analytics Center (NVAC) in the US, the authors demonstrated how they can be used in cartography, with the new term ‘geovisual analytics’. They wrote that geovisual analytics’ function is to “find, assimilate, and analyse continuously changing data about time-critical, evolving real-world situations” [KO3, 174].

Furthermore, they showed how the geovisual analytics process can work with the example of a running data record with locations and heart rate values.

Changes between the third and the fourth edition. Now they contained four definitions of cartography from four sources: Kraak & Ormeling, Taylor, ICA, and Guptill & Starr. Definitions from Taylor and themselves—introduced in the second edition—were still unchanged, but they updated the ICA’s older definition with the 2011 version that it is “the art, science and technology of making and using maps” [KO4, 44]. They also mentioned that the change in the definition of cartography “came about by the rise of the computer and GIS” [KO4, 44], and introduced the third definition by Guptill & Starr which reflects this paradigm shift.

Cartography is the information transfer that is centred about a spatial database which can be considered in itself a multifaceted model of geographical reality. Such a spatial database then serves as the central core of an entire sequence of cartographic processes, receiving various data inputs and dispersing various types of information products (Guptill & Starr, 1984) [KO4, 44].

Another change in the fourth edition is the introduction of the volunteered geographical information (VGI). The authors equated VGI with citizen science and wrote that “as fieldwork becomes increasingly expensive, mapping or nature conservation agencies become more and more dependent upon the information received from citizens” [KO4, 24]. VGI was mentioned frequently throughout the fourth edition, especially in the context of OpenStreetMaps (OSM).

(...) The volunteers use recent satellite imagery to, for instance, indicate where the damage is most severe, and the new maps, often part of the OSM, can be used to restore the basic infrastructure [KO4, 31].
 (...) the OSM initiative is now linked to other volunteer activities like the standards of the OGC, whereas commercial GIS programs usually have their own global coverages [KO4, 202].

Furthermore, in the subchapter *Copyright and Liability*, they added a new section ‘Creative Commons Licenses’. Pointing out that the perception about the copyright of digital data has been changed, they mention Creative Commons. While open data “can be freely used, reused, and redistributed by anyone” [KO4, 227], they are still subject to some constraints. These constraints are defined by the Creative Commons license, with four types of redistribution rights.

Summary. The authors have introduced a total of four definitions of cartography from different sources. While each of them has its own focus, they have all emerged during the rise and growth of the digital environment, therefore, all contain mentions about information technology. Also, along with the editions, they added new concepts and subfields that are relevant to cartography such as copyright, usability, visual analytics, and VGI. Similarly, all of these newly included concepts are closely related to digital production and dissemination of geodata. As a result, the above changes show how definitions and neighbours of cartography have changed in response to the increased relevance of digital technologies and their usefulness.

4.2.2 Evolution of Web Resources

In the first and the second edition, Kraak & Ormeling's book did not contain any hyperlinks. An exception is the preface of the second edition where they introduced a website built for providing all websites covered in the book and keeping them up-to-date. However, the preface does not count as the main body text, therefore, the first and the second edition are both counted as having zero hyperlinks. During the interview, Kraak said that they “have always had the idea of including the hyperlinks”, but since it is impossible to update them inside the textbook, they outsourced them to a separate website in the second edition.

On the other hand, the third edition included 20 hyperlinks, half of which belonged to two chapters «*Topography*» and «*Maps at Work: Presenting and Using Geospatial Data in Maps and Atlases*». Regarding the question about the decision to include hyperlinks inside the book, Kraak answered that “We thought we could keep relatively stable websites such as national mapping agencies” and “maps that come with interactive functions which cannot be shown in textbooks”.

This number increased to 55 in the fourth edition, most of which again belonged to topographic databases or web-based and electronic atlases. Also, since the volume of the books has not changed significantly, the relative occurrence of hyperlinks generally agrees with the absolute count; They both increased about 2.5 times between the third and the fourth edition. Thus, the evolution of web resources in *Visualization of (Geo)Spatial Data* can be seen from two aspects; Temporally, the introduction of hyperlinks happened since the third edition, and thematically, most of them refer to geodatabases, web maps, and electronic atlases.

Table 26. Absolute and relative occurrence of hyperlinks in *Cartography: Visualization of (Geo)Spatial Data*.

	1st ed. (1996)	2nd ed. (2003)	3rd ed. (2010)	4th ed. (2020)
Occurrence of Hyperlinks	0	0	20	55
Total Word Count	73,902	84,378	88,053	90,146
Relative Occurrence of Hyperlinks	0	0	22.7	61.0

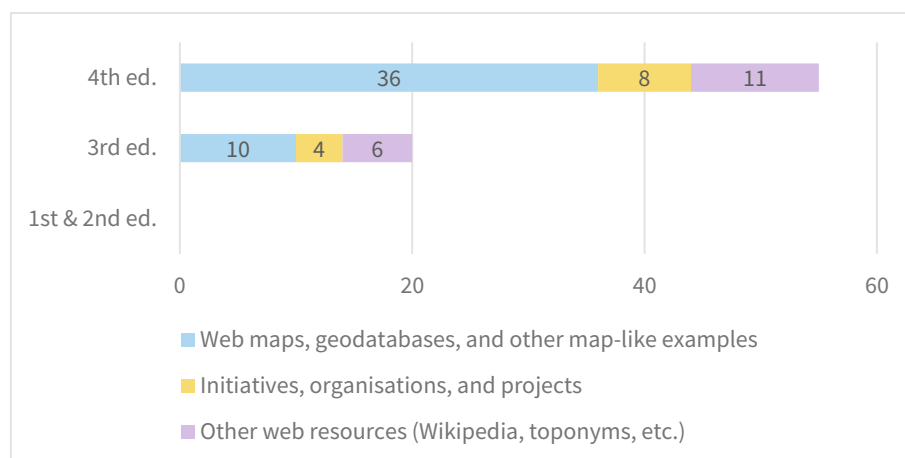


Figure 47. Classification of hyperlinks in *Cartography: Visualization of (Geo)Spatial Data*.

In both editions, the majority of the web resources belong to the first category: web maps, geodatabases, and other map-like examples. The steep increase in this category is mainly attributed to the figure captions rather than the main body text. The number of figures—a lot of which are screen captures of web maps and geodatabases—increased significantly between the third and the fourth edition, resulting in more captions that contain hyperlinks; For example, the number of figures alone in the chapter <<Topography>> increased from 36 to 51.

The other two categories experienced a milder increase. First is the hyperlinks of initiatives, organisations, and projects that are not directly related to web maps or geodatabases. Websites of INSPIRE and OGC belong to this category, despite their close relationship to GDI, because the websites themselves do not disseminate geodata nor do they are focused on providing map examples. Second is the hyperlinks of miscellaneous materials that do not precisely belong to the former two categories such as Wikipedia and national toponym websites. For both categories, the fourth edition added more websites such as Creative Commons, ICA, and a website where toponyms for Australia are listed.

However, it should be noted that this growth is largely due to the increased number of hyperlinks inside the captions rather than inside the text. And in Kraak & Ormeling's books, it was only the fourth edition where hyperlinks are provided inside the captions. In the third edition, none of the figures contained a hyperlink referring to the original source, and they only mentioned the name of the organisation or website. This indicates that there could be a lot of figures that are missing a hyperlink in the third edition, resulting in the less total occurrences. Furthermore, even within an edition, the authors selectively provided hyperlinks that are less likely to change (e.g., national geodatabases) as mentioned in the interview. This results in the question; 'How stable should a website be in order to be included as a hyperlink?' (Figure 48), which makes it difficult to make a comparison among different editions.



FIGURE 6.45 ESRI's Living Atlas, Map of Europe in National Geographic's map style

Figure 48. Example of a figure where the hyperlink to the original source is not provided, despite of its connection to web resources (Kraak & Ormeling, 2020).

As a result, it is true that the number of web resources increased in *Visualization of (Geo)Spatial Data* along with editions. It is especially noteworthy that the first and the second edition did not contain any hyperlinks, and the relative occurrence of the fourth edition rose to 61.0, which is comparable to the third edition of Slocum et al.'s book that recorded 69.9. However, the direct

comparison between the third and the fourth edition is difficult due to the inconsistent inclusion of hyperlinks and the possibility of omitted hyperlinks inside the captions of the third edition.

4.2.3 Derived Information from Text Mining

Here, results from two text mining methods are shown: word frequency and relationships among words using bigram. First, word frequencies are calculated to form a list of top-15 words and top-15 digital keywords. Top-15 words provide a general overview of what a book is about, and the top-15 digital keywords focus on topics that are more relevant to the overarching theme of digital cartography. Second, bigrams of the digital keywords are computed and visualised to investigate the relationships among these keywords and their neighbours.

4.2.3.1 Top-15 Words and Top-15 Digital Keywords

After removing stop words and conjugations, word frequencies are calculated per edition and ordered by occurrence (Table 27). Words that belong to digital keywords—extracted using the content analysis results—are highlighted in yellow.

In all editions, *map*, *data*, *information*, *scale*, and (*geo*)*spatial* ranked high on the list. The shift of the term from *spatial* to *geospatial* took place in the second edition and was maintained afterwards. The occurrence of *GIS* gradually decreased since the first edition, which agrees with the observation from the content analysis that the relevance of the web has constantly increased, replacing some descriptions about GIS applications on computer-based systems. Despite its increased relevance, *web* nor *internet* did not make it to the top-15 words in any of the editions. The occurrence of *user* and *atlas* leapt in the second edition. *User* is a word used often in the context of ‘graphical user interface’ or ‘map user’ (Figure 51). This suggests that the authors included more descriptions to explain the interactive functionalities of electronic and web maps since the second edition. Similar applies to *atlas* which was promoted to a separate chapter in the second edition where a lot of electronic and web-based maps are introduced.

Unlike words that can easily be explained, the frequency of some words such as *line*, *symbol*, and *object* rather pendulated and did not change consistently. For other words, even when the change is consistent, its increase or decrease can be interpreted ambiguously depending on which aspect it is looked at from. For example, the occurrence of *system* has decreased and was no longer on the top-15 words since the third edition. However, when filtered by digital keywords (Figure 49), its ranking has remained stable since the second edition. Thus, it is difficult to draw a conclusion concerning the status of *system* only relying on the word frequency.

Furthermore, some words possess broad use cases and it is difficult to infer their context without examining them. For example, while the word *time* can be used in describing animation—which would be relevant for this research—it can also appear in other analogue mapping techniques that deal with temporal variables or even be used as a simple noun outside the context of maps. This necessitates collecting more information aside from independent words, which is shown in 4.2.3.2 using bigrams that consider preceding and following words.

Table 27. Top-15 words of *Cartography: Visualization of (Geo)Spatial Data*.

1st ed. (1996)		2nd ed. (2003)		3rd ed. (2010)		4th ed. (2020)	
<i>Word</i>	<i>Occ.</i>	<i>Word</i>	<i>Occ.</i>	<i>Word</i>	<i>Occ.</i>	<i>Word</i>	<i>Occ.</i>
Map	1267	Map	1502	Map	1561	Map	1615
Data	824	Data	992	Data	1035	Data	1042
Information	266	Information	332	Information	350	Information	350
Spatial	223	Geospatial	221	Scale	220	Scale	212
GIS	197	Scale	213	Geospatial	210	Geospatial	208
Scale	196	Line	191	Atlas	195	Atlas	177
Line	183	Atlas	188	Line	177	Specific	174
System	162	GIS	180	GIS	175	Set	170
Colour	149	User	170	Set	171	Line	170
Symbol	145	Set	169	Specific	170	User	168
Object	144	Specific	159	User	169	GIS	166
Value	140	Symbol	159	Object	161	Object	166
Digital	137	Colour	158	Symbol	161	Symbol	158
User	137	System	156	Colour	157	Colour	157
Difference	134	Object	145	Time	152	Time	156

When the scope is narrowed to digital keywords, *GIS* ranked at the top until the third edition. However, this is reversed in the fourth edition by *user* which constantly rose since the beginning. Another obvious change is the advent of *web* in the second edition. *Web* was mentioned only six times in the entire first edition, but its occurrence skyrocketed to above 60 in every edition since the second. *Web* took over the place of *computer*; It ranked 14th in the first edition, however, it never made it back to the top-15 afterwards. The occurrence of *computer* still recorded at around 45 in the later editions, which is not a huge decrease in terms of absolute word counts, but it can be said that its relative importance decreased considering the ranking.

From the first to the fourth edition, the occurrence of *interactive* grew from 39 to 82. This can be explained in a similar way with the word *user*; The authors introduced more web-based maps and geodatabases that are equipped with interactive functionalities, resulting in more mentions of *interactive* and *user*. On the other hand, the occurrence of *digitize* plunged from 68 to 44, which aligns with the observation that a lot of descriptions about digitising technologies and hardware were continuously reduced. Interestingly, the mention of the word *digital* itself also gradually died down; Whereas *digital* used to be the third-most-frequently-used digital keyword in the first edition with 137 occurrences, it dropped to 96 in the fourth edition.

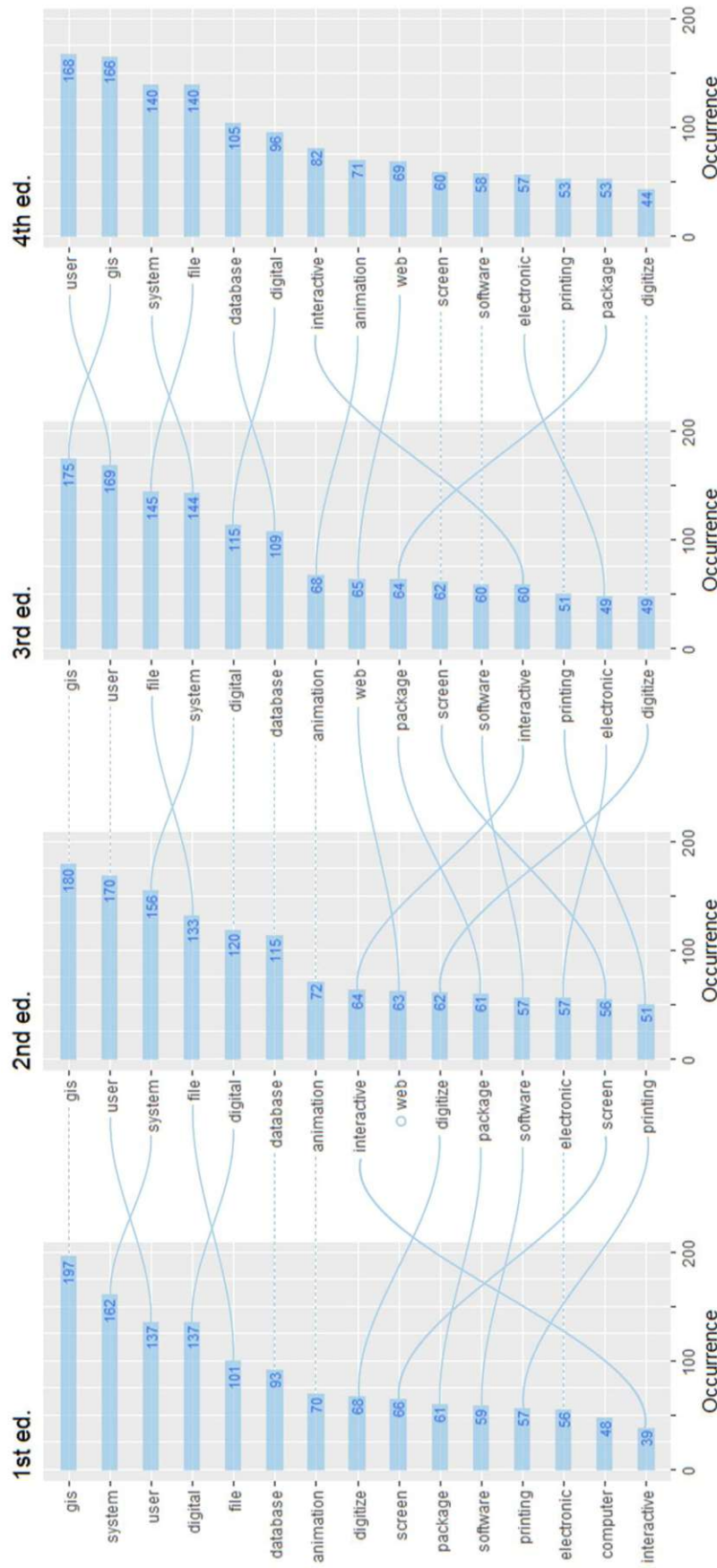


Figure 49. Flow diagram of bar charts showing changes in the rank of digital keywords from *Cartography: Visualization of (Geo)Spatial Data*. Keywords are connected to the next edition with a flow line. Dotted lines are drawn for keywords whose rank did not change between the editions. Keywords that are newly included in the top-15 are marked with a donut on the left side of the label. Keywords without any lines indicate that they were not included in the top-15 in the following edition.

It should be noted that the lower the rank is, the smaller the margin among occurrences tends to be. This means that even a few deletions or additions can affect the ranking, which is why the rankings in the lower half fluctuate whereas the upper half is relatively stable. For example, the trajectory of the ranking of the word *screen* appears as if its occurrence has changed drastically, however, the actual counts are 60, 56, 62, and 60, which is approximately 10% variance at maximum. This is a typical pitfall of working with a small word corpus such as Kraak & Ormeling's books, which produces word frequencies with a lot of similar values. Contrary to some cases where the evolution of relevance of a topic can easily be inferred by combining the knowledge from the content analysis (e.g., *digitize* and *web*), in many other cases, it is not clear and it is difficult to draw conclusions only based on the word frequency. Therefore, taking additional context into account can be beneficial to better understand these subtle changes that are not pronounced enough by simple word counts.

To summarise, the word count in Kraak & Ormeling's book well shows the pattern of changes in coarse topics of digital cartography such as printing, GIS, and the web. This is represented by the huge leap of *interactive* and *web*, contrasted with the fall of *digitize*. However, because of the small text corpus in all editions, a lot of word frequencies have similar values and the comparison based on their rank is prone to misinterpretation. This is demonstrated by the example of *screen*, where the trajectory inferred from the visualisation in *Figure 49* leads to a wrong conclusion. Furthermore, with some common words with various use cases such as *digital*, it is difficult to pinpoint the reason for their increase or decrease. As a result, making decisions based on simple word frequency is imprecise, suggesting that additional information should be considered to better understand the context of words.

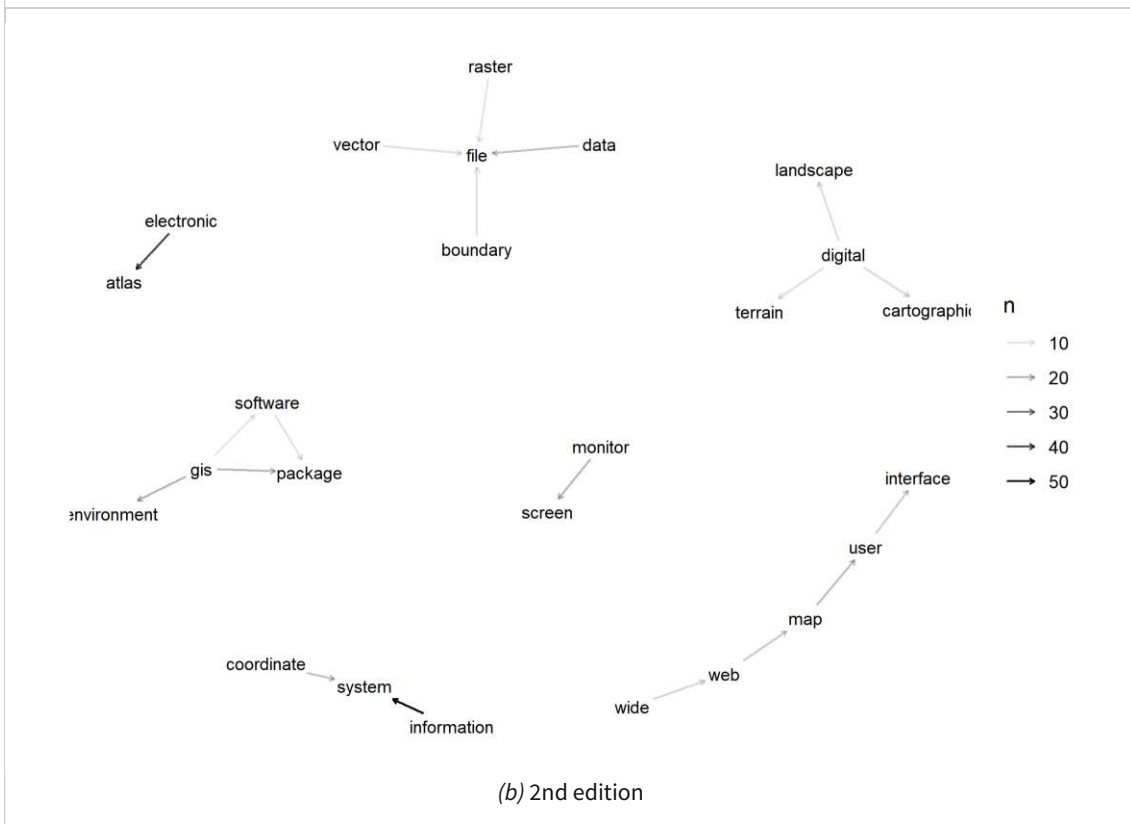
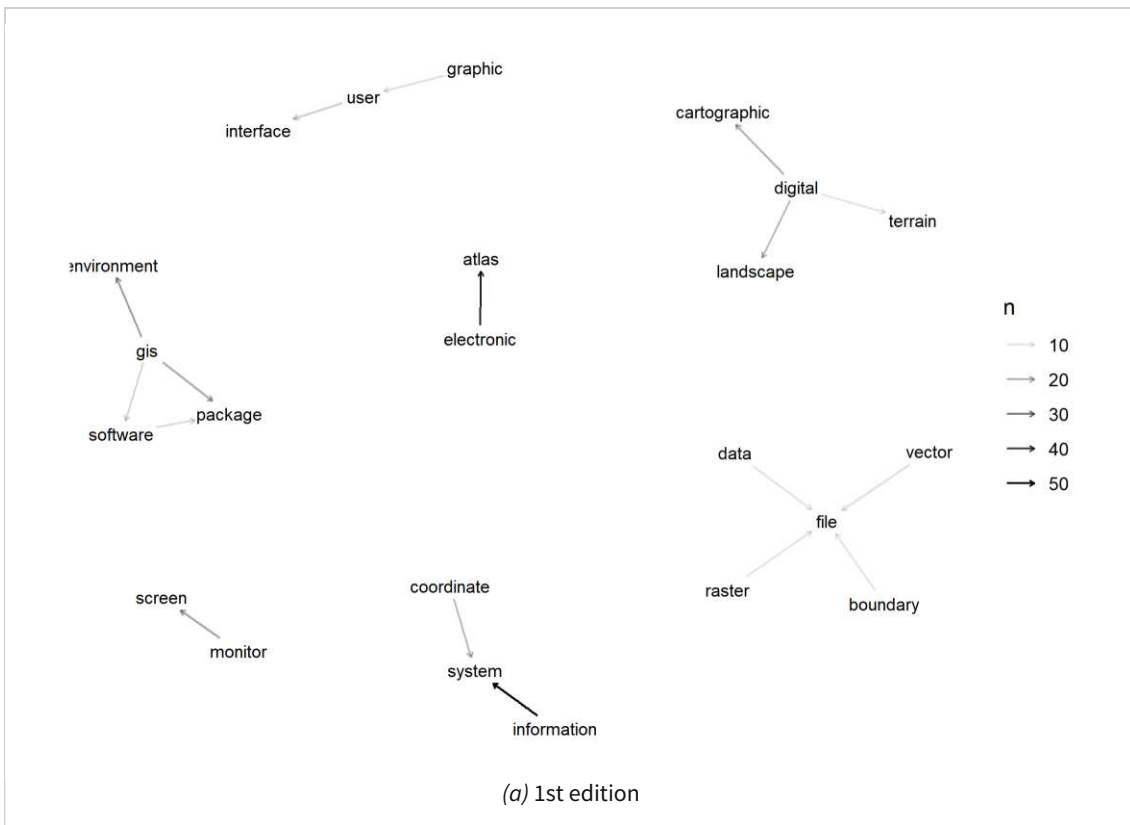
4.2.3.2 Relationships among Digital Keywords and Their Neighbours

Bigrams are used in order to consider the context of words, looking at the pair of words that co-occur within an edition. After calculating the occurrence of bigrams, they are filtered by digital keywords. Next, they are visualised into a directed network graph showing the frequency of bigrams that occurred at least 10 times in the edition (*Figure 50*). Finally, the graphs are aggregated per book so that the changes along the editions are better pronounced.

The strongest links throughout all four editions are (*geographic*) *information system* and *electronic atlas*. This is followed by less frequent but still consistently used bigrams such as *monitor screen*, *coordinate system*, *GIS package*, and *GIS environment*. Similarly, *file* has been preceded by *data*, *raster*, *vector*, and *boundary*, except for the fact that *boundary file* was less used in the fourth edition because of the OSM replacing some digital boundary geodatabases.

Since *graphic package* is a term that appeared since the third edition, there are no mentions of it in the first and the second edition—they used to be referred to as desktop publishing software (DTP). Also, its node in the fourth edition could not be visualised because parts describing the functionalities of a graphical package were deleted, resulting in less than 10 occurrences. Another use case of *graphic* is *graphic user interface (GUI)* which has occurred around 10 times from the first to the third edition, and six times in the fourth edition.

4 Results



4 Results

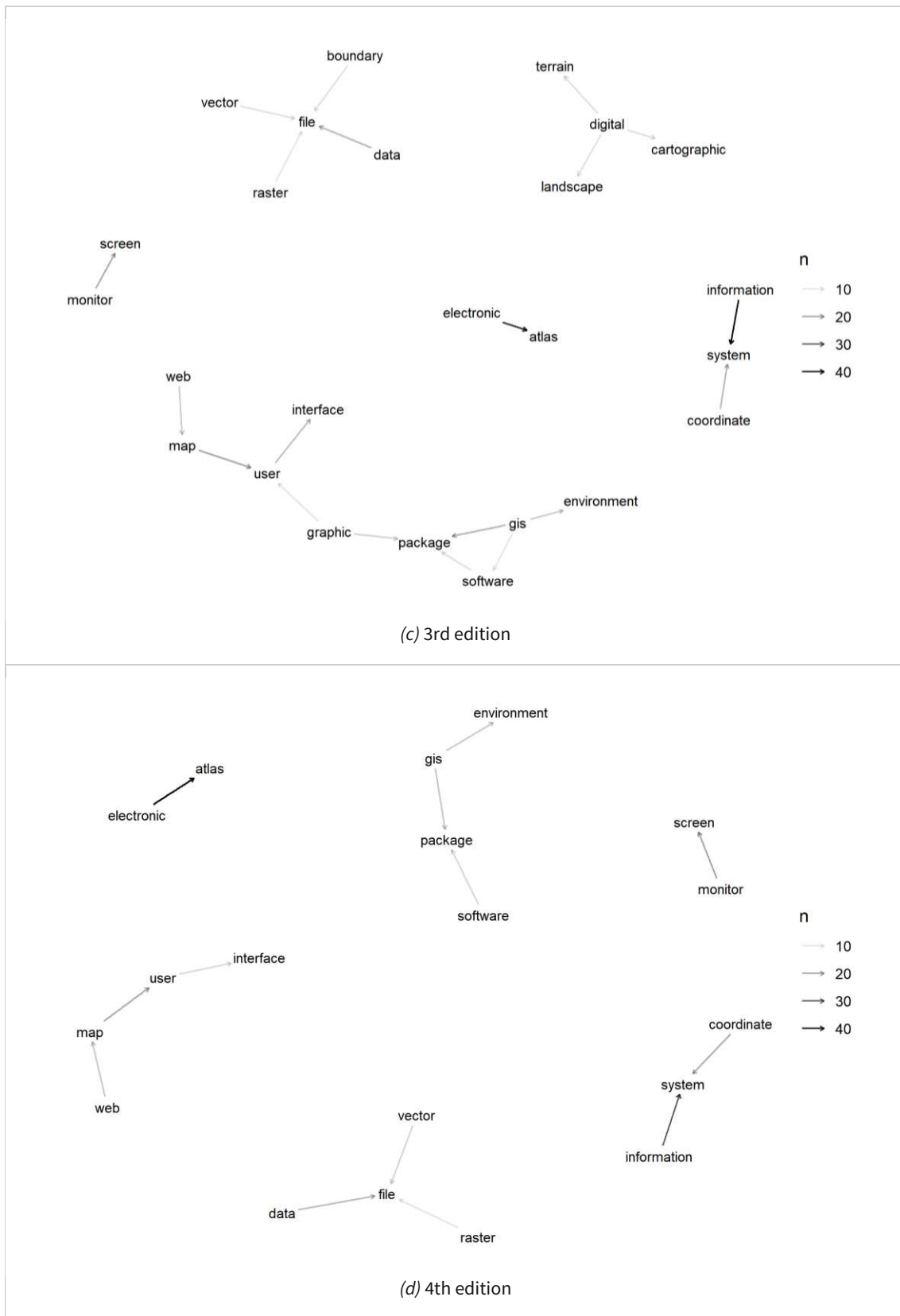


Figure 50. Directed network graph of bigrams of each edition of *Cartography: Visualization of (Geo)Spatial Data*.

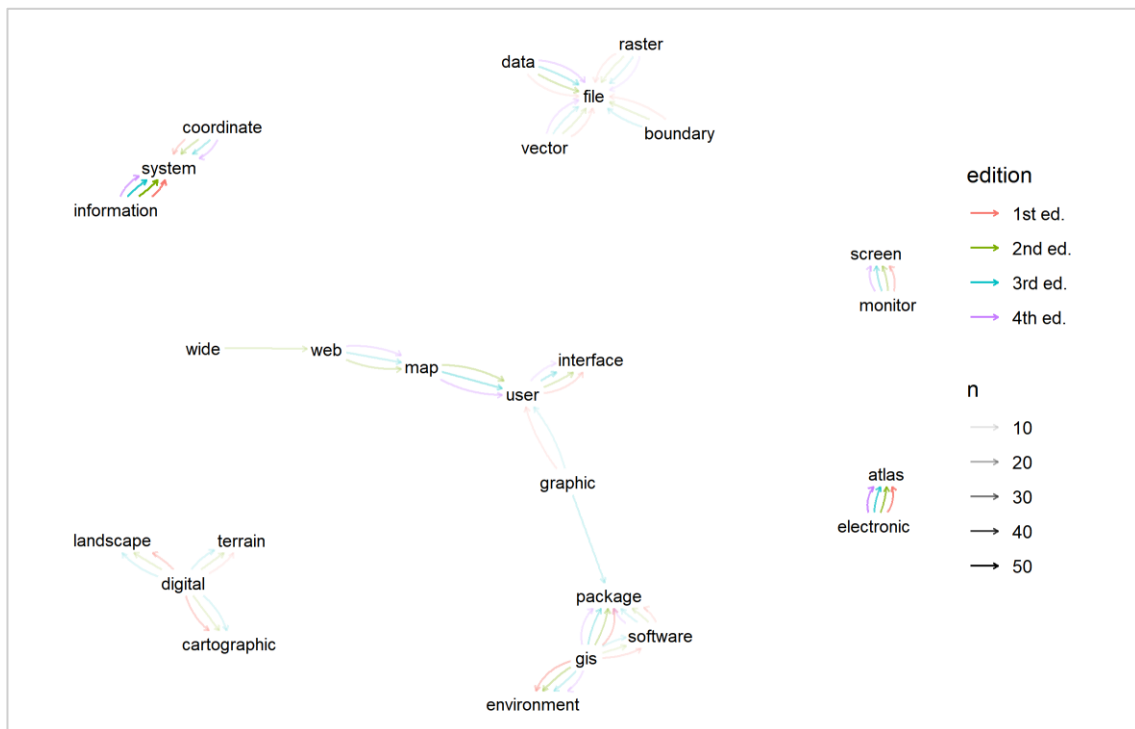


Figure 51. Directed network graph of bigrams of *Cartography: Visualization of (Geo)Spatial Data*.

Also, using bigram analysis, it can be partially explained how the word *digital* has diminished over time. Its popular use cases were *digital cartographic model (DCM)*, *digital terrain model (DTM)*, and *digital landscape model (DLM)* in the first edition, usually mentioned in the context of national geodatabases such as ATKIS in Germany and USGS in the US. However, they were less frequently mentioned in the following editions since some national geodatabases were replaced by international projects, and eventually occurred less than 10 times in the fourth edition.

Web map rose as a common bigram in the second edition, contrary to the first edition where *web* only occurred six times and the bigram *web map* did not exist. Also, whereas *world wide web* was frequently used in the second edition, most of them were replaced with *the Web* since the third edition, resulting in no vertex between *wide* and *web* after the second edition. Similarly with *web map*, the occurrence of *map user* doubled between the first and the second edition—from 9 to 18. This agrees with the previous observation that the increased relevance of web maps has entailed increased descriptions of their interactive functionalities, which often contain the expression *map user*.

To summarise, the bigram analysis provides another viewpoint from the word frequency. Whereas the largest pillars of changes such as *web* and *user* are also visible in the bigram analysis, other changes are often concealed because of the even sparser number of occurrences. Instead, some unknown relationships are revealed; For example, the change of the term from ‘World Wide Web’ to ‘web’ is found. Also, whereas it was nearly impossible to infer the reason for the decrease in the word *digital*, the bigram analysis provides possible hypotheses by looking at its most common use cases.

5 Synthesis and Discussion

In this section, the results of the analyses will be summarised together to answer the research questions. It will also include limitations of the methodologies that were revealed during the research and provide further suggestions to overcome the shortcomings.

RQ 1a Which codes related to digital technologies can be derived from cartography textbooks?

The documented changes from Slocum et al.'s textbook were tagged using a total of 25 codes. For Kraak & Ormeling's textbook, a total of 22 codes were found. They had 17 codes in common which represent topics that both books share (Figure 52). In contrast, the unique codes show how each textbook's scope differs from the other; For example, Slocum et al.'s book had a large emphasis on thematic mapping and quantitative methods, which resulted in codes such as colour, ESDA, and statistics. Also, topics such as 3D and virtual environments were important which were not thoroughly covered in the other book. On the other hand, Kraak & Ormeling's book contained unique contents concerning copyright, semantics, GDI, and toponyms.

While codes were useful in classifying the content for further abstraction, they did not have quantitative values. For instance, the frequency of codes could not be counted because documented changes highly varied in size; A documented change could be a single sentence or the entire subchapter. Although this was not adopted in this project, a possible solution is to slice every change to a sentence level and code each sentence, which can help understand the thematic distribution of changes and gain additional insights.

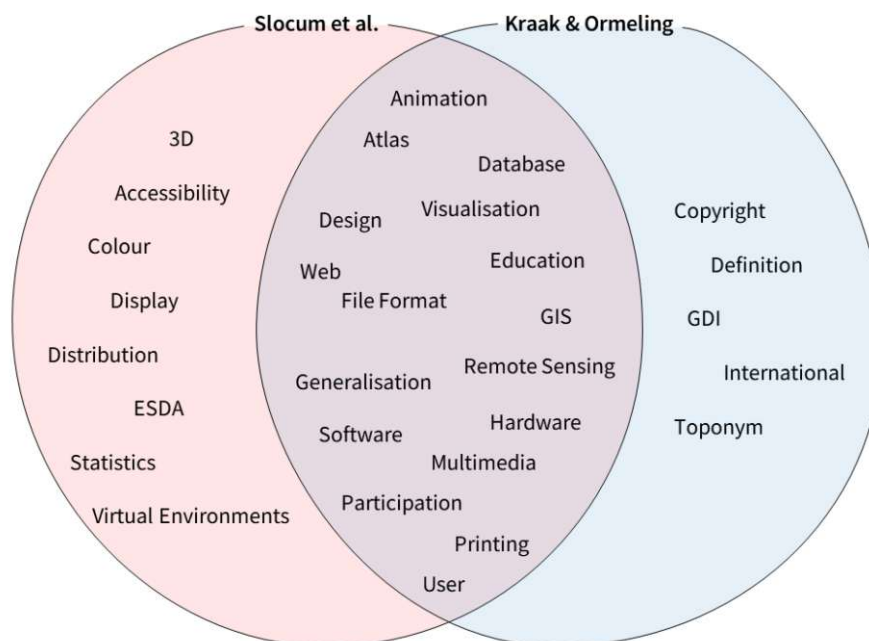


Figure 52. Venn diagram of codes extracted from two textbooks.

RQ 1b What kind of patterns and trends are observed concerning digital technologies?

In Slocum et al.'s books, a total of eight categories were formed by assorting multiple codes, which help understand the patterns and trends from the documented changes. In terms of *map production methods and colour*, their evolution was explained in accordance with the life cycle of digital technologies involved in map production. For example, map dissemination using television and printers dwindled as computer-based and web-based methods emerged. Also, the democratisation of colour display and printing soon became normal and was no longer considered an innovation; Instead, web technology replaced this anticipation.

Regarding *GIS and map design*, the authors first perceived it as a crisis that the border of traditional cartography is dissolving due to the rapid development of GIS. However, since the third edition, they started to affirm the new possibilities of integration witnessing the rise of geovisualisation and related fields. Also, enhanced functionalities of GIS software did not result in absolute dominance as a mapping tool, but its development trickled down to the growth of other software that can assist in map design.

The automation of *generalisation* used to be a difficult task due to the limited capability of the algorithms and the complexity of the databases but was largely resolved at the time of the third edition. Also, the meaning of *multimedia* has been broadened along with the editions; from sound and visuals to other senses, including touch and smell.

In *topographic mapping*, the evolution was seen in the second edition regarding how to display 3D landscapes and generate a 3D physical model. In *thematic mapping*, improvements were observed in various mapping methods such as dasymetric mapping, cartograms, and flow maps. This is supported by software applications that can handle complex statistics (e.g., SpaceStat, ChoroWare), and a novel geodatabase LandScan which shows how technologies associated with cartography (i.e., remote sensing, the web) can refine thematic mapping methods.

Since the second edition, *virtual environments* have been anticipated as a technique that can create realistic representations of the environment. Its evolution could be summarised into two streams: technological development in software and hardware, and the increased accessibility represented by the emergence of Google Earth.

Animation was initially spotlighted in cartography as a tool to assist in *data exploration*. However, the perspective of data exploration was diversified since the second edition with the emergence of the ESDA concept, and other functionalities other than animation became integrated into the data exploration. Also, while they have long existed as a highly anticipated visualisation technique, user studies were followed relatively later to evaluate the effectiveness and cognitive limitations of animation. Also, animation and ESDA influenced new neighbouring fields of cartography such as information visualisation, spatial data mining, and visual analytics.

Electronic atlases and web maps show the transitional period where a lot of electronic atlases in CD-ROMs were transferred to the internet, and then eventually replaced by web maps. Along with editions, descriptions of their capabilities and limitations have changed. Also, the influence of the web has propagated further and resulted in novel ways of producing and consuming geo-data such as location-based services (LBS) and mobile mapping.

In Kraak & Ormeling's books, a total of six categories were formed by assorting multiple codes. The central concepts in understanding the evolution of digital cartography in their books are *GDI and geodatabases*, since they represent how online sources of geodata have been diversified and democratised, supported by initiatives and policies that foster the standardisation and the exchange of geodata. Changes observed in the USGS and the Ordnance Survey show how the national organisations have adapted to the changing environment and needs. Also, descriptions in each edition about international projects such as SABE and Global Map provide a snapshot of how global databases are designed and implemented.

The development of *electronic atlases* provides historical insights through the intersection of map products, computers, and the web. Whereas the earlier editions focused on functionalities that computer-based systems can offer, the development of geodatabases has opened new possibilities for accessing and manipulating the original data using atlases. The potential of the web is also largely reflected in *animation and multimedia*. Each edition has shown novel data exploration techniques of the time and various plug-ins that enable interactive functionalities.

Regarding *map production*, it can be observed how GIS and graphic software have been increasingly integrated and compatible with each other. The increased importance of digital dissemination resulted in a lot of printing hardware being replaced. In addition, as a web browser was introduced as another method of producing digital maps (e.g., WMS, WFS, WCS), the boundary between web mapping and computer-based mapping became blurred.

Concerning *remote sensing*, it was demonstrated how far remote sensing technologies have evolved and how relevant they have become in cartography. This was largely enabled by web-based tools that allow easier access and use of location-based data and satellite imagery.

Furthermore, updates in *semantics and policies* are observed, revealing how definitions and neighbours of cartography have changed in response to the increased relevance of digital technologies. For example, the authors have introduced a total of four definitions of cartography from different sources, all of which emerged during the rise and growth of the digital environment. Also, along with the editions, they added new concepts and subfields that are relevant to cartography such as copyright, usability, visual analytics, and VGI.

The most pronounced trend found in both books is the increased relevance of computer- and web-based mapping methods. This has transformed the way how cartographers and users access the data, produce maps, and disseminate them. So far, GIS and graphic software have risen as the predominant mapping tool, which caused the demise of printed maps. The development of computer-based map production has led to the growth and diversification of functionalities in digital mapping techniques such as animation, multimedia, data exploration, and virtual environments. Here, the emergence of the web has brought new possibilities for handling geodata; Numerous geodatabases and international initiatives were established which fostered the standardisation and the exchange of geodata. Its development is also closely related to neighbouring fields of cartography including visual analytics and location-based services. Furthermore, the web is increasingly used as a mapping tool—not only as a vessel for data exchange—and is again substantially altering the pipeline of map production.

RQ 2a *How have the usage and occurrences of web resources in cartography textbooks changed over time?*

Web resources were included in the content analysis since they are a direct proxy of how digital technologies are integrated into the textbook. In both books, their occurrence has constantly increased along with the edition, which reflects the phenomenon where more maps and databases are published online, resulting in the need to introduce them using hyperlinks.

In Slocum et al.'s books, this increased integration of web resources was achieved in two ways; First, they introduced new digital maps and software packages, and second, they complemented the existing content with web resources. Furthermore, it was found that the authors started to include auxiliary sources such as Wikipedia and project websites, which shows the changes in textbook writing concerning how it has become more common and convenient to outsource the 'further reading' by referencing external web resources.

In Kraak & Ormeling's books, the stark contrast was observed between the second and the third edition; They did not include any hyperlinks in the main body text until the second edition, however, the number increased to 20 in the third edition and eventually to 55 in the fourth. While it is true that the overall integration of web resources has been enhanced along with the editions, some inconsistencies were found concerning the inclusion of hyperlinks inside the caption. And since it is the authors' decision about which websites to introduce and which to omit, comparing occurrences of websites in different books could lead to a wrong conclusion.

The largest hurdle in the analysis of web resources was that they could not be qualitatively compared. This is because websites' links are often altered, or their content is relocated to other repositories without notice. For example, approximately half of the websites introduced in the fourth edition of Kraak & Ormeling's book—which was published in 2020—were already obsolete and could not be accessed. The simplest solution would be the authors consistently keeping track of their examples and updating the hyperlinks, which is not plausible because it is highly time-consuming. Moreover, even when the link is accessible, it is likely to display only the latest version of the website, making it nearly impossible to observe the changes in the content over time. As a result, a more sophisticated and robust workflow should be designed if the actual content of the web resources were to be accessed and compared.

RQ 2b *How have the occurrences and neighbours of digital keywords changed over time?*

Using the result from the content analysis, the top-15 keywords related to digital technologies were extracted from each edition of a textbook (3.4.1). When aggregated per book, this resulted in a total of 21 digital keywords in Slocum et al.'s book, and 16 in Kraak & Ormeling's book. Here, it can be inferred that the content of Slocum et al.'s book went through more extensive changes compared to Kraak & Ormeling's, because the keywords per edition have a larger variance. Among these, 10 of the keywords overlapped in both textbooks: *animation, computer, electronic, user, software, GIS, digital, web, printing, and interactive*. On the other hand, some keywords were unique to each book, showing the different scope of the book similarly to the codes.

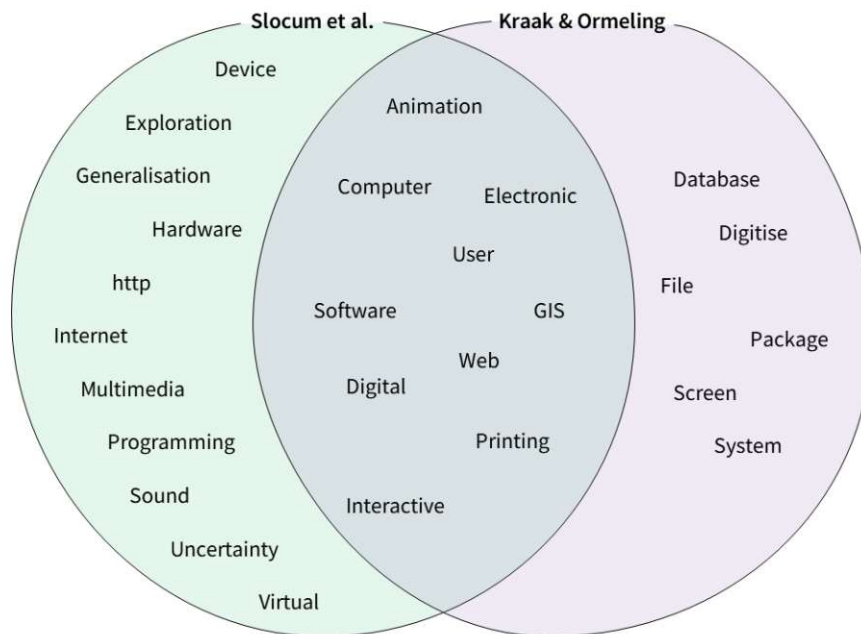


Figure 53. Venn diagram of digital keywords from two textbooks.

In Slocum et al.'s book, the changes in the occurrence of digital keywords coincided with the major restructuring between the two consequent editions. Occurrences of *generalization*, *virtual*, *device*, and *web* in the second edition, and *uncertainty* and *http* in the third edition increased, reflecting their heightened relevance. *Web* and *user* showed a disproportionately large increase both in the second and the third edition, which means their importance has become phenomenally larger. Also, the result of the bigram analysis concurs with the findings from content analysis and the word count; The relevance of the web and its interactive functionalities have increased. Furthermore, by looking at the combination of words, new terms such as *wearable computer* and the changes in terms such as *virtual reality* versus *virtual environment* were identified.

In Kraak & Ormeling's book, the occurrence of digital keywords also agrees with the general pattern of changes in larger topics of digital cartography such as printing, GIS, and the web. This is represented by the huge leap of *interactive* and *web*, contrasted with the fall of *digitize*. Also, some unknown relationships are revealed through the bigram analysis; For example, the change of the term from 'World Wide Web' to 'web' was found. Also, whereas it was nearly impossible to infer the reason for the decrease in the word *digital*, the bigram analysis provided possible hypotheses by looking at its most common use cases.

Although text mining provided another perspective to view the evolution of digital cartography, two limitations were identified. First, it produces unreliable results when the text corpus is small, which is particularly applicable in the case of Kraak & Ormeling's book. This results in a lot of word frequencies or bigrams having similar values, indicating that the comparison based on their rank is prone to misinterpretation. Also, the first edition of Slocum et al.'s book was significantly smaller in volume compared to the latter two, making it difficult to compare the

absolute occurrence of bigrams among editions. Second, the text mining methods adopted in this research could not accurately distinguish the context of words. As a result, some words are accurately counted because of their homogenous use cases, while some are counted less than their use since they are referred to in various forms and are used in different contexts. Although the researcher tried to remove as many conjugations as possible if they are likely to be used in the same context (3.4.1), a lot of words could not be stemmed (e.g., explore). This can be resolved by working with textbooks that have a larger volume and by using more complex and robust text mining methods based on machine learning and deep learning.

RQ 3a *How useful is content analysis in comparing and understanding changes in the textual content of cartography textbooks?*

Content analysis was selected as the main methodology to analyse the changes throughout multiple editions of cartography textbooks. Its strengths are that it can treat textual data without editing the content and that it involves a systemic abstraction. These two strengths were adapted in this research so that; first, all relevant changes are seamlessly documented, and second, the documented changes are coded and categorised. The first aspect helped the researcher observe subtle changes in descriptions that could have been lost when only seen from a higher perspective, and the second approach was useful in extracting patterns and trends in digital cartography by breaking down the overarching theme into more specific topics that have gone through frequent updates along the editions of a textbook. Furthermore, although it was largely limited to the rudimentary quantitative analysis, content analysis was also proven to be useful in documenting the changes in semi-textual content such as hyperlinks.

In contrast to the documentation and comparison, the content analysis showed several limitations in deriving further information from the changes. These limitations can be reduced to the fact that a lot of updates in a textbook remain unexplained, making it difficult to infer the reasoning behind them. First, the content analysis only allows one-to-one comparison of two editions of a book. This means that the extracted trend from the first to the last edition of a book is the result of the interpolation of a researcher, which could entail the loss of information. For example, if a description about a certain technology appeared in the first edition, was deleted in the second, reappeared in the third, and was again deleted in the fourth edition, the researcher is likely to draw a conclusion that its relevance has eventually decreased, which holds only when connecting the first and the last edition.

Second, it is difficult to account for the genesis of the content in the first edition. Although the emphasis of this research was on the changes among editions, it could be as equally important as why some contents were included in the first edition. However, since the first edition does not have a precursor to be compared to, it was rather considered a static reference. Therefore, some contents from the first edition of textbooks were not thoroughly described in the results section if their content has barely changed throughout the editions. This again inevitably results in the loss of information, possibly omitting some content that could have offered meaningful insights into the theme of digital cartography.

RQ 3b What other methods can be applied to complement the results of content analysis?

First, the authors of the textbooks were interviewed to comprehend the reasoning behind the changes. It was expected that these unexplained changes as well as the general book writing process can be better contextualised through interviews. Most of their answers supported the inference of the content analysis; Deletions were largely due to the decreased relevance of the topic and new insertions and vice versa. The interviews were also helpful in learning new background information such as the Dutch version *Kartografie* which functioned as the basis of the first edition of *Visualization of (Geo)Spatial Data*.

Second, text mining was adopted. Since the content analysis of the changes was mainly qualitative, a more quantitative approach was expected to support and extend the findings. The results showed that the word occurrence and bigram analysis generally agree well with the conclusion derived from the content analysis while allowing the researcher to view the changes from another aspect. However, its contribution to revealing unknown facts was rather marginal. This can be attributed to the limited ability of bigrams to figure out the context of words since it only considers a pair of words. Although it was not feasible in this research because of the small text corpus, a possible solution is to increase the N of n-grams (e.g., trigram) to look at more sequencing words, thereby approximating the specific use case of words.

RQ 3c Can content analysis and other methods be further used to compare different books?

This research focused on documenting and analysing the changes within a textbook without comparing one to the other. The methodologies applied in this research (i.e., content analysis, author interviews, text mining) are theoretically apt for comparing different books and are often utilised for that purpose, however, there are several points to be considered.

First, each book has an intended audience. For example, while it was assumed that two textbooks would share a similar purpose since they are both used in universities, it was revealed during the interview that they were quite different. Kraak said that their book is intended for students with advanced knowledge such as “end of bachelor’s or the beginning of master’s”. On the other hand, Slocum wrote in the first edition’s preface that “I have written it so that it can be understood by those who have had no prior training in cartography” (Slocum, 1999). This explains why Slocum et al.’s book is more extensive and covers a wider range of topics. Thus, findings such as ‘Kraak & Ormelings’ book did not cover basic statistics’ are not significant because the authors intended to do so, and these differences do not represent the relevance of a topic.

Second, publishing is a complex issue influenced by various actors and motivations. For example, regarding the full-colour printed copy of the fourth edition, Kraak mentioned the competition in the cartography textbook market as one of the reasons. Similarly, Slocum said that “the publisher encouraged them to update the book because of the commercial success of the previous edition”. This suggests that the content analysis can fall short in comprehending the update of a book and that trying to derive a conclusion by the internal comparison of different books would be inaccurate because the influence of external factors is crucial.

Third, writing is a personal matter, and therefore, the usage of words is highly subjective. This makes it difficult to compare the results from the text mining, especially when the text corpus is not large enough. While the most common and established terms (e.g., map, data) can be relatively reliable, other words are more variable. For example, an author may prefer the noun ‘display’ over ‘screen’—or vice versa—without contemplating the precise definition of those two. Therefore, even when the context of a word is figured out, the issue remains that the quantitative summary of words cannot readily be used as a tool for comparing different books.

The comparison of different books is feasible when the scope of the books aligns well with each other, and the topic of interest is relatively narrow so it is less needed to deal with the authors’ intention. A good example is a work by Kessler (2018) where they compared various cartography textbooks in terms of map projection. On the other hand, it was rather inappropriate for this research which had a much broader theme, and the focus was on the changes in the content of which a large portion relied on the authors’ decision. Thus, while it is possible to extend the methodologies to further compare different books, it was not adopted in this research due to several reasons that could produce inaccurate results.

6 Conclusion and Outlook

This project sought to trace the evolution of digital cartography and map the history of digital technologies used in cartography. To achieve this goal, changes regarding digital cartography found in multiple editions of cartography textbooks were analysed. Through the content analysis of three editions of *Thematic Cartography and (Geo)Visualization* and four editions of *Cartography: Visualization of (Geo)Spatial Data*, the changes were documented and exhibited using codes and categories. The results suggest that the evolution of digital cartography can be observed in various realms, from printing and map design to virtual environments and digital policies. While the selected textbooks had different scopes, commonly identified is the emergence and proliferation of computer-based GIS and the web, which have led to a substantial shift in cartography in terms of how geodata and maps are produced and disseminated. Their influence is also accountable for the recent development of neighbouring fields including visual analytics, location-based services (LBS), volunteered geographical information (VGI), et cetera.

The increased relevance of digital technologies was also confirmed by the analysis of web resources included in textbooks and the results from text mining. In both textbooks, the number of hyperlinks has constantly increased, which indicates that more map examples and geodatabases are being published online and the authors of the textbooks have responded to this phenomenon. The results from text mining supported this as well; Word occurrences found that terms related to web mapping (e.g., user, web, interactive) are significantly used more often with the progress of editions. The bigram analysis showed that some evolution of the textual content in a textbook can be better understood by considering the context of words. Also, some limitations were identified during the quantitative analysis; These are mainly attributed to the relatively small text corpus of the selected textbooks, which results in similar statistic values which are subject to change with the slightest alterations.

The application of content analysis in mapping changes in a certain topic was proven to be useful in formulating various changes into multiple trends while preserving individual changes. Here, author interviews provided background information to updates that could not solely be explained by content analysis and have helped grasp the complex issue of writing and publishing a textbook. Furthermore, it was demonstrated how the findings from the qualitative analysis can be supported and extended by coupling more quantitative approaches such as counting occurrences of hyperlinks and text mining. However, extending the methodologies for cross-book comparison requires precaution due to the complex issue of writing a book. It should be noted that each book is written in consideration of a specific audience and that numerous external factors can intervene in the writing process. Therefore, in order to extend the methodologies for comparing different books, it is recommended that the researcher ensures that the scope and the purpose of selected textbooks are highly similar and that each text corpus is large enough to be compared with other books.

The key challenge of implementing different approaches was the disparity in the size of datasets that each research method requires. Whereas documenting and analysing the changes in

textbooks consisting of hundreds of pages was a reasonable workload from a qualitative research perspective, some editions' text corpus was not large enough for text mining to derive concrete conclusions. Therefore, future research is needed on identifying the optimal intersection of the dataset size between content analysis and text mining. If textbooks with larger volumes were to be chosen, this would consequently require designing detailed criteria about which changes are to be counted and which are to be omitted, so that only broader changes are consistently documented whereas the smaller ones are systemically removed. Other shortcomings mentioned above could also be considered to refine the methodologies; minimising the loss of information caused by content analysis and applying advanced text mining methods that are more robust to various context of words.

Also, it could be explored how to exhibit the changes in textual data. This research largely relied on illustrating the changes in a written form, since it was considered the most non-destructive method of displaying them. Whereas this helped preserve subtle changes that also play a role in comprehending the evolution, it resulted in a time-consuming process of writing and a complex thesis structure. With various interactive web-based libraries available, future research can experiment with how to automatically aggregate the changes, assort them according to the theme, and visualise them in a way in which readers can easily understand the patterns.

Furthermore, due to the emphasis on the textual content, other materials such as diagrams and map examples were not analysed in this research. Considering that the foremost purpose of cartography textbooks is to teach maps, visual materials are as important as text descriptions and their changes can provide meaningful insights. Therefore, further research could set focus on non-textual contents and design a framework for documenting and comparing the changes in them, providing another perspective from which the evolution of the topic could be seen.

Given the importance of digital technologies in cartography and the need to augment its discipline, historical research on digital cartography can contribute to knowing how far the development has come and setting our goals anew. To instantiate this, this project suggested cartography textbooks as a lens through which evolution can be observed. While it was demonstrated how the changes of their textual content can be documented, compared, and analysed to learn the evolution, there are improvements that can be made and realms that are to be explored. With further research that would result in various perspectives, the aspects and directions of the evolution of digital cartography can be more holistically understood.

Bibliography

- American Psychological Association. (2019). *Publication Manual of the American Psychological Association*. American Psychological Association.
- Bell, E., Jansen, H., & Young, L. (2007). Sustaining a Dynasty in Alberta: The 2004 Provincial Election. *Canadian Political Science Review*, 1(2), 27–49.
- Berelson, B. (1971). *Content Analysis in Communication Research*. Hafner.
- Besel, A., Zimmerman, T. S., Fruhauf, C. A., Pepin, J., & Banning, J. H. (2009). Here comes the bride: An ethnographic content analysis of bridal books. *Journal of Feminist Family Therapy*, 21(2), 98–124.
- Bryman, A., & Bell, E. (2011). *Business Research Methods*. OUP Oxford.
- Bryman, A., & Bell, E. (2019). *Social research methods*. OUP Oxford.
- Burrows, R., Nettleton, S., Pleace, N., Loader, B., & Muncer, S. (2000). Virtual community care? Social policy and the emergence of computer mediated social support. *Information, Communication & Society*, 3(1), 95–121.
- Cobb, G. W. (1987). Introductory Textbooks: A Framework for Evaluation. *Journal of the American Statistical Association*, 82(397), 321–339.
<https://doi.org/10.1080/01621459.1987.10478436>
- Cummings, J., & Mittelbach, A. (2010). The Holinshed Project: Comparing and linking two editions of Holinshed's Chronicle. *International Journal of Humanities and Arts Computing*, 4(1–2), 39–53. <https://doi.org/10.3366/ijhac.2011.0006>
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1), 107–115.
- Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing content analysis. *African Journal of Emergency Medicine*, 7(3), 93–99. <https://doi.org/10.1016/j.afjem.2017.08.001>
- Feldman, R., & Sanger, J. (2006). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511546914>
- Foucault, M. (1984). *The Foucault Reader*. Pantheon Books.
- Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11), 1129–1164.
<https://doi.org/10.1002/spe.4380211102>
- Goodchild, M. F. (1991). Geographic information systems. *Progress in Human Geography*, 15(2), 194–200. <https://doi.org/10.1177/030913259101500205>
- Goodchild, M. F. (2000). Cartographic Futures On A Digital Earth. *Cartographic Perspectives*, 36, 3–11. <https://doi.org/10.14714/CP36.821>
- Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, 24(2), 105–112.

- Greer, J., & Mensing, D. (2006). The evolution of online newspapers: A longitudinal content analysis, 1997-2003. *Internet Newspapers: The Making of a Mainstream Medium*, 13–32.
- Harley, J. B., & Woodward, D. (1989). Why Cartography Needs Its History. *The American Cartographer*, 16(1), 5–15. <https://doi.org/10.1559/152304089783875604>
- Harman, D. (1991). How effective is suffixing. *Journal of the American Society for Information Science*, 42, 7–15.
- Hatton, E., & Trautner, M. N. (2011). Equal opportunity objectification? The sexualization of men and women on the cover of Rolling Stone. *Sexuality & Culture*, 15(3), 256–278.
- Herring, S. C., Scheidt, L. A., Kouper, I., & Wright, E. (2012). Longitudinal content analysis of blogs: 2003–2004. In *Bloggling, citizenship, and the future of media* (pp. 19–36). Routledge.
- Hotho, A., Nürnberger, A., & Paass, G. (2005). A Brief Survey of Text Mining. *LDV Forum - GLDV Journal for Computational Linguistics and Language Technology*, 20, 19–62.
- Hsieh, H.-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>
- Jackson, M. J. (1987). Digital Cartography, Image Analysis, and Remote Sensing: Towards an Integrated Approach. *Interdisciplinary Science Reviews*, 12(1), 33–40. <https://doi.org/10.1179/isr.1987.12.1.33>
- Jo, I., & Bednarz, S. (2011). Textbook Questions to Support Spatial Thinking: Differences in Spatiality by Question Location. *Journal of Geography*, 110, 70–80. <https://doi.org/10.1080/00221341.2011.521848>
- Jurafsky, D., Martin, J. H., Norvig, P., & Russell, S. (2014). *Speech and Language Processing*. Pearson Education.
- Kessler, F. (2018). Map Projection Education in General Cartography Textbooks: A Content Analysis. *Cartographic Perspectives*, 90, 6–30. <https://doi.org/10.14714/CP90.1449>
- Kessler, F. C., & Slocum, T. A. (2011). Analysis of Thematic Maps Published in Two Geographical Journals in the Twentieth Century. *Annals of the Association of American Geographers*, 101(2), 292–317. <https://doi.org/10.1080/00045608.2010.544947>
- Korat, O., & Shamir, A. (2004). Do Hebrew electronic books differ from Dutch electronic books? A replication of a Dutch content analysis. *Journal of Computer Assisted Learning*, 20(4), 257–268. <https://doi.org/10.1111/j.1365-2729.2004.00078.x>
- Kraak, M. J., & Ormeling, F. (1996). *Cartography: Visualization of Spatial Data*. Longman.
- Kraak, M. J., & Ormeling, F. J. (2020). *Cartography: Visualization of Geospatial Data* (4th ed.). CRC Press.
- Kraak, M.-J., & Ormeling, F. (2010). *Cartography, Third Edition: Visualization of Spatial Data* (3rd ed.). Guilford Publications.
- Kraak, M.-J., & Ormeling, F. J. (2003). *Cartography: Visualization of Geospatial Data* (2nd ed.). Pearson Education.
- Krippendorff, K. (2018). *Content Analysis: An Introduction to Its Methodology*. SAGE Publications.
- Larson, S. R. (1974). *A Study of Some Textual Variations in the Book of Mormon Comparing the Original and the Printer's Manuscripts and the 1830, the 1837, and the 1840 Editions*.

- Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2015). *Geographic Information Science and Systems*. John Wiley & Sons.
- Mayring, P. (2019). Qualitative Inhaltsanalyse – Abgrenzungen, Spielarten, Weiterentwicklungen. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 20(3), Article 3. <https://doi.org/10.17169/fqs-20.3.3343>
- Meng, L. (2003). *Missing Theories and Methods in Digital Cartography (English)*. CD-Proceedings of the 21st International Cartographic Conference Durban 2003. <https://media-tum.ub.tum.de/1353500>
- Merriam-Webster. (n.d.). *Definition of EDITION*. Merriam-Webster Dictionary. Retrieved 21 August 2022, from <https://www.merriam-webster.com/dictionary/edition>
- Milson, A. J., & Alibrandi, M. (2008). *Digital Geography: Geospatial Technologies in the Social Studies Classroom*. IAP.
- Mittelstädt, F.-G. (1989). Vorworte in deutschen Schulatlanten. *Kartographische Nachrichten*, 39(6), 212–216.
- Monmonier, M. (1982a). *Computer-assisted cartography: Principles and prospects*. Prentice Hall.
- Monmonier, M. (1982b). Cartography, geographic information, and public policy. *Journal of Geography in Higher Education*, 6(2), 99–107. <https://doi.org/10.1080/03098268208708847>
- Monmonier, M. (2015). *The History of Cartography, Volume 6: Cartography in the Twentieth Century*. University of Chicago Press.
- Muehlenhaus, I. (2011). Genealogy That Counts: Using Content Analysis to Explore the Evolution of Persuasive Cartography. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 46(1), 28–40. <https://doi.org/10.3138/carto.46.1.28>
- Oxford Dictionary. (n.d.). *Definition of EDITION*. Oxford Advanced Learner's Dictionary. Retrieved 21 August 2022, from <https://www.oxfordlearnersdictionaries.com/definition/english/edition?q=edition>
- Pickles, J. (1999). Cartography, digital transitions, and questions of history. *Proceedings of 19th International Cartographic Conference*.
- Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3), 130–137. <https://doi.org/10.1108/eb046814>
- Putu, S., Surata, S., Sri, G., Jayantini, R., & Lansing, J. (2014). *Exploring Community Capital of the Balinese Subak Cultural Heritage: A Content Analysis of Participatory Maps*. 2320–8163.
- Robinson, A. H. (1982). *Early Thematic Mapping in the History of Cartography*. University of Chicago Press.
- Salemans, B. J. P. (1994). Comparing text editions with the aid of the computer. *Computers and the Humanities*, 28(3), 133–139. <https://doi.org/10.1007/BF01830734>
- Schilling, J. (2006). On the pragmatics of qualitative assessment: Designing the process for content analysis. *European Journal of Psychological Assessment*, 22(1), 28.
- Scholz, M. A., Huynh, N. T., Brysch, C. P., & Scholz, R. W. (2014). An Evaluation of University World Geography Textbook Questions for Components of Spatial Thinking. *Journal of Geography*, 113(5), 208–219. <https://doi.org/10.1080/00221341.2013.872692>

- Schreier, M. (2012). *Qualitative Content Analysis in Practice*. SAGE.
- Shen, Z., Zhang, R., Dell, M., Lee, B. C. G., Carlson, J., & Li, W. (2021). *LayoutParser: A Unified Toolkit for Deep Learning Based Document Image Analysis* (arXiv:2103.15348). arXiv. <https://doi.org/10.48550/arXiv.2103.15348>
- Silge, J., & Robinson, D. (2016). tidytext: Text Mining and Analysis Using Tidy Data Principles in R. *Journal of Open Source Software*, 1(3), 37. <https://doi.org/10.21105/joss.00037>
- Silge, J., & Robinson, D. (2017). *Text Mining with R: A Tidy Approach*. O'Reilly Media, Inc.
- Slocum, T. A. (1999). *Thematic Cartography and Visualization* (1st ed.). Prentice Hall.
- Slocum, T. A., McMaster, R. B., Kessler, F. C., & Howard, H. H. (2005). *Thematic Cartography and Geographic Visualization* (2nd ed.). Pearson/Prentice Hall.
- Slocum, T. A., McMaster, R. B., Kessler, F. C., & Howard, H. H. (2022). *Thematic Cartography and Geovisualization* (4th ed.). CRC Press. <https://doi.org/10.1201/9781003150527>
- Slocum, T. A., Slocum, T. A., McMaster, R. B., Kessler, F. C., & Howard, H. H. (2013). *Thematic Cartography and Geovisualization* (3rd ed.). Pearson Education Limited.
- Taylor, D. R. F. (1985). The educational challenges of a new cartography. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 22(4), 19–37.
- Taylor, D. R. F. (1991). A conceptual basis for cartography: New directions for the information era. *The Cartographic Journal*, 28(2), 213–216. <https://doi.org/10.1179/000870491787859214>
- The History of Cartography Project. (2015, May 18). *Volume 6, Cartography in the Twentieth Century*. <https://geography.wisc.edu/histcart/>
- Tobler, W. R. (1959). Automation and cartography. *Geographical Review*, 49(4), 526–534.
- Tomlinson, R. F. (1968). *A Geographic Information System for Regional Planning*. Department of Forestry and Rural Development, Government of Canada.
- Tomlinson, R. F. (1987). Current and potential uses of geographical information systems The North American experience. *International Journal of Geographical Information Systems*, 1(3), 203–218. <https://doi.org/10.1080/02693798708927808>
- Tyner, J. A. (2005). Elements of Cartography: Tracing Fifty Years of Academic Cartography. *Cartographic Perspectives*, 51, 4–13. <https://doi.org/10.14714/CP51.392>
- Visvalingam, M. (1990). Trends and concerns in digital cartography. *Computer-Aided Design*, 22(3), 115–130. [https://doi.org/10.1016/0010-4485\(90\)90070-S](https://doi.org/10.1016/0010-4485(90)90070-S)
- Wendling, A. E. (2005). Comparing Two Editions of Marx–Engels Collected Works. *Socialism and Democracy*, 19(1), 181–189.
- Wickham, H. (2014). Tidy Data. *Journal of Statistical Software*, 59, 1–23. <https://doi.org/10.18637/jss.v059.i10>
- Wilson, M. W. (2015). New lines? Enacting a social history of GIS. *The Canadian Geographer / Le Géographe Canadien*, 59(1), 29–34. <https://doi.org/10.1111/cag.12118>

Appendix A

Interview questions for the authors of *Thematic Cartography and (Geo)Visualization*.

Section I

Below are the points that I would like to hear your opinion about.
They are general and open questions concerning your book.

1. What made you decide to write this textbook in the first place?
2. What kind of materials did you have available before starting to write the book?
3. What was the key motivation for each later edition to be published?
4. What was the writing process like when examining and updating the contents?
5. Have you recognised any changes in the cartographic teaching environment that the rapid growth of digital technologies has brought?
6. How has the relationship between cartography and the digital technology evolved throughout the editions of the book?

Section II

Below are the updates in the book that I would like to contextualise why they were added or deleted. Unlike the ones above, I do not expect a lot of discussion about these, and the answers can be succinct (e.g., “because the website was deprecated”). It is also okay if you cannot recall the reasons.

7. [Deleted in the third ed.]
Why was the subchapter «*Low-Volume Print Reproduction*» (i.e., laser, ink-jet, thermal-wax transfer, and dye-sublimation printers) deleted?
8. [Added in the third ed.]
Why was the example LandScan added to the chapter «*Dot and Dasymmetric Mapping*» only in the third edition, when its earliest version was introduced in 1998?
9. [Deleted in the third ed.]
Why was the subchapter «*Multimedia Systems*» deleted?
10. [Deleted in the third ed.]
Why was the example Animation of Climate Change (Weber & Buttenfield) deleted?
11. [Deleted in the third ed.]
Why was the example of the physical model of Death Valley deleted?
12. [Added in the second ed.]
Was the project Virtual Los Angeles ever successfully transferred to Windows-based systems and become accessible to the public?

Appendix B

Interview questions for the authors of *Cartography: Visualization of (Geo)Spatial Data*.

Section I

Below are the points that I would like to hear your opinion about.
They are general and open questions concerning your book.

1. What made you decide to write this textbook in the first place?
2. What kind of materials did you have available before starting to write the book?
3. What was the key motivation for each later edition to be published?
4. What was the writing process like when examining and updating the contents?
5. Have you recognised any changes in the cartographic teaching environment that the rapid growth of digital technologies has brought?
6. How has the relationship between cartography and the digital technology evolved throughout the editions of the book?

Section II

Below are the updates in the book that I would like to contextualise why they were added or deleted. Unlike the ones above, I do not expect a lot of discussion about these, and the answers can be succinct (e.g., “because the website was deprecated”). It is also okay if you cannot recall the reasons.

7. [Deleted in the second ed.]
Why were two examples ‘TOP10vector’ from the Netherlands and ‘ATKIS’ from Germany deleted?
8. [Added in the third ed.]
Why did you newly introduce the subchapter “Geographical Names”?
9. [Deleted in the third ed.]
Why was the website ‘Odden’s Bookmarks’ deleted?
10. What was the motivation to start providing hyperlinks of web resources in the textbook since the third edition?
11. Were there any criteria about which hyperlinks to add and which to leave out in the textbook?

Appendix C

The R codes used for calculating the word frequency of *Thematic Cartography and (Geo)Visualization* are uploaded to this repository:

https://github.com/sanghostevielee/thesis/blob/main/word_count_Slocum.R

The R codes used for calculating the word frequency of *Cartography: Visualization of (Geo)Spatial Data* are uploaded to this repository:

https://github.com/sanghostevielee/thesis/blob/main/word_count_Kraak.R

The R codes used for performing the bigram analysis of *Thematic Cartography and (Geo)Visualization* are uploaded to this repository:

https://github.com/sanghostevielee/thesis/blob/main/bigram_Slocum.R

The R codes used for performing the bigram analysis of *Cartography: Visualization of (Geo)Spatial Data* are uploaded to this repository:

https://github.com/sanghostevielee/thesis/blob/main/bigram_Kraak.R