PhD-FHSE-2024-012
The Faculty of Humanities, Education and Social Sciences

DISSERTATION

Defence held on 08/04/2024 in Esch-sur-Alzette
to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN *HISTOIRE*

by

Aida HORANIET IBAÑEZ

Born on 26 October 1985 in Madrid (Spain)

THE LUXEMBOURG TIME MACHINE: AN INTERDISCIPLINARY EXPLORATION INTO THE VISUALIZATION OF COMPLEX DATA FROM THE PAST

Dissertation defence committee

Dr Andreas Fickers, dissertation supervisor
Professor, Université du Luxembourg

Dr Reinhard Schneider, Chairman
Professor, University of Luxembourg

Dr Sean Takacs, Vice Chairman
Professor, University of Luxembourg

Dr Johanna Drucker
Breslauer Professor, University of California

Dr. Christophe Schuwey
Associate Professor, Université Bretagne Sud

The work presented in this dissertation was conducted at the:

Luxembourg Centre for Contemporary and Digital History

The work presented was sponsored by:

Institute for Advanced Studies
University of Luxembourg

Doctoral Program in History
University of Luxembourg

Affidavit

I, Aida Horaniet Ibañez, hereby confirm that the PhD dissertation entitle “The LuxTIME Machine: An interdisciplinary exploration into the visualization of complex data from the past” has been written independently and without any other sources than cited.

Luxembourg, Luxembourg 15th of February 2024

Aida Horaniet

Aida H.I

Acknowledgements

First, I would like to thank myself for taking the courageous decision to pursue a PhD after many years working in the industry, a decision that has allowed me to dedicate the necessary time and enjoy such an enriching experience. Looking back from the end of the path, it is clear to me, without any doubt, I would do it again. I want to thank my husband, Raimon, and my daughter, Maeva, for embracing the adjustments required by this project and for patiently listening to me talk endlessly about data visualization.

I would like to express my sincere gratitude to my supervisor, Andreas Fickers, for guiding me through the writing of this thesis with tireless optimism. I really appreciate your consistent support, always respecting my interests and work approach. Thank you for opening doors, that I would not have dared to open, and which have inspired me and allowed me to share my work with others. It has been a real pleasure to work with you these three years. This is how interdisciplinary work should always be, based on real interest, flexibility, and mutual learning. I would also like to thank Brigitte Melchior, the guiding compass on this road, always available, resourceful, kind, and knowledgeable. I extend my sincere gratitude to Johanna Drucker, not only for the wealth of invaluable knowledge shared over the years, but also for dedicating time to provide detailed and motivating feedback. Additionally, I appreciate her availability for meetings across different time zones, even at very early hours. Thank you for being part of my thesis committee together with Reinhard Schneider, whom I would also like to thank.

Thanks to my “adoptive team” during these years, the Digital Research Infrastructure (DRI) team at the Centre for Contemporary and Digital History (C^{2}DH). The DRI has been for me the safe space to share the progress of my project, ask questions about technology and research, and keep updated about other projects at the centre. More importantly, the DRI meetings have been a forum for sharing joys and frustrations along the way, essential conversations for me in this interdisciplinary adventure. Many thanks to Lars, Elisabeth, Daniele and everyone else. Another important pillar of this process has been the other PhDs, thanks to Carmen, Suzanna, Daniel, Johanna, Veronique, Sarah, Irene, and all the others, for our endless coffees, debates, and encouragement. I have learnt from all of you, and I hope to work with many of you in the future, because you are a group of brilliant people.

A special thank you to all the members of the LuxTIME Machine Project (LuxTIME). First, I would like to thank Dagny Aurich, the other PhD student with whom I have shared an exciting interdisciplinary journey, along which we have learnt from each other, and navigated difficulties with great success. Second, I would like to thank the other LuxTIME members at LCSB and LIST: Emma Schymanski, Laurent Pfister and Christophe Hissler. Thank you for all your feedback and participation in the LuxTIME seminar series and the writing workshops. Finally, thanks to all the other participants that have in some way contributed to the LuxTIME project.

Furthermore, I extend my gratitude to the University of Luxembourg, and the Institute for Advanced Studies for funding this project; and to all the employees at the C^{2}DH. You are an unlimited source of inspiration, and I could not have carried out my thesis in a better context. With a background in engineering and data science, developing my thesis at the C^{2}DH, in collaboration with historians, chemists and hydrologists, seemed challenging to say the least. However, it has been a very rewarding learning process. Especially, being exposed to humanities practices, has made me discover a new part of myself and my research in data visualization. I believe more than ever in interdisciplinary collaborations, I am *an*

interdisciplinarian, especially interested in collaborations among epistemologically distant disciplines.

Finally, I am thankful to my family, friends, and neighbours, who have been that support network, on a day-to-day basis when I have needed them. Thank you all.

Abstract

This research studies and implements multiple approaches to data visualization in an interdisciplinary context. This includes the use of visualization to explore and communicate data, but also to collect, interpret and reflect on the process itself. Different types of data visualization techniques are studied along the epistemological continuum that encompasses quantitative, qualitative, critical, and creative methods; and after, put into practice beyond the discipline of origin in our case study, the LuxTIME Machine (LuxTIME). LuxTIME studies the exposome in the south of Luxembourg, the Minett region, during the industrialization period (1890 – 1990). Exposome research studies the impact of different exposures (e.g., lifestyle, air pollution, work) on the health of the local population. In collaboration with researchers in history, environmental cheminformatics, and eco-hydrology; we collect historical data sources from natural and social archives across different data sources in Luxembourg (e.g., libraries, national and local archives, scientific data from research centres, government agencies, digital newspapers). Based on this, we propose the concept of the *historical exposome*, as an alternative way to study the impact of different exposures on the health of the population. Natural archives are physical objects, collected, processed, and deposited in the environment without the interference of the human species, while social archives are the result of a conscious collection of past evidence, be it for cultural, political, or economic reasons. Since finding evidence in natural archives has proven to be challenging, looking at social archives opens new possibilities to encounter historical evidence. Throughout this research, we draw on two concepts: *Trading Zones* and *Thinkering*. The interrogation of the trading zone between the fields of history, ecohydrology, environmental cheminformatics and, of course, data visualization, allows us to perform methodological and epistemological reflections on how to critically test the analytical potential of a multi-layered research design. Thinkering – the action of playful experimentation with technological and digital tools (for the interpretation and presentation of history) – provides us with the framework to experiment with different data visualization techniques from distant epistemologies.

Table of Contents

List of Figures .... *viii*

List of Publications .... *ix*

Introduction *1*

- **The Luxembourg Time Machine project: Interrogating the trading zone between the exposome, history and data visualization.** 1
- **History of data visualization** 5
- **Data visualization across the epistemological continuum** 11
- **The Minett region: environmental history and exposomics research** 12
- **The data visualization toolbox** 15
- **Dissertation outline** 15

Chapter 1: Digital humanities at the intersection of three approaches to data visualisation: statistical graphics, data humanism, and humanistic interpretation *17*

Chapter 2: Historical exposomics: a manifesto *37*

Chapter 3: Simulating and visualizing data in environmental history: Airbone dust concentration from the Belval plant in Luxembourg (1911-1997) *52*

Chapter 4: How can data visualization support interdisciplinary research? LuxTIME: studying historical exposomics in Belval *85*

Conclusion and further perspectives *107*

- **Breaking out of duality** 107
- **Advancing data visualization and interdisciplinary research** 108
- **LuxTIME: From Belval to a national case study** 110
- **Humanities theories and data visualization** 112

Bibliography *114*

Appendix I: Information inventory and exposome mapping *121*

Appendix II: Entity Relationship Diagram *125*

List of Figures

Figure 1 Positioning of the different disciplines involved in LuxTIME along the epistemological continuum.........2
Figure 2 Conceptual map of the knowledge within the different disciplines and their contribution to the scope of LuxTIME. Extracted from Chapter 4.........3
Figure 3 Evolution of different topics across domains which demonstrates changing practices. Extracted from Chapter 4.........5
Figure 4 Examples of data visualization with predominant elements of different methods.....11
Figure 5 Examples of historical data visualizations.........7
Figure 6 Alternative approaches to data visualization.........9
Figure 7 Data physicalization examples.........10
Figure 8 LuxTIME Data Model.........141

List of Publications

The 4 publications listed below are the core of my thesis:

1. **Research article (first authorship). Status: Submitted**

Aida Horaniet Ibañez, Digital Humanities at the intersection of three approaches to data visualisation: statistical graphics, data humanism, and humanistic interpretation, *Journal of Digital Humanities Benelux 2023*.

2. **Research article (first co-authorship). Status: Published**

Dagny Aurich, Aida Horaniet Ibañez, Christophe Hissler, Simon Kreipl, Laurent Pfister, Emma L Schymanski, Andreas Fickers, Historical exposomics: a manifesto, *Exposome*, Volume 3, Issue 1, 2023, osad007, https://doi.org/10.1093/exposome/osad007

3. **Research article (first co-authorship). Status: Submitted**

Dagny Aurich, Aida Horaniet Ibanez, Jens van de Maele, Simulating and visualising data in environmental history: Airborne dust concentration from the Belval plant in Luxembourg (1911-1997), *Journal of Digital History*.

4. **Research article (first co-authorship). Status: Published**

Dagny Aurich, Aida Horaniet Ibanez, How can data visualization support interdisciplinary research? LuxTIME: studying historical exposomics in Belval Frontiers in Big Data, Frontiers in Big Data 6:1164885, doi: 10.3389/fdata.2023.1164885.

As part of this PhD, the following publications were also completed:

5. **Encyclopaedia article. Status: Accepted.**

Hilke Brockmann, Roger Fernandez-Urbano (Eds.) Encyclopedia of Happiness, Quality of Life and Subjective Wellbeing. Edward Elgar Publishing, 2024, doi: https://doi.org/10.4337/9781800889675. Aida Horaniet Ibañez, Dagny Aurich *Environmental Science and Happiness*.

6. **Online journal article. Status: Published.**

Aida Horaniet Ibañez, Daniel Richter, Joëlla Van Donkersgoed, From The Historical Archive To The Citizens: Visualizing Census Data From Brill Street in 1922, Journal of the Data Visualization Society Nightingale, March 2023, https://nightingaledvs.com/from-the-historical-archive-to-the-citizens-visualizing-census-data-from-brill-street-in-1922/ (accessed 7th December 2023).

7. **Contribution to the Dagstuhl report about visualization literacy (follow-up article ongoing).**

Derya Akbaba, Alejandro Benito-Santos, Jeremy Douglass, Jean-Daniel Fekete, Aida Horaniet Ibañez, Aida; Jan Horstmann, Matthieu Jacomy, Steffen Koch, Yanni Loukissas, Isabel Meirelles, David Pao, Florien Windhager, Visualization Literacy, Dagstuhl Reports, Dagstuhl Seminar 23381, 2023.

8. **A survey of data visualisation evaluation methods in the Digital Humanities (ongoing).**

Alejandro Benito-Santos, Alfie Abdulrahman, Florian Windhager, Eva Mayr, Rabea Kleymann.

Introduction

The Luxembourg Time Machine project: Interrogating the trading zone between the exposome, history and data visualization.

The value of interdisciplinary research[1] as well as all the associated challenges are well known. To better understand complex issues, studied from different perspectives, researchers must navigate through the rules, beliefs, and perceptions of the participants about knowledge production (Campbell, 2005). Not just laying one discipline alongside another (Gibbons & Nowotny, 2001) in a superficial manner where "traditional modes of work are patched together under a new label" (Rhoten, 2004, p. 6), but truly "learning, unlearning and relearning across disciplines" (Rhoten, 2004, p. 11). In "Becoming Interdisciplinary", Willard MacCarty says that "if disciplines are epistemic cultures in the anthropological sense, then we have not just silos or islands of knowledge but islands populated by communities of knowers, their languages, histories and artifacts" (Schreibman et al., 2016, p. 76). He states that "becoming interdisciplinary both rides the urge to know and struggles to hang against the possibility of being thrown at it" (Schreibman et al., 2016, p. 79). Just as there are researchers who are resistant to venture outside of their disciplinary boundaries, there are many others "driven to the edges of their fields by a shift in their epistemological values and intellectual interests" (Rhoten, 2004, p. 8), as the participants in this project, The Luxembourg Time Machine Project (LuxTIME), demonstrated.

LuxTIME is an interdisciplinary research project funded by the Institute for Advanced Studies (IAS) of the University of Luxembourg[2]. Three research institutes are involved in this so-called "Audacity Project", the Centre for Contemporary and Digital History (C²DH), the Luxembourg Centre for Systems Biomedicine (LCSB) and the Luxembourg Institute of Science and Technology (LIST). LuxTIME uses the industrialization of the Minett region as a testbed for methodological and epistemological reflections on how to study the impact of environmental changes on the health of the local population in a long-term perspective. The Belval-case, on which this research is focused, critically tests the analytical potential of a multi-layered research design that could be expanded into a national case study, including different types of data on this and other topics, from many institutions. In the conclusions and future perspectives, I will discuss in detail how this project lays the groundwork for a future national project. The Belval LuxTIME is one of the local time machines of the European Time Machine project[3], which aims to leverage digital technologies and infrastructures to create a collective digital information system that maps Europe's history.

Scholars are unanimous about the fact that there is no single phenomenon of interdisciplinarity, but multiple interdisciplinarities exist (Fickers & Tatarinov, 2022; Katri Huutoniemi, 2017). This manuscript is the result of the experience in an interdisciplinary project involving researchers in the natural (eco-hydrology, environmental cheminformatics) and applied sciences (data visualization), and the humanities (environmental history). As I will discuss throughout the different chapters, LuxTIME expands along the entire epistemological

---

[1] Later in this section, I discuss which parts of the project are multidisciplinary, interdisciplinary, or transdisciplinary. Until then, I use the term interdisciplinary in a broad sense, as the integration of different disciplines, fields, or bodies of knowledge in research activities.

[2] https://wwwfr.uni.lu/ias

[3] https://www.timemachine.eu

continuum, manoeuvring across quantitative, qualitative, critical, and creative methods. This involves understanding different definitions of what data is, the use of different research methods and therefore the research questions they can answer (e.g., how much, what, how, why?), as well as the different evaluation systems.

To create a project relevant to all the disciplines involved, it is necessary to accommodate different (often conflicting) definitions of quality (Katri Huutoniemi, 2017), publication standards (e.g., reference standards, article structure, use of footnotes) and practices (e.g., authorship, journals, publication pace). Most importantly, to find a common vocabulary that allows for exchange and useful peer review. Sometimes, positions on key issues are so fundamental to the researcher's beliefs that they are not acknowledged until conflict arises (Campbell, 2005). Some of these beliefs in the context of LuxTIME, as we will demonstrate in the next chapters, include the definition of data itself, at what pace data and data visualization must be consumed and understood, what is considered significant and how it is measured, or to what extent ambiguity and multiple narratives are accepted. In Figure 1 we show the position of the different disciplines involved in LuxTIME along the epistemological continuum.

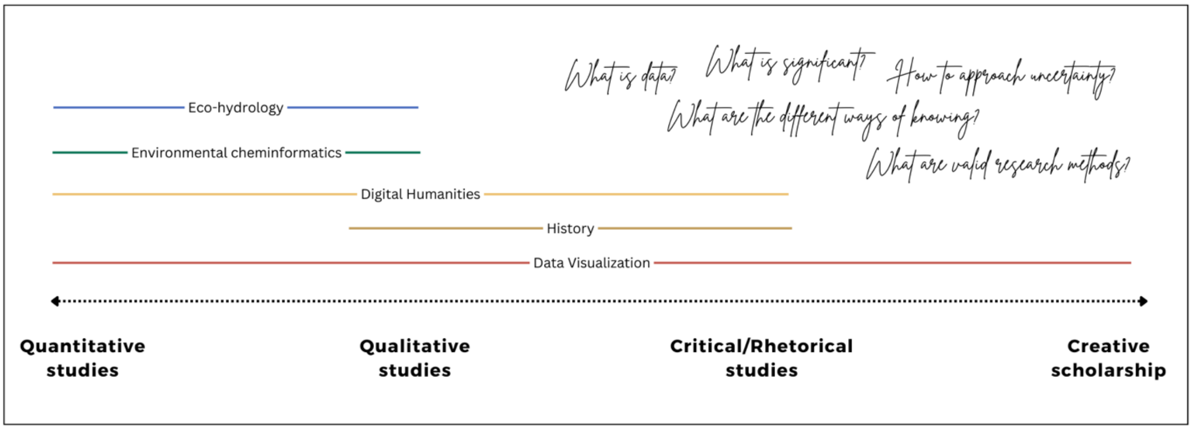


Figure 1 Positioning of the different disciplines involved in LuxTIME along the epistemological continuum.

All these challenges create undeniable opportunities for the researchers to grow beyond their own disciplines and epistemological standards, through information sharing networks (Rhoten, 2004). However, all this takes time. Time to read unfamiliar literature and to share knowledge, that often implies starting with the basics. More specifically, we are talking about a historian explaining what a historical archive is and how it is consulted. A chemist explaining where it is possible to find samples in nature to measure pollution, the types of pollution and the health impact of different concentrations. A hydrologist explaining what ground water is and what can be measured. And a data visualization researcher, explaining how to select a chart or how to check a visualization for colour-blindness. These conversations would never take place within the disciplines themselves, and furthermore, they do not happen once but repeatedly throughout the project, as it is a learning process. The implementation also takes more time. For example, in the case of historical archives, the number of visits and the methods of collecting information require more time. It clearly takes a data scientist or a chemist longer to explore an archive than a trained historian, but at the same time, they approach the reading with a new perspective and potentially find new insights (e.g., when reading reports about pollution that include chemical concentrations).

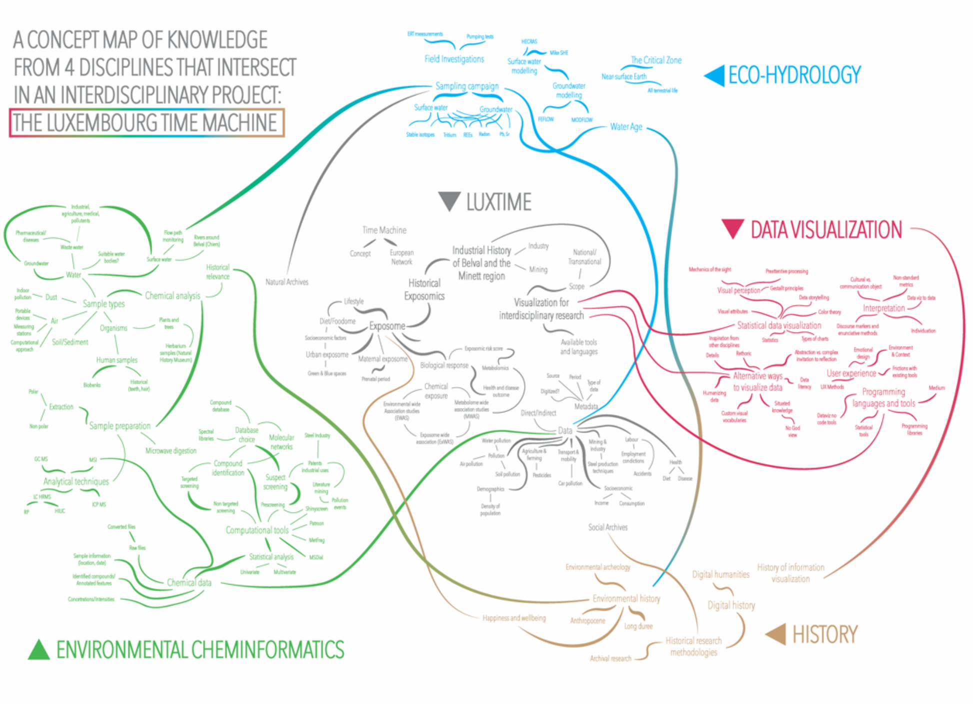


Figure 2 Conceptual map of the knowledge within the different disciplines and their contribution to the scope of LuxTIME. Extracted from Chapter 4.

It takes time to develop collegiality, to establish communication methods, to agree on publishing expectations and protocols, to write together, to organize research seminars and to participate in conferences and events across fields. Figure 2, which will be discussed in detail in Chapter 4, gives a first impression of the complex exchange that has to take place between disciplines to define the scope of an interdisciplinary project like LuxTIME, as shown in the centre of the image. All this effort is hard to measure and ultimately it must be reflected in the research output, the academic journals, and the conferences; creating value for all the disciplines involved. Stefan Jänicke refers to this effort as "the balancing act" (Jänicke, 2016), or how to generate beneficial research results for the fields of data visualization and digital humanities. In this research, the LuxTIME Machine, that *balancing act* extends to the fields of environmental cheminformatics and eco-hydrology, to study the exposome, and more specifically historical exposomics, a concept that will be discussed in Chapter 2.

This research builds on the idea of the "trading zone", initially defined by Peter Galison as "an area in which radically different activities could be locally, but not globally, coordinated" (Galison, 1996, p. 119). Max Kemman analyzed the trading zone between historians and computer experts according to three dimensions: *engagement* i.e., the extent to which the communities to meet and interact; *power relations* i.e., the measure in which a community has a stronger negotiation power to decide goals and practices; and *changing practices* i.e., whether the trading zone remains an interaction of distinct communities, or merges into a singular community of shared practices (Kemman, 2021). Below is a brief analysis of this research, according to these three dimensions.

Engagement – In LuxTIME, a community of practice, i.e., group of people who engage in a process of collective learning in a shared domain (Wenger-Trayner, 2008), was created through regular interaction taking place in hybrid mode during project meetings, collaborative writing sessions, a seminar series on various aspects of the project, and informal discussions. During these meetings, there was a constant negotiation of common objectives and the responsibility of the participants, and the development of a shared jargon and practices.

Power relations – Research, as a group activity, is often influenced by power differences between members e.g., career stage, project leadership (Campbell, 2005) that are also approached differently in different organizations, for example concerning how formal structures and hierarchical layers define tasks, roles, and behaviors (Laloux & Wilber, 2014).

The already mentioned negotiations of the project goals and the individual responsibilities do not exist in a vacuum. "Negotiations are positioned in a broader system that influences the collaboration" (e.g., research centres in which the participants are employed, disciplinary backgrounds, funding structures) (Kemman, 2021, p. 45).

In the case of LuxTIME, this system included factors such as the background of the different researchers in four different disciplines, the different practices of the three research centres involved, or the different levels of experience.

Changing practices – As in most projects, in LuxTIME the collaboration moved from heterogeneous to more homogeneous as the project progressed, as we aligned with respect to the project's goals, terminology, and desired results. We gradually defined certain core areas of the project where the interaction was more fluid, we all learnt the common vocabulary and could develop further together. Figure 3, extracted from the visualizations discussed in Chapter 4, provides an early glimpse of these changing practices, increasingly homogeneous and aligned across disciplines. However, throughout this process of *changing practices*, we also observed which areas would remain at a level of interaction where we could all read articles from other disciplines (with considerable additional effort) but could not participate in an in-depth discussion or write an article about it. This is mainly observed in Figure 2, in relation to those topics that remain under the specific disciplines. In those cases, the research was developed in a multidisciplinary way, the disciplines operated side by side with little or no exchange. For example, *Historical Exposomics and High Resolution Mass Spectrometry (Aurich et al., 2021)*[4] is a publication written by the environmental cheminformatics and eco-hydrologists in LuxTIME; while *Digital Humanities at the intersection of three approaches to data visualization: statistical graphics, data humanism, and humanistic interpretation*, in Chapter 1, was written by the data visualization researcher. These two publications established the theoretical basis in the fields of visualization and the exposome, that informed the rest of the project, but were difficult to carry in an interdisciplinary manner due to the technical depth within each discipline.

---

[4] This publication is published as part of the Doctoral Thesis of the authors.

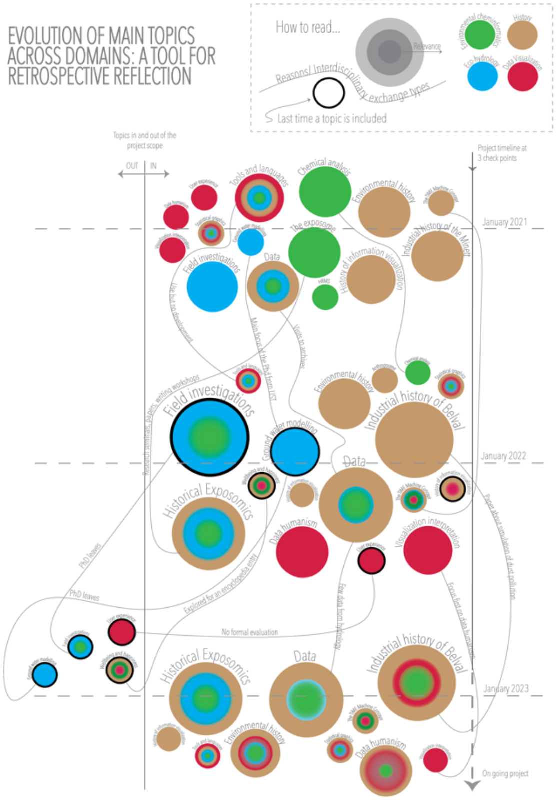


Figure 3 Evolution of different topics across domains which demonstrates changing practices. Extracted from Chapter 4.

On the other hand, *Historical Exposomics: A Manifesto*, in Chapter 2; *How can data visualization support interdisciplinary research? LuxTIME: Studying historical exposomics in Belval*, in Chapter 3; and *Simulating and visualizing data in environmental history: Airborne dust concentration from the Belval plant in Luxembourg (1911-1997)*, in Chapter 4, are examples of the result of a truly shared practice, where the different disciplines blend. Lastly, two transdisciplinary collaborations, not only crossing disciplines but also sectorial boundaries (i.e., industry collaborations), can be highlighted. First, in Chapter 1, the proposed classification is based on the study of practices beyond research, such as data journalism, education, scientific communication, or art; and informed by the exchange with stakeholders outside academia during conferences, science communication activities, and the author's work experience prior to this project. Second, the collaboration with Inspyro, a metallurgy consultant, discussed in Chapter 3, to select the adequate dispersion model and generate the initial simulations. In Chapter 3, we explore in detail and through visual means these three dimensions throughout the entire project.

What is the role of data visualization in the project? What does this balancing act consist of specifically in this research? To contribute to the field of data visualization, it is important to understand its history.

History of data visualization

Any data visualization requires a dialogue between the urge to simplify and the need to embrace complexity. This dialogue is influenced not only by the specific needs of the audience, but also by the paradigms of the discipline in which it is developed, the access to the data, the medium, the tools available to implement and disseminate the visualization and their technical complexity, and the usage context. The history of information visualization, that is the history of data visualization, has deep roots that reach into cartography, statistical graphics, and data visualization. It was initially hand-drawn, later etched on copperplate and manually colored, still later lithographed or photoetched, and finally produced using computer software (Friendly, 2008). The trade-off between detail and abstraction, functionality and aesthetics, reductionism and complexity, is not a debate exclusive to this field. We can observe it in the different movements of art history, natural sciences, or philosophy.

Numerous examples of visualizations are presented throughout history focusing on detail, granularity, and expression of complexity. Maps, genealogies, astronomical texts, theological texts, medical treatises, technical sketches of war instruments, botanical and zoological taxonomies, historical chronologies, economic changes, hydrology and changes in water levels, music theory, architectural designs, or mortality rates[5], among many others, show the finest details about the theme of the visualization, the associated variables, the narrative and annotations, the context, and even the production process. Some examples (see Figure 4) include the earliest known attempt to show changes in the position of the sun, the moon, and the planets through the year, from 950; Ramon Llull's Tree of Knowledge from 1305, a diagram to support reasoning; the Rose of Bohemia from 1677, using cartographic techniques to interpret the political situation; the Tableau Poléométrique from 1782, attributed to Fourcroy de Ramecourt, comparing cities according to their sizes; the Plan of organization of New York and Erie Railroad from 1855, representing the division of administrative duties and the number and class of employees engaged; Charles-Joseph Minard's flow map graphic on Napoleon's March on Moscow from 1869.

---

[5] List based on the collections published by (Rendgen, 2019) in History of Information Graphics, (Lima, 2014) in The Book of Trees: Visualizing Branches of Knowledge and (Lima, 2017) The Book of Circles: Visualizing Spheres of Knowledge.

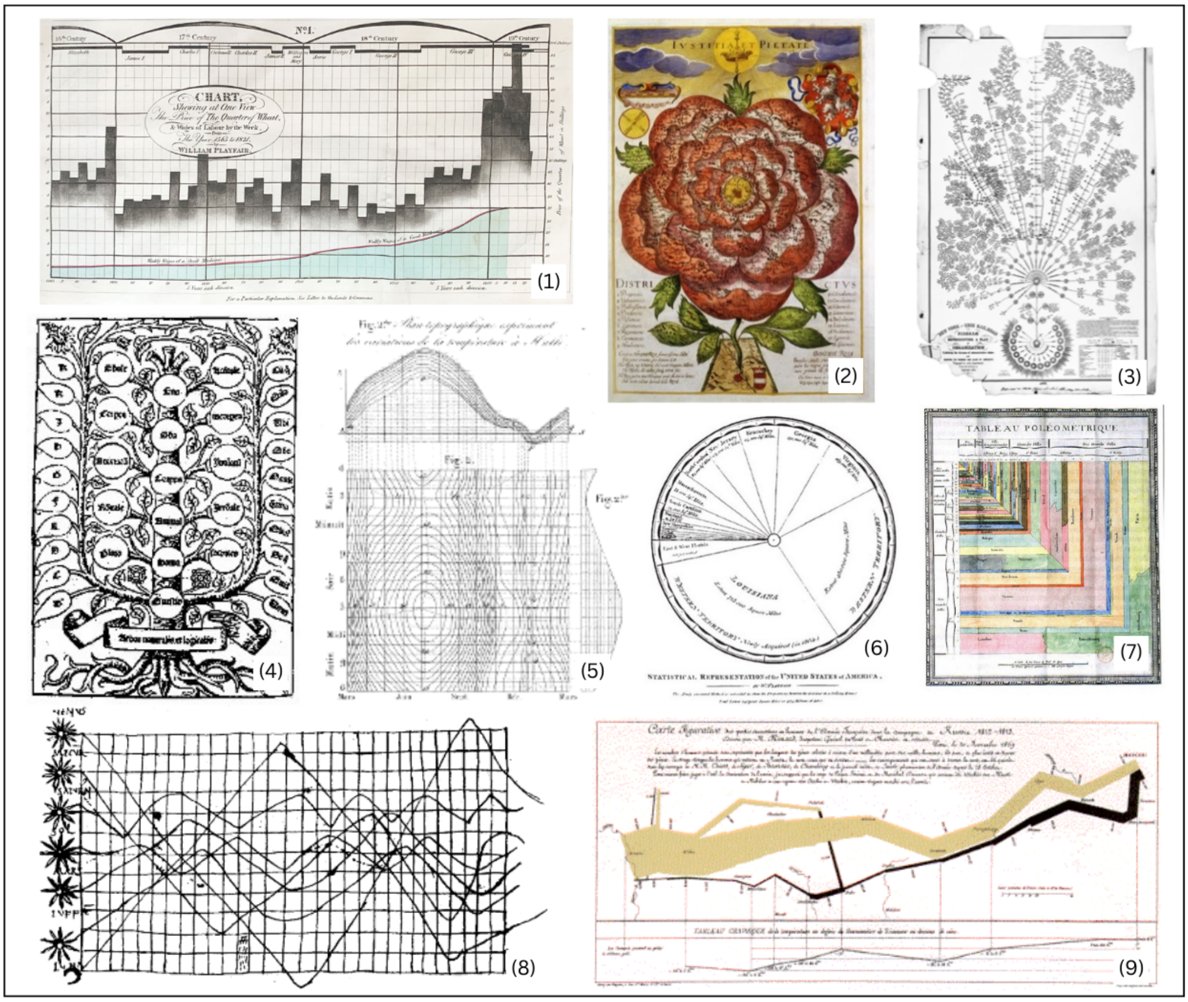


Figure 4 Examples of historical data visualizations.
(1) Combination of bar chart and line chart showing the evolution of wheat prices by Playfair (1821)(Bellhouse, 2023), (2) The Rose of Bohemia (1677) – extracted from History of Information Graphics (Rendgen, 2019), (3) Plan of Organization of New York and Erie Railroad by McCallum (1855) – extracted from The Book of Trees: Visualizing Branches of Knowledge (Lima, 2014), (4) Llull's Tree of Knowledge (1305), (5) Lalanne contour diagram (1845), (6) Playfair's 1805 Statistical Representation of the U.S.A. (1805), (7) Fourcroy's Proportional Squares (1782), (8) Planetary Movements Diagram (950), (9) Minard's March on Moscow (1869). Figures 4-9 have been extracted from The Milestones Project datavis.ca (Friendly & Denis, n.d.)

During the first half of the nineteenth century, thanks to the technical and design innovations, all the modern forms of data display in statistical graphics were invented, such as bar charts, line graphs, pie charts, histograms, time-series plots, or contour plots (Bederson & Shneiderman, 2003). These forms of data display still dominate the data visualization field today. Statistical graphics developed around four main problems: spatial organization, discrete comparison, continuous distribution, and multivariate distribution and correlation (Beniger & Robyn, 1978). Numerous studies have since proposed different taxonomies and frameworks for its study, use and evaluation[6], many of them will be referenced in the following chapters when discussing specific techniques.

---

[6] Among many other, the following can be highlighted: Exploratory data analysis (Tukey, 1977), Bertin's Semiology of Graphics: Diagrams, Networks and Maps (Bertin & Berg, 2010), Visual Explanations (Tufte, 1997) and the Visual Display of Quantitative Information (Tufte, 1999), The Craft of Information Visualization: Readings and Reflections (Bederson & Shneiderman, 2003), Information Visualization: Perception for Design (Ware, 2004), Show me the numbers: designing tables and graphs to enlighten (Few, 2012), Data Visualization: A handbook for data driven design (Kirk, 2016), Visualization analysis and design (Munzner, 2015), The Truthful Art: data,

During the last quarter of the 20th century, the development of interactive statistical computing systems opened new paradigms of interactivity and automation. A range of visualization methods and data types were available in desktop computers. The storage as well as the computing capacity of those computers was however quite limited, making it necessary to work only with aggregations and summary measures, using predefined charts. With the internet growth and the increased access to personal computers by the end of the 20th century, the generation and access to data rapidly increased. Digital data storage became more cost-efficient than paper-based storage, but the limitations to process the data and visualize it were still present. During the 21st century the systems that allow computers to store and process large data sets were developed and web-based computing infrastructure was available. Only then, the data analysis and visualization tools could be developed. However, after centuries of constraints, they were developed to visualize statistical graphics based on aggregated metrics for the most part.

As I will discuss in depth in Chapter 1, in the last decade, nevertheless, we have observed the increasing presence of alternative approaches to data visualization (see Figure 5), some based on statistical representations with slight variations of the layout and design, additional layers of information, or defining new visual vocabularies. Among many other information designers, data illustrators, visualization developers or data artists[7], we can highlight the work of Lupi, Posavec, Fragapane, Bremer, Wu, or Kuijpers. These types of visualizations arise from the use of data in alternative contexts and mediums, where more time is allowed for reading and understanding data visualizations, such as the data journalism from the South China Morning Post, the Financial Times Visual and Data Journalism, and the visual infographics of La Lettura (Corriere della Sera), all of which often use visualizations as an alternative or complement to the traditional articles. Data visualization, often referred to as data art[8] in this specific context, is also present in museums and exhibitions such as the Dear Data Exhibition at MoMA (Lupi & Posavec, 2015), featuring a year-long, analogue data drawing project; the Data Items: A Fashion Landscape Exhibition at MoMA (Lupi, 2017), a data-driven installation interpreting the exhibition, that combines standard charts with original, organic graphics; or the Mood Test (Domestic Data Streamers, 2013), a street data wall to analyse people’s attitudes towards life. Ultimately, this approach arises from a need to collect, explore, and communicate beyond the summary, in a wide range of domains and topics addressing social issues, historical research, or music, just to name a few examples.

---

charts, and maps for communication (Cairo, 2016), or Better data visualizations: a guide for scholars, researchers, and wonks (Schwabish, 2021).

[7] These are just some of the terms used to refer to the data visualization practitioners.

[8] The use of the term “data art” and its implications will be discussed in the Conclusions and future perspectives.

Figure 5 Alternative approaches to data visualization.
(1) Arrested Hong Kong Protesters (South China Morning Post, 2020), (2) Week 7 of the Dear Data project (Lupi & Posavec, 2016), (3) Space Wars (Nadieh Bremer, 2020), (4) Data Items: A Fashion Landscape (Lupi, 2017), (5) What's Cooking: Sequential data illustrations about the State of the industry in Data Visualization (Clark, 2022), (6) Submission to Viz for Social Good for the project Build up Nepal (Nwosu, 2022), (7) Journal of the Digital History Article Fingerprint (C2DH, University of Luxembourg, 2021), (8) Barriers to Reporting Sexual Violence (Frederica Fragapane, 2022), (9) Pebble visualization about the victims of the Shoah (C2DH, University of Luxembourg, 2023), (10) The Mood Test (Domestic Data Streamers, 2013), (11) The Poet's Journey (Michela Lazzaroni, 2023), (12) Updating Happiness: Wellcome Collection (Stefanie Posavec, 2021), (13) The Hotel New Hampshire (Sonja Kujpers, n.d.)

Another related area of research, data physicalization (see some examples in Figure 6) uses physical artifacts to convey data using the geometry and material properties, with the aim to help people explore, understand, and communicate data (Dataphys, 2021). Extensive research has been conducted in this field[9]. For example, Hostile Terrain 94: Deaths at the US/Mexico Border shows a prototype of a physical map made of toe tags representing the recovered bodies of migrants who died while crossing the US/Mexico border in the Sonoran Desert between 2000 and 2020. Multimodality is an emerging area of research in the digital humanities, especially in the field of data visualization, that aims at exploring visualization beyond the visual. It raises questions about how to augment visualizations with other experiential modes, not only using data physicalization but also sonification, AR/VR, olfaction or edibilization; and its impact on accessibility, viscerality or the data itself.

---

[9] Some examples of research in data physicalization include: Supporting the design and fabrication of physical visualizations (Swaminathan et al., 2014), Exploring the Challenges of Making Data Physical (Alexander et al., 2015), Opportunities and Challenges for Data Physicalization (Jansen et al., 2015), Tangible thinking: Materializing how we imagine and understand systems, experiences and relationships (Lockton et al., 2019), Narrative Physicalization: Supporting Interactive Engagement with Personal Data (Karyda et al., 2021) (Perovich et al., 2021) and Data to Physicalization: A survey of the Physical Rendering Process (Djavaherpour et al., 2021).

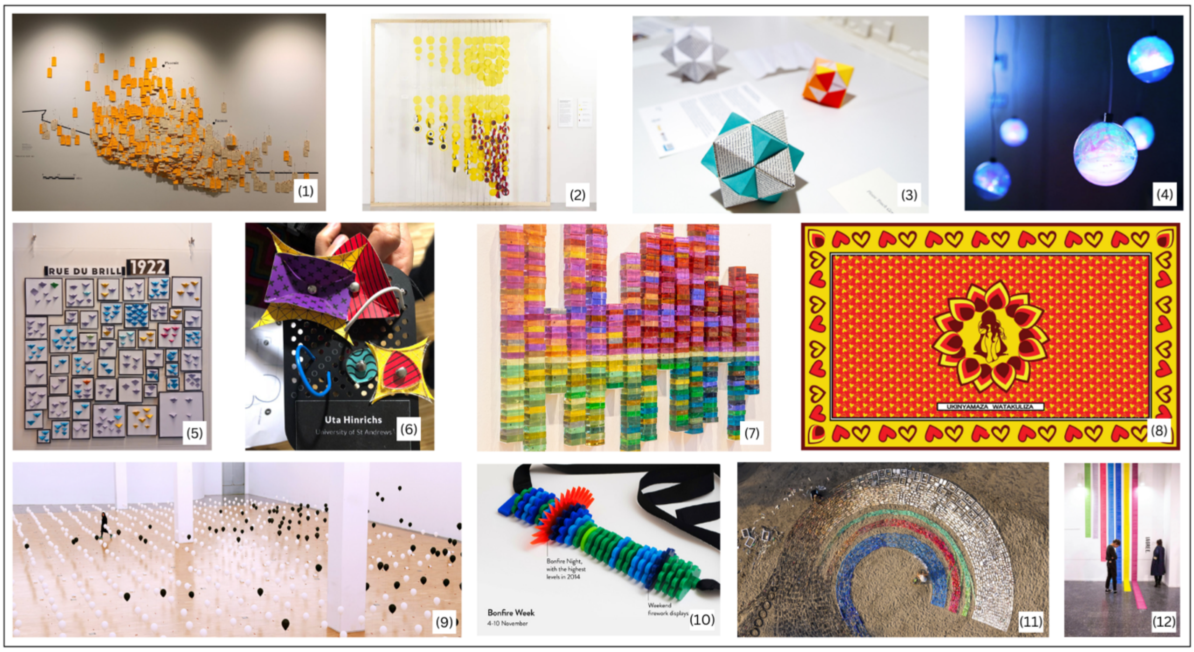


Figure 6 Data physicalization examples.
(1) Hostile Terrain 94: Deaths at the US/Mexico Border (Undocumented Migration Project, 2019), (2) Connecting Dots on Income Inequality (Giorgia Lupi, 2018), (3) Data Beyond Vision (Koeser et al., 2020), (4) One Amongst Many (Shirly Wu, 2019), (5) From the Historical Archive to the Citizens: Visualizing Census Data from Brill Street in 1922 (Aida Horaniet Ibañez et al., 2023), (6) Data Badges (Panagiotidou et al., 2020), (7) Childhood Hunger (Laurie Frick, n.d.), (8) Data-informed Khangas (Tanzania dLab, 2018), (9) A Visualization of Life and Death (Domestic Data Streamers, 2014), (10) Air Transformed: Wearable Data Object (Stefanie Posavec & Miriam Quick, 2015), (11) Perpetual Plastic (Liina Klauss et al., 2021), (12) Visualizing Terms and Conditions (Dima Yarovinsky, 2018)

As we will see in detail in Chapter 1, there are numerous tools to generate statistical graphics, without the need for advanced programming or design skills. On the other hand, tools that allow the production of more complex human-centric data visualizations that foster emotional connections, introduce multiple narratives or express interpretation, are limited or even non-existent without advanced technical knowledge. This imbalance has an impact on the development of the field of data visualization itself, as well as on the disciplines whose paradigms are more distant from the idea of reduction and simplification inherent to statistical graphics. Notably, the humanities, which are constrained to use the same methods and tools to visualize their data.

Data visualization across the epistemological continuum

As every history, the history of data visualization is made up of more than one past. It's "a process of layering, with previous layers shaping the structure of later layers and later layers determining whether older layers continue to influence the present" (Wimmer, 2023, p. 2). These layers draw on the history of computer science, visual studies, communication, design, statistics, and all the fields from which the data studied originate (e.g., hydrology, chemistry, politics, law) and therefore spreading along the entire epistemological continuum. Data visualization is a highly interdisciplinary discipline just by itself. In Figure 7, we see four examples of data visualizations, which, although using various methods (e.g., quantitative and qualitative, rhetorical and creative), present predominant elements of one or the other, to illustrate the idea of data visualization along the epistemological continuum.

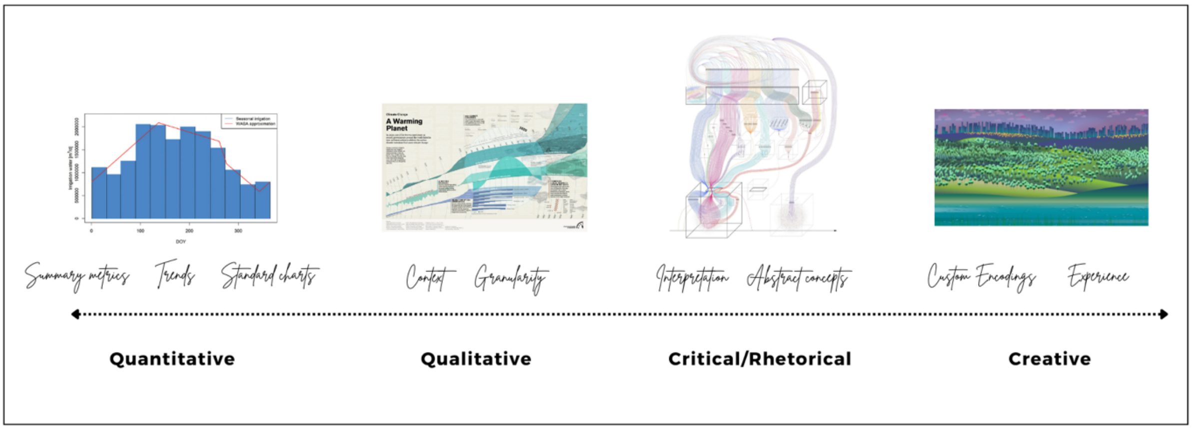


Figure 7 Examples of data visualization with predominant elements of different methods.
From left to right images extracted from (Pentagram, 2021; Sonja Kujpers, 2019; Stefanie Posavec, 2023; Voit et al., 2023).

One of the challenges when looking at the current state of the field is that much of the practice does not occur within academia and it is not limited to a specific medium. Data visualization is used in journalism, scientific dissemination, education, research, public administrations, non-profits' communication, personal projects, and art, among others; in digital media, in print, or created with physical materials (i.e., data physicalisation). Due to the variety of sources and their ephemeral nature (e.g., unarchived data visualization blogs that disappear when the domain is not maintained), it is difficult to collect and archive them systematically for study. However, we cannot ignore the presence of such material if we are to understand the state of the art. In Chapter 1, the effort is aimed at understanding the current situation, where these combined histories have led us to, the impact on the different disciplines, what needs are being met, and what gaps still exist. This analysis allows me to propose a classification of current approaches to data visualization, which I believe will help navigate the continuum of possibilities between quantitative and interpretative methods.

Current practices, as well as the current offerings of data visualization tools, are largely anchored in a type of graphics that supports mostly quantitative and some qualitative methods. In the humanities, the predominant use of such techniques has been widely discussed, with a critical approach to the definition of data (Drucker, 2020; Gitelman, 2013; Loukissas, 2019), the politics and ethics of representation (D'Ignazio & Klein, 2020), the challenges to communicate uncertainty and ambiguity (Panagiotidou, Lamqaddam, et al., 2022; Sánchez et al., 2019), the design process (Hinrichs et al., 2019), the role of the interfaces (Drucker, 2011), temporality (Drucker et al., 2022), non-representational approaches (Drucker, 2017b), the use of close (and distant) reading (Jänicke et al., 2015), and trust (Boyd Davis et al., 2021), among others. However, although some of the previous references refer to applied projects, the practice is still limited. Throughout four publications, which represent the chapters of this manuscript, we will explore theoretical and practical aspects of the different practices.

Chapter 1, and to some extent this entire manuscript, aims to guide us on several fronts. First, that of the “Visualization for the Digital Humanities", predominantly consisting of survey articles using quantitative (and some qualitative) methods, and carefully justified search queries; and the evaluation of existing and new techniques based on efficiency and error minimization. Second, that of the “Digital Humanities for Data Visualization”, which criticizes the current use of data visualization and its limitations when applied to the humanities and proposes the use of humanities theories and methods to create more suitable data visualizations. But also, that community of practitioners, who apply existing techniques, adapt, and reinvent them according to the needs of their projects, in different fields. To propose a classification to help us navigate all these scenarios, we cannot define a precise search query or perform a quantitative analysis on research publications only. Nor can we limit ourselves only to analysing and criticizing the use of data visualization techniques on existing collections and articles, or applied in the humanities field, because the data visualization collections are limited (or non-existent), and the field could benefit from approaches taking place in other disciplines, within and outside academia. Therefore, in Chapter 1, I use a hybrid methodology that allows me to cover a wider range of visualizations, in a less systematic way, but which opens new possibilities to understand the field across disciplines inside and outside academia. Based on a collection of 400 non-statistical data visualizations, I explore the prevailing techniques, the limitations of the current tools to produce them, and the potential to fill those gaps and create more inclusive data visualization tools, stretching along the entire epistemological axis.

The Minett region: environmental history and exposomics research

LuxTIME is inspired in the growing field of environmental history. According to Testot and Throssell, environmental history can take three forms: one that aims to historicize it by bringing nature into history, one that looks at the impact of humanity on the environment; and finally, one that studies the impact of the environment on humanity (e.g., health, trajectories of societies) (Testot & Throssell, 2020). The Anthropocene, the epoch of humans, has become a form of conceiving this age in which humanity has a major geological power (Bonneuil & Fressoz, 2016), characterized by the fact that “the human imprint on the global environment has now become so large and active that it rivals some of the great forces of Nature in its impact on the functioning of the Earth system” (Steffen et al., 2011, p. 1). Nevertheless, the Anthropocene also represents a moment in the development of knowledge about the environment, “found in measurements taken in the oceans, in the atmosphere, of changes in biota, of radiation and the detritus of consumption and construction”, a change only discoverable via the aggregated expertise (Warde et al., 2018, p. 181). When looking at environmental history, especially at the developments in the Anthropocene, interconnections of environment and health can be observed, where, as described in Chapter 2, expensive technological studies using current biological samples often fail to find them. In 2005, Christopher Wild described the complementary concept to the genome as “the exposome” including “life-course environmental exposures (including lifestyle factors) from the prenatal period onwards” (Wild, 2005, p. 1848). A definition later extended by Miller and Jones that expanded the concept adding the aspects of diet, behaviour and endogenous processes, with a particular focus on the biological responses to those exposures (Miller & Jones, 2014). In

Figure 8, we can see a list of factors that are part of the exposome as described by Vermeulen et al. (Vermeulen et al., 2020).

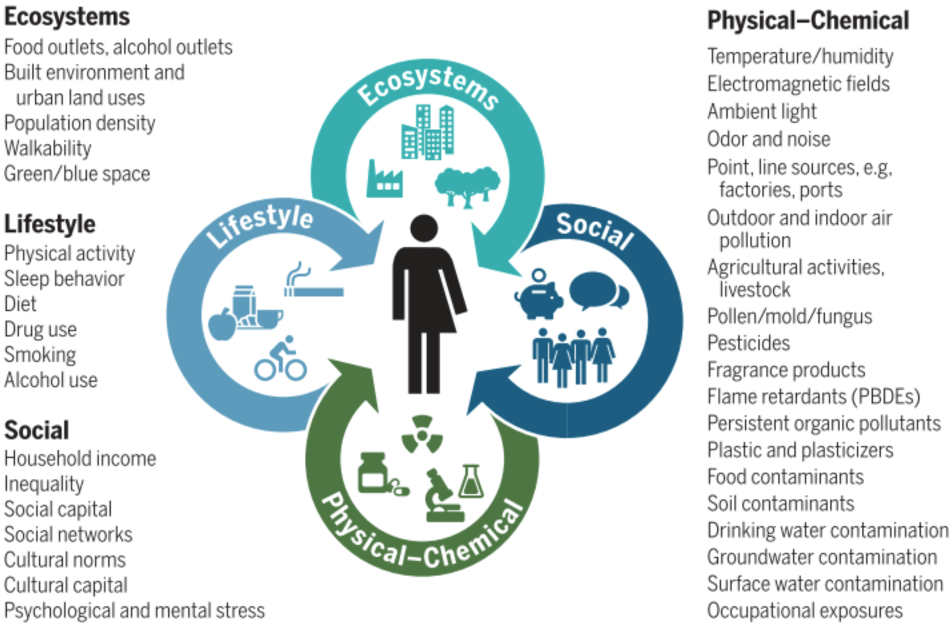


Figure 8 The Exposome and Health (Vermeulen et al., 2020)

Chapter 2 introduces historical exposomics and reflects on the value of using an interdisciplinary approach to study the exposome. It focuses on the contribution to the field of exposome research, based on the concept of “historical exposomics”, where it is argued that adding a historical view can offer interesting perspectives for both the natural sciences and the humanities. The study of historical sources so far unexplored by the sciences, might contribute to reconstruct the human exposome, covering all the external influences and internal (biological) responses of a human being from conception onwards (Canali, 2020). Through the concept of historical exposomics, we interrogate the trading zone between the exposome, history and data science. In this chapter, we critique the current approach to the exposome research, based on increasingly expensive resources and sophisticated techniques driven by metabolomics approaches, and the use of databases that mostly contain “known” knowledge, while new synthetic compounds are quickly appearing and might influence the human metabolism differently. Based on the definitions proposed by Pfister and Wanner , we describe multiple examples of natural (i.e., physical objects collected, processed, and deposited in the environment under natural circumstances without the interference of human species) and social archives (i.e., archives that are the result of a conscious and direct collection of past evidence for culture, political or economic purposes), arguing for enlarging the scope of current research using historical evidence from social archives. As previously discussed, interdisciplinary research aims at producing “significant” contributions for all the fields involved. This “balancing act”, this interrogation of the “trading zone” in the context of historical exposomics, implies discussing the practices and evaluation methods to assess significance in each discipline, as well as the facilitating role of data visualization during the exploration of the data, and the knowledge exchange. This chapter is the outcome of more than two years of joint writing workshops with all the members of the LuxTIME team. It is the result of the engagement in a collective learning process about the exposome.

Our case study is the industrial history of the Minett region. "Minett" is the name given to the south of Luxembourg, whose identity is the red soil, interspersed with iron ore, "Minett" (Thierry Kruchten et al., 2021). While the financial sector has played a key role in Luxembourg’s prosperity since the 1990s, the country’s wealth is rooted in its industrial past (Denis Scuto, 2019). The exploitation of oolitic ores to produce steel was a key episode in Luxembourg's industrial conversion in the 19th century. From the 1840s, ironmasters gradually abandoned the alluvial ores in the centre of the country in favour of the oolitic ores in the south-west. Luxembourg's iron ore deposits are part of the large Lorraine mining basin, the largest in Europe. The iron content of Luxembourg oolitic ore is relatively low (hence the diminutive “Minett”), at around 30% (compared to 35-55% for alluvial ore), but it is the considerable quantity that compensates for its low quality (Knebeler & Scuto, 2010), and the geological advantage that the deposits were close to the surface. Other factors that transformed the production conditions include Luxembourg’s entry in the “Zollverein” i.e., the German customs union, the construction of railroads, and the technological reorientation in the iron and steel industry (Leboutte et al., 1998). The development of the region was shaped by recessions and strong growth phases, including the industrial expansion crisis and the boom in the 1890s, the poor supply and the economic uncertainty of the interwar years, 30 years of growth after World War II, and finally the economic crisis and deindustrialization from the 1980s onwards. The region's agricultural landscape became an industrial one shaped by mining and smelting activities, and with the industrialization the environmental pollution .

In Chapter 3, we aim at creating a long-term picture of past dust pollution levels around the Belval plant. The Belval plant, built from 1909 to 1911, just before the First World War and during a phase of continuous expansion, became the flagship of the steel industry in Luxembourg, due to its size and modernism (Goedert et al., 2023). We use simulated data of airborne dust concentrations in the vicinity of the plant, between 1911 and 1997, when the last blast furnace was shut down. The data is simulated using a Gaussian dispersion model with input parameters: the production volumes, the processes (e.g., Thomas), the typical usage of filtering systems, and the prevailing wind directions over time. Based on a selection of historical sources, current knowledge in industrial engineering for the selection of the model and input parameters, the interpretation of a researcher in environmental cheminformatics to understand the impact on health, and the use of data visualization; we discuss an estimate of the amount of dust breathed, and how far it reached at different times. This transdisciplinary research made it possible to generate new, non-existent (historical) data and to isolate a selected source of pollution, which would not be possible even by observing measurements if they existed.

Throughout the process of discovering and learning about the historical context in this research, the existing literature on the history of Luxembourg, and more specifically on the history of the Minett region has provided us key indications about the areas of the exposome (see Figure 8) that could be further explored from a historical perspective. We can highlight the collection edited by Charles Barthel and Josée Kirps, Terres Rouges: History of the Luxembourg Steel Industry (Barthel et al., 2009), that has helped us better understand the economic, political, social and cultural heritage and therefore, the social exposure of the people living in the region. Terres Rouges by Marcel Schroder introduced us to the “not always harmonious” (Marcel SCHROEDER, 1961, p. 10) growth of the cities and the urban development around the industrial activities, as well as the cultural norms of its inhabitants, helping us to get a better understanding of their ecosystem, lifestyle and social exposure. Several other publications enrich the history of urban planning, especially about the city of Esch-sur-Alzette, “the iron metropole”, across different periods (Jean Goedert et al., 1991; Scuto, 1993; Victor Eischen &

Roger Fournelle, 1986). 100 Joer Esch (Biltgen, 2005) discusses social history, traditions, childhood, and living conditions, giving us an overview of social, lifestyle and ecosystems. In the book Luxembourg's workers' colonies and cheap housing (Lorang, 1994), Antoinette Lorang describes the living conditions between 1840 and 1940. The ARBED (Aciéries Réunies de Burbach-Eich-Dudelange), founded in 1911, initially one of several iron and steel making companies, gradually became the country's sole steel producer, and had great impact on the society and their living conditions. Gilbert Trausch discussed this influence in his book ARBED in the Luxembourg society (Trausch, 2000). Later publications such as the Environmental Atlas for Luxembourg (Dulli Frühauf & Rene Kollwelter, 1987) gave us some hints to explore physical-chemical exposures. With these references as a starting point, we have been able to explore other historical and scientific sources, to approach the different types of exposures of the population from different perspectives, and to help us understand the potential of historical exposomics research.

The data visualization toolbox

At the beginning of this research, my ambition was to create data repositories including the digitization of all sources found. These repositories would be part of a digital platform that would also integrate tools with all the discussed data visualization techniques ready to be used. This is still my ambition, but after three years of research, I have been able to understand where we are and how to carry it out, step by step, which goes far beyond the scope of this PhDs. Despite not having the repositories entirely digitized, and the visualization tools that allow us to experiment with different techniques across approaches, if there is one thing I have learnt, it is that before we get to that point, we need to extend the integration of data visualization as a tool for reflection and discussion through the entire project. If needed, in an analogous way. It is the questions that arise, the discussions about details and context, the need to think in other terms and from different perspectives, the discussions about established methods, that are of most value at this time. Data visualization is a very valuable tool to support that process. In Chapter 4, we reflect on the use of data visualization concepts and techniques through the project and propose a *data visualization toolbox* that could be implemented as a starting point in other interdisciplinary projects. In this chapter, we explore specific concepts and techniques (some of which were first introduced in Chapter 1), such as variations of statistical graphs, concept maps, the use of visual rhetoric, data humanism, or the use of multivariate glyphs, among others. We discuss their application in different cases of our project such as knowledge mapping, project evolution, data and metadata exploration or the analysis of the participants' experience. Data visualization is used not only a means to an end but also to embrace the idea of sandcastles (Hinrichs et al., 2019) using a speculative and process-oriented approach to advance knowledge.

Dissertation outline

In Chapter 1, I explore three data visualization approaches across different epistemological practices. The hypothesis is that a classification that is not limited to one discipline, use or context, might allow a better understanding of the data visualization field, the possibilities outside disciplinary boundaries, the current limitations, and how to bridge them. As a result of this research, I identify a series of recurrent techniques that could be integrated in data visualization tools to promote their use. I also conclude that the digital humanities is in a privileged position to bridge this gap, as they are accustomed to working interdisciplinary, using quantitative and qualitative methods, computational analysis, rhetorical analysis and interpretation, as well as aesthetic methods for the creation of knowledge through experience; and therefore, covering a wide range of epistemological methods.

In Chapter 2, we explore the use of social archives to study the exposome. The hypothesis is that the concept of *historical exposomics*, might offer a useful conceptual framework for studying the Anthropocene in an interdisciplinary way. As a result, we discuss the diversity of natural and social archives found and their mapping to the categories and subcategories of the exposome, based on which we discuss the concept of historical exposomics and its grounds for future exposomics research.

In Chapter 3, we reconstruct and visualize the airborne dust concentration generated by the Belval Steelworks between 1911 and 1997. The hypothesis is that history, and in this specific case study, Luxembourg's industrial history, can be reconstructed using historical primary sources containing quantitative and qualitative industrial and environmental data, which then can be analysed and visualized digitally. As a result, we create data (that does not exist otherwise) and that even though the concentrations are lower than the available estimations, it allows to isolate the contribution of a specific source of pollution, only based on historical primary sources. In addition, thanks to the data visualization, we can also observe the impact on the residents of Esch-sur-Alzette, as well as its cross-border reach.

In Chapter 4, we explore different ways to visualize objective and subjective phenomena in LuxTIME (e.g., knowledge exchange, interdisciplinary blending, data and metadata exploration, participants' contribution, experience). The hypothesis is that experimenting with data visualization concepts and techniques along the epistemological continuum, facilitates exchange and brings new perspectives in interdisciplinary collaborations. As a result, we propose a *data visualization toolbox*, that could be used, adapted, or enlarged in other interdisciplinary projects.

Chapter 1: Digital humanities at the intersection of three approaches to data visualisation: statistical graphics, data humanism, and humanistic interpretation

Contribution statement**:** As the only author of this article, I have conceptualized, reviewed, and edited the document.

Submission statement**:** This article has been submitted to the Benelux Journal of Digital Humanities 2023, as a long paper, following a presentation at the conference in June 2023.

Digital Humanities at the intersection of three approaches to data visualisation: statistical graphics, data humanism, and humanistic interpretation

Aida Horaniet Ibañez[1]

[1]Center for Contemporary and Digital History (C^2DH),
University of Luxembourg

Digital humanities has a critical role in the progress of the data visualisation field. First, humanities scholars engage with the concept of knowledge as interpretation. Second, they leverage computational tools and statistical methods for the analysis and visualisation of data and metadata. And finally, digital humanities projects are a space for experimentation where different epistemic cultures (which often go beyond the humanities and computer science) negotiate new forms of knowledge. Digital humanities are therefore in a privileged position to benefit from the full potential of data visualisation, using and improving existing methods and tools, and participating in the development of those that are much needed but do not yet exist. This article presents a classification of current approaches to data visualisation into (I) statistical graphics, (II) data humanism and (III) humanistic interpretation. These three approaches are based on four main aspects: a) the intellectual habits around the definition of data, b) the characteristics of the data visualisation and its objectives, c) the relation between the data and the data visualisation, and d) the expected user interaction. The classification is not intended to narrowly categorise any data visualisation, but rather to help navigate the visualisation continuum across epistemological practices. Moreover, based on a large collection of data visualisation examples compiled in an online gallery, several techniques are identified, that allow to make the transition between the different approaches, potentially facilitating interdisciplinary projects such as the LuxTIME Machine, that leverage multiple types of sources across languages and modalities.

Keywords: digital humanities; data visualisation; statistical graphics; data humanism; interpretation

1. Introduction

This research is developed in the context of the LuxTIME Machine (LuxTIME), an interdisciplinary project about the industrial history of Belval, Luxembourg. We use data visualisation to explore, understand and communicate about data from historical archives, scientific data from the fields of eco-hydrology and cheminformatics, and socio-cultural data. In LuxTIME, we hypothesize that data visualisation can facilitate interdisciplinary work, this research provides one study for understanding the needs of the different disciplines, the common grounds, and the options for visualising data integrating different approaches.

Leveraging data visualisation for research involving multimodal and interdisciplinary data requires understanding the needs and current practices along the epistemological continuum across quantitative, qualitative, rhetorical, and creative practices. In this research I aim at answering the following questions: Is there a data visualisation classification that could help us understand the needs and navigate the different solutions across epistemological practices? What differentiates the different approaches to data visualisation? Which specific data visualisation techniques are used to transition between those approaches? Could they be standardized and integrated into interdisciplinary projects and data visualisation tools? Which disciplines are best positioned to pursue these developments?

In section 3.1, I propose a classification into *statistical graphics*, *data humanism* and *humanistic interpretation*, based on a) the intellectual habits around the definition of data, b) the characteristics of the data visualisation and its objectives, c) the relation between the data and the data visualisation, and d) the expected user interaction. This classification is intended to help us understand the main needs and how they are reflected in the practice of data visualisation in the different fields. In section 3.2, with the aim of understanding which specific data visualisation techniques enable the transition from *statistical graphics* to *data humanism* and *humanistic interpretation*, I curate and analyse a collection of 400 data visualisations. Finally, in section 4, I discuss the current gaps between the approaches, the possibilities for future development to facilitate the use of a variety of data visualisation techniques across fields of research, and the role of the digital humanities.

2. Methods

This part of the research project, which lays the theoretical foundation for further practical experimentation[1], is conducted in three phases. First, I explored a collection of data visualisations from research articles, mainly published in journals in the related fields of the project: hydrology (e.g., Hydrology Sciences Journal, Advances in Water Resources), environmental cheminformatics (e.g., Exposome Journal), and digital humanities (e.g., Digital Humanities Quarterly, Journal of Digital History, Digital Scholarship in the Humanities). This initial exploration resulted in a predominance of visualisations using standard statistical graphs including bar charts, histograms, line plots, heatmaps, scatterplots, pie charts, contour plots, networks graphs and choropleth maps, among others. This is probably due, among other factors, to the fact that numerous data visualisation tools mostly support this type of standard graphs, accessible to all types of users with minimal technical and statistical knowledge. During this initial phase of the research, some data visualisations, later classified as *humanistic interpretation*, were also found. This initial collection gathered about 200 data visualisations[2] (see Figure 1), extracted from the different volumes of each journal starting with the most recent publications.

---

[1] In this article only the theoretical framework is discussed. The practical application to the project has been discussed in a separate article (Aida Horaniet Ibañez and Dagny Aurich 2023).

[2] This collection of data visualisations is collected in a Miro board. The separation into *statistical graphics* and *humanistic interpretation* was done later to integrate the latter in the second gallery and analysis.

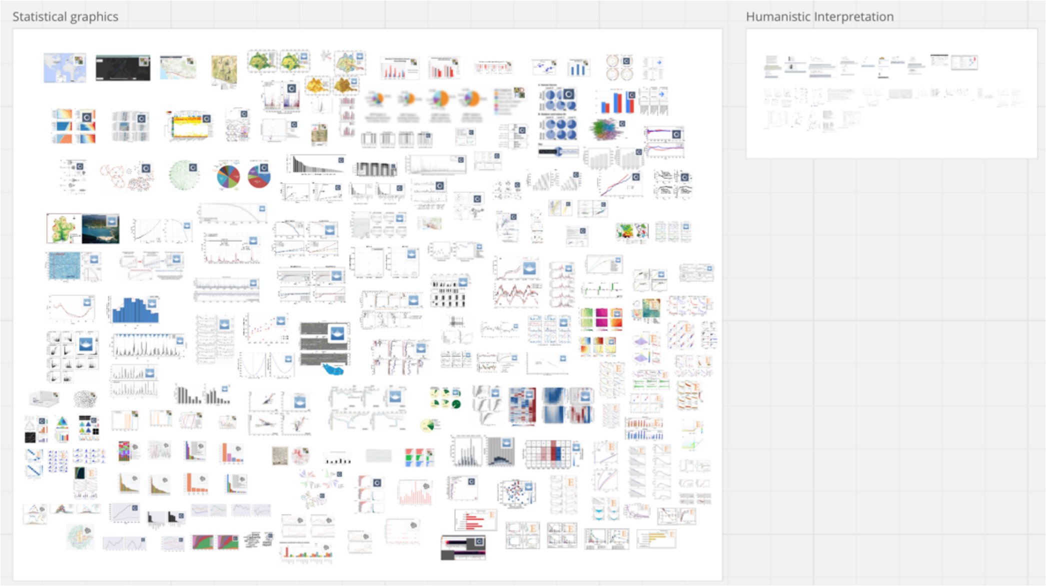


Figure 1 Initial Collection

Secondly, I gathered and explored a collection of about 400 data visualisations[3] (see Figure 2) that excluded those using strictly standard charts. As it will be discussed in section 4.2, the number of visualisations currently published in the online gallery is 400 because many links are no longer valid, but the number has been higher throughout the project. In addition to research articles, this second collection also includes non-research data sources such as online newspapers and magazines (e.g., Nightingale, National Geographic, South China Morning Post, Scientific American), social media posts (e.g., Twitter/X, Instagram), information designers' portfolios, data visualisation tool galleries (e.g., Tableau Public Gallery), community projects (e.g., Viz for Social Good, #journaldataviz) and data visualisation awards (e.g., Information is Beautiful Awards). Due to the variety of sources, this collection has been compiled manually, with the sole criterion of excluding standard statistical graphs.

---

[3] This collection of 400 data visualisations is available on Github [https://aidahi-unilu.github.io/AlternativeDataViz], including author(s), title, description, source, and the applied techniques, as discussed in the results section. This collection may be enlarged in the future with submissions from other practitioners in the field, through a call for contributions.

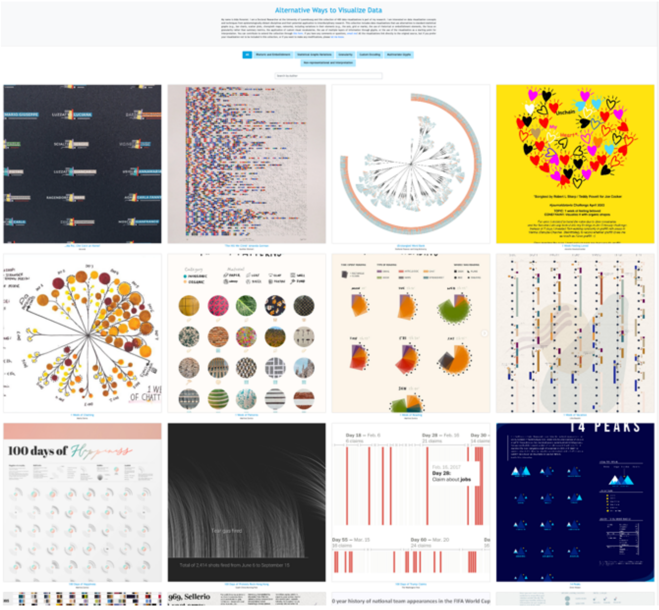


Figure 2 Non-standard charts gallery

Finally, based on these two collections, I propose a classification of approaches to data visualisation, and describe their main characteristics. Moreover, I identify a list of techniques used to move between these approaches. In the second collection, I have labelled the techniques used in the visualisations, and not in which approach they should be classified, because in some cases different combinations of the techniques allow to create visualisations belonging to different approaches. In other cases, the classifications are ambiguous, as the visualisations apply concepts from different approaches.

The classification is intended to help us to read the current landscape in general terms, while the techniques, give us the practical tools to move between the different approaches. These could be potentially used to explore epistemologically distant techniques in interdisciplinary setups and to be standardized and integrated in data visualisation tools.

3. Results

3.1. Classification Overview

[Table 1] presents a very concise summary of the classification that will be described in detail in the following sections. It includes the suggested name for the approach, the intellectual habits around the definition of data as *given* or *taken*, the main objectives of the data visualisation and its characteristics, the relation data to data visualisation, the expected user interaction, and the fields where it is most frequently used.

Table 1 Summary of Approaches to Data Visualisation

Approach	***Data***	***Objective***	***Characteristics***	***Data-Visualisation Relation***	***User Interaction***	***Main Fields***
#1: Statistical Graphics	*Given*	Complexity reduction, decision-making	Summary statistics, one clear message	Representational	Known charts, formatted answers	Business, Engineering, Natural Sciences
#2: Data Humanism	*Given (with emphasis on context)*	Emotional connection, engagement	Details, context, multiple narratives	Representational	Custom visual vocabularies, open exploration	Journalism, Education, Information Design
#3: Humanistic Interpretation	*Taken*	Model interpretation	Non-standard metrics	Non-representational	Close reading, marking, interpretation	History, Art, Literature

3.1.1. Approach names

To facilitate the subsequent analysis, each approach has been identified with a name that describes it as closely as possible: Approach #1: *Statistical graphics*, Approach #2: *Data humanism* and Approach #3: *Humanistic interpretation*.

Statistical graphics includes all the known forms of graphical representation based on statistical principles used to describe and summarize data, including plots such as scatter plots, histograms, box plots, choropleth maps, bar charts, line charts or heatmaps. The name chosen for the category is the most frequent name for this set of charts. I have included networks in this approach, because although they depict individual entities and relationships; they are standard charts (unless they include variations) based on mathematical structures, often used to gather statistics. In Figure 1, a collection of standard statistical graphs was presented.

Data humanism is based on the concept defined by (Lupi 2017) where she claims that “whenever the main purpose of data visualisation is to open people’s eyes to fresh knowledge, it is impractical to avoid a certain level of visual complexity”. This approach refers to a practice that embraces complexity in data visualisation, slowness, engagement looking at details – the grain, understanding the individual cases behind the numbers, avoiding standards when they do not fit the purpose, and adding context to understand the uniqueness of the representation. In Figure 3, we can see a collection of data visualisations that introduce some of these ideas.

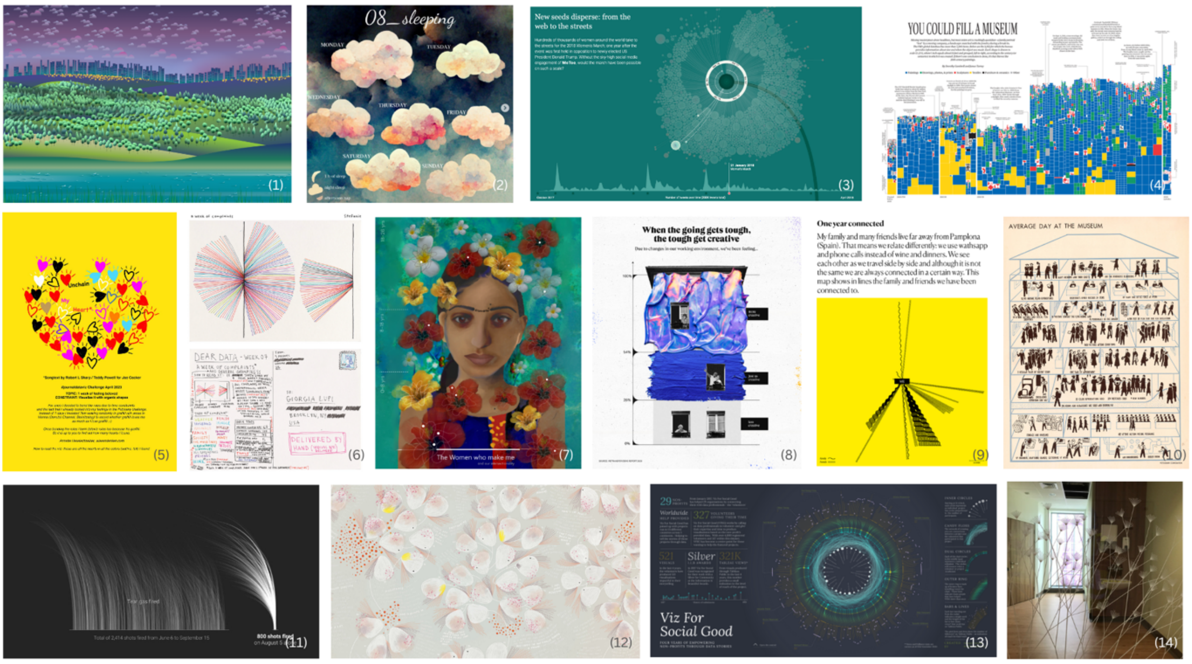


Figure 3 Data Humanism. (1)(Sonja Kujpers 2019), (2)(Alenka Gucek 2023), (3)(Valentina D'Efilippo and Lucia Kocincova 2017), (4)(Gambrell 2018), (5)(Annette Hexelschneider 2023), (6)(Lupi and Posavec 2016), (7)(Kadambari Komandur 2022), (8)(Gabrielle Merite 2020a), (9)(Alberto Molina Arce 2022), (10)(MoMA 1946), (11)(Pablo Robles, Darren Long, Dennis Wong 2019), (12)(Giorgia Lupi 2017), (13)(Samuel Parsons 2020), (14)(Stefanie Posavec 2019). All the visualisations are part of the collection described in Figure 2.

Humanistic interpretation refers to the approach that uses data visualisation to model interpretation and produce knowledge. A non-representational approach, in which data visualisation is not a surrogate of the dataset but the primary act of knowledge production (Drucker 2020). Besides the basic graphical features used in other approaches, it uses activators and inflectors and dimensions of interpretation, such as point of view or relative scales. This concept is based on the work of Drucker[4]. In Figure 4, we can see a less homogeneous collection of data visualisations representing humanistic interpretation. As we will discuss later, the examples are fewer, and thus we observe several attempts at interpretation through visualisation, from a mixture of proposed sketches from articles, physicalizations, and personal projects, among others.

---

[4] (Drucker and Bethany Nowviskie, n.d.; Drucker 2011; 2017; 2020)

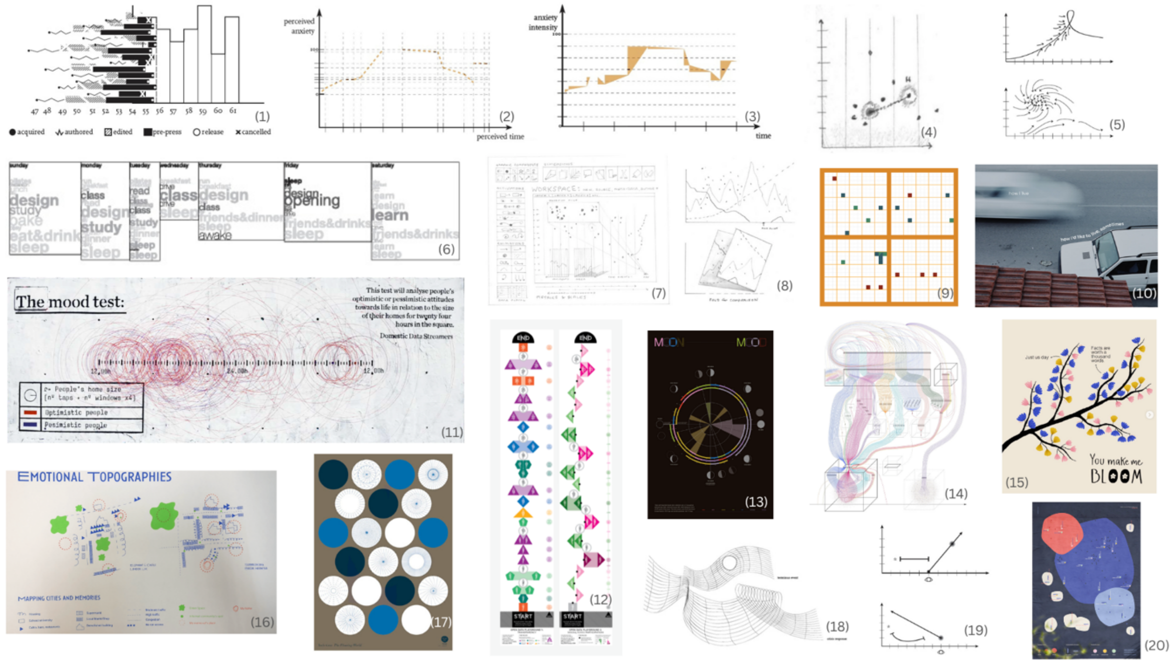


Figure 4 Humanistic Interpretation. (1, 2, 3, 5, 6, 18, 19)(Drucker 2011), (4, 7, 8)(Drucker 2017), (9)(Lauren Klein et al. 2016), (10)(Alessia Musio 2022), (11)(Domestic Data Streamers 2013), (12)(Stefanie Posavec 2014), (13)(Sarah Stern 2022), (14)(Stefanie Posavec 2023), (15)(Martina Zunica 2023), (16)(Syennie Valeria 2022), (17)(Stefanie Posavec 2011), (20)(Camilla De Amicis 2022)

3.1.2. Data

The intellectual habits around the definition of data are the first element of differentiation among approaches: *given* or *taken*[5]. I refer to the *intellectual habits*, and not to the definition of data itself, because all data is *taken* (as opposed to *found* or *given*). However, in the three approaches presented here, we observe different levels of engagement with the data construction process, or how that data is *taken* or *produced*, before being *reused*.

In *statistical graphics*, the intellectual habit is to use data as the starting point, a pre-existing fact, used practically as if it was *given*. In science and business, a data-driven culture suggests to always "start with the data" (Patil and Mason 2015), what somehow implies that the data is "the truth", a value-neutral object. The visualisations cite the data source, and if that source is considered reliable in the context in which it is presented, that is sufficient. Information about uncertainty is collected statistically in the form of confidence intervals and mathematical errors. Data is observer independent.

Data visualisation, as we predominantly know it today is the result of two main transformations: first, the data collection that is the mapping of a phenomenon to data, and second, the data visualisation or generation of a visual display from that data. This process involves a series of decisions from the occurrence of the phenomenon to the data visualisation, and beyond to the understanding of it: what elements do we collect depending on the physical or emotional context, the purpose, the person's previous knowledge or the configuration of the machine; what is grouped together; where the data is stored; which elements we choose to visualise, using which representations, what level of granularity or aggregation we choose and why. Even before adding the complexity of more advanced statistical models, the process involves many decisions, and each combination of them leads to a different result, a different

---

[5] Based on Drucker's concept of data as capta being taken actively instead of given (Drucker 2011).

dataset, and a different data visualisation. “Rethinking histories of data requires not only better answers to existing questions, but also better questions” (Strasser and Edwards 2017). In Big Data is the Answer…But What is the Question? Strasser and Edwards suggest several questions such as “What counts as data?”, “How are objects related to data?”, or “Why do we keep data, and how do we decide which data to lose or forget?”.

In *data humanism*, the intellectual habit is still to use data mostly as if it was *given*. However, a greater effort is made to situate it and to understand the reductive process that occurs during its generation. This will be visible later in the exploration of the data visualisation that becomes co-dependent with the user. In this approach, citing the data source, regardless of how trustworthy it is, is not enough. It requires a problematization of the provenance of the data, an identification of the stakeholders, and other more human elements necessary to understand “the where, why, and how of data” (D’Ignazio 2017). It interrogates the context, limitations, and validity of the same. It could be argued that such problematization allows to classify *data humanism* in the intellectual habit of using data as *taken* and not *given*. Nevertheless, in many of the examples analyzed, we see that the discussion occurs mostly at the level of the visualisation, and to a lesser extent at the data level. It remains true that some visualisations classified within this approach could indeed fit into the category of *taken*.

In *humanistic interpretation,* the intellectual habit is to always approach data as *taken*. Loukissas states that “things become data within interpretative acts” (Loukissas 2019). Rosenberg emphasizes on the rhetorical meaning of data: “Data has no truth (…). It may be that the data collected has no relation to truth or reality whatsoever beyond the reality that data help us to construct” (Gitelman 2013). Drucker refers to data as *capta*, ‘taken’ actively, while *data* is assumed to be a ‘given’ able to be recorded and observed” (Drucker 2011). In this approach to data visualisation, special attention is paid to the process itself, to the sources and their cultures because “when data are assembled from heterogeneous sources, each with their own local conditions, multiple settings are juxtaposed, creating a clash between discordant originating data cultures (Loukissas 2019); where it is stored since “making data means bringing a subject into a pre-existing system, defined by durable conditions of data collection as well as storage, analysis, and dissemination” (Loukissas 2019) Manovich refer to datasets as not just any collection of information but “objects structured in ways that allow them to exist within a computational medium and be analyzable by particular methods” (Manovich 2020). It looks at the limitations of the data itself, what exactly it represents, and the definition of social identities created through the definition of categories (Suchman 1993).

While the intellectual habits around the definition of data in *statistical graphics* are clearly different from the other two approaches, the difference between *data humanism* and *humanistic interpretation* is subtle. Perhaps because a large part of the visualisations classified as *data humanism* come from statistical graphics practitioners who have started to experiment with new visual vocabularies and high levels of granularity to avoid the limitations of the graphics they know, but without the critique of the process more rooted in the humanities.

3.1.3. Objectives and characteristics

The second factor of differentiation between approaches concerns the objectives of the data visualisation and how they characterize the graphical representation.

Statistical graphics aim at simplifying complex topics, they follow a reductionist approach that allows to convey high amounts of information into digestible charts very quickly. The objective is to support effective decision making and to communicate a clear message to a given audience. To achieve this goal, the visualisations are characterized by their abstraction, reduction, standardization, representation, and legibility. Using predefined charts and standard practices facilitates the fast processing of the content. These charts are selected based on the number of dimensions and measures, the type of data to be represented (e.g., continuous, discrete, categorical) and the data relationships that need to be displayed (e.g., deviation, ranking, flow, part-to-whole). The extensive research about graphic semiology and use of visual elements (e.g., color, shapes, size), the mechanics of the sight (e.g., Gestalt principles) and user experience, is applied to optimize the process, that is to ensure that the user receives the message as effectively as possible. It adheres to the concept of data-ink ratio introduced by Tufte (Tufte 1999), according to which any visual elements that interfere with the attention to the data are clutter, and therefore they should be avoided.

Data humanism focuses on promoting human values to foster emotional connection with the user. Data visualisations show complex topics in depth. The objective is to engage the user in the interpretation of the visualisation, introducing multiple narratives when necessary. Granularity, specificity, full coverage, high realism, physicality and situatedness are the main characteristics of this approach. Special importance is given to showing the gaps in the data and the complexity of the visualisation process itself. Unique visual vocabularies adapted to the visualisation of a particular theme are introduced, using organic shapes and original elements. These are defined from any combination of visual elements such as color, shape, angle, or size; and they are often explained in a "how to read" section. The use of *non-essential* or embellishment visuals is common. Communication and storytelling theory is applied to facilitate the navigation through the visualisation of the multiple narratives (e.g., audience analysis, story structure). It embraces pluralism and includes multiple perspectives to increase reflexivity (D'Ignazio and Klein 2020).

In *humanistic interpretation*, the goal is the interpretative process itself. A co-dependent constructivism allows to directly construct arguments through visual means. The visualisations are characterized by the use of non-standard visual elements such as non-discrete categories, unequal scale divisions, metrics as a factor of a point of view, of assumptions, presumptions or convention; non-continuous, non-homogeneous and multidirectional temporalities (Drucker 2017). [Figure 5] revisits Snow's cholera chart presented above, where Drucker suggests to look at the individual profiles of the people behind the dots (what would be considered as data humanism), but she goes further and suggests to "take the rate deaths, their frequency, and chart them on a temporal axis inflected by increasing panic". Redrawing the urban streetscape to express the emotional landscape is *humanistic interpretation*.

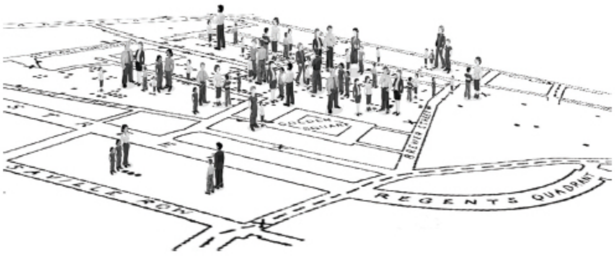


Figure 5 Snow's chart revisited to show the individuals behind the dots, emphasizing that "each dot represents a life, and none of them are identical" (Drucker 2011).

3.1.4. Relation between data and visualisation

The third differentiator explores the relation between the data and the visualisation. In *statistical graphics* and *data humanism*, data is the starting point and from the data we generate the data visualisation. They are both based on a representational approach in which "the relation between data and display is uni-directional, the data precede the display, and the data are presumed to have some reliable representational relation to the phenomena from which they have been abstracted" (Drucker 2017). *Humanistic interpretation* focuses on a non-representational approach, that uses the "graphical input as primary means of interpretative work" (Drucker 2017). The interaction with the data visualisation changes the data or create new versions of it.

3.1.5. User Interaction

Last, in *statistical graphics*, the user is expected to know the charts used in the visualisation. It requires a certain data literacy, including knowledge about statistical graphs (e.g., selecting the right charts or knowing how to read them). A user who knows how to read a type of graph can quickly read a graph related to any data set, ask predefined questions (e.g., in the case of a histogram: where most values fall and how much the variation is among values), and generate formatted answers. A critical approach is expected in terms of assessing whether the rules of statistical graphs are respected such as using the wrong graph, manipulating the axes, partially selecting data, or going against conventions to mislead the user. The user can interact with the visualisation by using filters or drilling up and down to different levels of aggregation (e.g., by date – monthly, yearly; by category). The visualisation is designed to effectively answer the a set of predefined questions.

In *data humanism*, the user does not know beforehand what to expect from the visualisation. The design and the topic stimulate the user's interest in exploring it further. No standard graphics are used, and if they are, they have been significantly redesigned in some way. Based on the user's own understanding of the data representation, guided by the frequently used 'how to read' section, the visualisation allows to ask open questions and explore the answers. The user is expected to take ample time to explore it, as effectiveness is not an important factor. We do not know if special skills or some level of data literacy is required to read the visualisations, as they are not standardized and are usually accompanied by instructions. Lupi, in her manifesto, *Data Humanism, The Revolution will be visualized* states that "the first wave [referring to *cheap marketing infographics*] was successful in making others more familiar with new terms and visual languages", implying that some kind of learning process is required to read these data visualisations (Lupi 2017). We "might want to move beyond literacy with datasets and towards literacies with infrastructure” since “ datasets do not simply neutrally designate aspects of the world, they also render the world in accordance with different visions, values and cultures, making it navigable through data” (Gray, Gerlitz, and Bounegru 2018). To gain an in-depth understanding of the implications, data literacy research needs to be extended to *data humanism* and *humanistic interpretation*.

In *humanistic interpretation*, the user interacts with the visualisation to build knowledge, the process happens through a series of interactions in which the user can mark and annotate the data. A set of tools allow the interaction with the visualisation, such as activators to introduce attraction and repulsion, weight, force, or sequence; or the use of graphical elements of interpretation such as point of view, layers, slices, tilt, fold, parallax or split (Drucker 2017). The user of this type of visualisations usually has a humanistic background, accustomed to the principles of ambiguity, complexity, and interpretation.

3.2. Data visualisation techniques: from *statistical graphics* to *data humanism* and *humanistic interpretation*.

Based on the second collection described in section 2. I identified several techniques to move between *statistical graphics* and *data humanism,* and *humanistic interpretation*. They range from subtle variations of standard charts (e.g., eliminating the grid or the tick marks in a bar chart), decomposing data showing the granularity behind the aggregated values, using custom visual vocabularies such as organic shapes, overlaying multiple layers of information using data glyphs, to completely challenging the principles of standard charts such as homogeneously spaced time units. All of them allow us to walk that continuum across epistemological practices. In the following sections, they are discussed separately in detail. Furthermore, these techniques are not mutually exclusive, but can be applied simultaneously, as can be seen in the online gallery through the multifilter selection.

3.2.1. Variations of statistical graphs

When using standard statistical charts, a series of rules are applied to convey the information efficiently. These rules include, among others, the selection of a graph adapted to the data and the analytical purpose, the right level of aggregation, the precise display of axes, labels, gridlines, legends, or annotations; or a low data-ink ratio avoiding distracting elements. In this analysis, I consider *variations of statistical graphs*, all those data visualisations including graphs where the standard graph is still recognizable, and knowing the rules that apply to such a graph, one or more rules are not applied for a specific purpose.

The options are multiple, but from the collection studied, it is worth noting the total elimination of grids, axes, or tick marks, to draw attention to the changes and not the specific values (Kim Albrecht 2016; Lopez, Williams, and Berne 2019); the turning, superposition and use of curved axes (Pentagram 2021a); and the over-encoding, using multiple visual attributes to encode the same information (Lazzaroni 2020). It is not considered a variation of a statistical graph when the graphs are misleading (e.g., extending the y-axis to minimize a variance), or when for no specific purpose the rules of standard charts do not apply (e.g., the legend is missing and there is no way to understand the content). These visualisations must be based on decisions made deliberately by the designer, to move between different approaches. Although the techniques presented below could be generally described in some cases as other types of variations, in the online gallery, I have only identified in this category those where the reference graph is fully recognizable.

3.2.2. Custom encoding

The use of custom visual vocabularies implies leaving aside the standard visual attributes associated with statistical graphs such as lines, rectangles and circles and creating an alternative way of encoding data variables, from replacing some of the elements of a standard statistical graph to completely redesigning all the elements of the visualisation.

Frequently observed encodings include organic forms (Martina Zunica 2023; Kadambari Komandur 2022), symbols (Kuijpers 2018) and drawings (Clark 2022), but also new combinations of geometric forms (Nwosu 2022). Among others, the use of custom visual vocabularies allows us to disaggregate information (by not having the constraint of standard graphs based on aggregate metrics), to display simultaneously multiple layers of information by integrating multiple variables in the visual vocabulary (see multivariate data glyphs below), to use rhetoric to attract the reader's attention (see mapping rhetoric below), and to potentially remove barriers to data literacy, since the reader does not need prior knowledge in statistics to read the visualisation[6]. This type of visualisation requires a "how to read" section, so that learning to read it is part of the process.

3.2.3. Rhetoric and embellishment

Contrary to data visualisations with a low data ink ratio, where "extra" elements are avoided, there are also data visualisations where these "embellishing" elements are used with a rhetorical purpose, to inform, persuade or motivate a given audience in a particular situation. These elements can be external to the graphics and therefore not linked to the data, in the form of images or drawings. Furthermore, rhetoric and embellishment elements can be part of the graph, not directly to encode the data but for example by using colour to create a specific effect (Scientific American 2015; Gabrielle Merite 2020b), customizing the layout (Valentina D'Efilippo and Lucia Kocincova 2017), rotating and superimposing graphs or through annotations (Bravo 2020). Last, rhetoric can also be applied directly in connection to the data, often in combination with the use of custom encodings through rhetorical mapping, where the visual attributes used to encode the data, apply visual metaphors to present the data in terms of the topic, the context, or the implications (Kate Wong, Liz Wahid, and Jan Willem Tulp 2021; Ervin Vinzon 2021).

3.2.4. Granularity

Granularity is defined in this research as the intent to disaggregate summary data typically presented in standard statistical graphs. The "grain" in a dataset is the smallest unit collected. In a data visualisation focused on granularity, importance is given to individual and unique cases behind the aggregated numbers. There might be different levels of aggregation/disaggregation, and they might be combined in the same data visualisation (drill-up/down approach). For example, in the data visualisation about the history of national team appearences in a world cup (Sonja Kujpers 2018), the grain would be the data related to each national team's appearance (e.g., year, result), the semi-aggregated data would present summary metrics by country (e.g., number of appearances, total number of points), and the aggregated data would describe the worldwide data (e.g., total number of tournaments, average

---

[6] From The Historical Archive To The Citizens: Visualizing Census Data From Brill Street in 1922 (Aida Horaniet Ibañez, Daniel Richter, and Joëlla Van Donkersgoed 2023) presents an example of the use of custom encodings to make historical data accessible to citizens.

number of points, total number of regions). These levels of aggregation are relative to the objectives of the visualisation, since the data can be broken down from many different perspectives. The granularity presented in a data visualisation gives us an important clue as to where it is situated in the epistemological practices. The closer to the "grain" the closer to the humanities, the rhetorical analysis and the interpretation; the more aggregated, the closer to the natural and applied sciences, standardization and generalization.

Based on the collection studied, one of the most commonly used techniques is displaying individual points, icons or drawings to humanize data, such as arrests during a protest (South China Morning Post 2020), Covid-19 deaths (Pentagram 2020), personal experiences (Lupi and Posavec 2016) or cells in a Jupyter notebook to create a unique article fingerprint (C2DH, University of Luxembourg 2021). Another approach consists of replacing (or overlaying) traditional shapes, such as the bars in a bar chart, with the data points that constitute them (Gambrell 2018), showing both the aggregated and the disagregated view at the same time; or using direct visualisation (Manovich 2011), preserving to some extent the original form of the grain (Flavio Gortana et al. 2018; Ferry 2020).

3.2.5. Multivariate data glyphs

A way to display multiple layers of information simultaneously, while maintaining a certain degree of granularity, is the use of multivariate data glyphs, a visual representation of data where the attributes of the graphical entity are defined by the attributes of the data record. To form the data glyphs, the encoding might rely on standard graphs and different levels of variations (Silvia Romanelli 2022), customized visual vocabularies (Kimly Scott 2023), or a combination of both. Every data glyph defines the lowest level of detail and they usually appear as small multiples, where each representation is derived from the same design structure, to leverage visual consistency, economy of perception and uninterrupted visual reasoning (Chuah and Eick 1998).

3.2.6. Non-representational and interpretation

In contrast to representational approaches, in a non-representational approach, the existence of data or other representations is not assumed prior to the interpretative work. Examples include data visualisations that use visual attributes to collect data not based on measurements or observations, such as personal experience or perception (Pentagram 2021b; Domestic Data Streamers 2013; Panagiotidou, Gorucu, and Vande Moere 2020). Another example is the use of visual vocabularies for interpretation, such as the uncertainty about a time duration (Nadieh Bremer 2016) or more complex concepts such as point of view, layers, scales or folds (Drucker 2017). Finally, it includes data visualisations where not only the graph elements can be aesthetically modified, but also the underlying principles, using for example ambiguous categories, multidirectional or perceived timeliness (Drucker 2011). Although these elements could be evaluated separately, I have decided to group them together because there are very few visualisations that implement them.

4. Discussion

4.1. Approaches and gaps

The classification proposed in this article helps us to understand the current landscape in data visualisation beyond disciplinary boundaries. The collection of data visualisations belonging to *data humanism* and *humanistic interpretation* allows us to identify, through numerous examples, a series of techniques linking these three classifications along the continuum of epistemological practices.

Data humanism and *humanistic interpretation,* both present options for shifting beyond *statistical graphics*. However, it is also interesting to discuss the gaps between the two as possibilities for future research. First, both promote the use of new visual vocabularies, however the "interpretative" visual language described in *humanistic interpretation* is perhaps more along the lines of a new standardization than the freedom to use any visual language, i.e., a specific symbol would mean attraction, as opposed to the designer deciding which representation of attraction is the most appropriate to a specific case, as it would be the case in *data humanism*. Second, both use variations of statistical graphs, in *data humanism* the aesthetics of the graph elements are modified, while in *humanistic interpretation*, the basic principles of their construction are challenged. Even so, aesthetically, the latter seem closer to the statistical graphs. In the literature, *data humanism* and *humanistic interpretation* are closer to each other than to *statistical graphics* (e.g., relevance of context to situate the data, transparency of the design process). However, based on the examples analyzed in the gallery, visually, it is *data humanism* that is more distant. In addition, we find numerous examples that symbolize the gradual transition between *statistical graphics* and *data humanism*, as well as between *statistical graphics* and *humanistic interpretation* (in proportion to the lower number of examples available in this category). We do not find so many examples situated between *data humanism* and *humanistic interpretation*, though. This leads to the question of what case studies would require such a combination, and how to bridge this gap between the two approaches, which are not the same.

4.2.Future applications and development

The techniques discussed above, among others, have been applied in the continuation of the LuxTIME project to facilitate interdisciplinary work (Aida Horaniet Ibañez and Dagny Aurich 2023). In addition, the identified techniques could be used as a reference to standardize certain functionalities in the existing data visualisation tools, since as we can observe, most of them have been created either with design programs (e.g. Adobe Illustrator), sacrificing the connection with the data and the interactivity; or with programming libraries such as (d3.js), limiting the accessibility to users without development knowledge. With some basic statistical and digital knowledge, any user can create standard statistical charts using tools like Excel or Tableau. However, moving away from *statistical graphics* requires specialized tools. This implies that the natural and applied sciences do not experiment with other approaches, and that the humanities, which need these techniques for interpretative work, are limited to quantitative analysis and standard visualisations. Without any doubt, some of the elements discussed above, especially regarding rhetorical elements and customized visual vocabularies, have a degree of variability that makes their development challenging, but offering some options would already be a step forward.

In this analysis, I have decided to group together certain techniques that could be analyzed extensively on their own. However, the lack of a gallery of data visualisations that would allow us to do so, and in cases, such as in *humanistic interpretation*, the lack of applied examples, make the disaggregation difficult at this stage. This gallery has been created with the objective of having a reference point in the analysis of data visualisation across disciplines and beyond the standard statistical graphs. There are other galleries (Manuel Lima 2014; Yan Holtz 2022; Maarten Lambrechts 2023), but to the best of my knowledge, none focused on the analysis of the approaches described in this research. Future research can leverage this collection and extend the metadadata (e.g., time period, sources, interactivity) to perform further analyses, or the identified techniques used could be studied independently in more detail. The collection could be enlarged collaboratively, or automatically thorugh the use of image similarity algorithms. It is worth noting that from the beginning of the project, when I started collecting data visualisations about two and a half years ago, almost one fifth of the links have been lost. Many of these projects come from social networks, personal blogs, company websites, newspapers and magazines, and other media with very variable lifespans. Consequently, it would be necessary to archive all the visualisation websites in a lasting way (e.g., using web archives), that allow us to maintain and do further research taking into account their specific context, and not only the images.

4.3. The role of the digital humanities

In recent years, numerous research projects in the field of digital humanities have focused on other ways to visualise data (Lauren Klein et al. 2016; Viktoria Brüggemann, Mark-Jan Bludau, and Marian Dörk 2020; Hinrichs, Forlini, and Moynihan 2019). However, in research articles, as discussed above, the use of standard *statistical graphics* presenting summary metrics, and network diagrams to depict relationships, continues to predominate. It is possible that by gradually integrating techniques from different disciplines that allow us to walk along this line between objectivity and subjectivity, first in a "manual" way, and then standardizing them in the tools, the possibilities available to all the disciplines will increase. The development of these tools requires researchers accustomed to working in an interdisciplinary way encompassing quantitative and qualitative methods, computational analysis and development, but also familiar with rhetorical analysis and interpretation, as well as the use of aesthetic methods and the creation of knowledge through experience. All these factors are present in the digital humanities, which places them in a strong position to advance the field of data visualisation.

References

Aida Horaniet Ibañez and Dagny Aurich. 2023. ‘How Can Data Visualization Support Interdisciplinary Research? LuxTIME: Studying Historical Exposomics in Belval [Submitted for Publication]’. *Frontiers*.

Aida Horaniet Ibañez, Daniel Richter, and Joëlla Van Donkersgoed. 2023. ‘From The Historical Archive To The Citizens: Visualizing Census Data From Brill Street in 1922’. 2023. https://nightingaledvs.com/from-the-historical-archive-to-the-citizens-visualizing-census-data-from-brill-street-in-1922/.

Alberto Molina Arce. 2022. ‘One Year Connected’. 2022. https://www.informationisbeautifulawards.com/showcase/5130-one-year-connected.

Alenka Gucek. 2023. ‘Sleeping’. 2023. https://www.instagram.com/p/CnKFvAvI-cJ/?img_index=1.

Alessia Musìo. 2022. ‘Slowness’. 2022. https://www.behance.net/gallery/148388993/Slowness.

Annette Hexelschneider. 2023. ‘1 Week Feeling Loved’. 2023. https://twitter.com/ainvienna/status/1651825405815291904.

Bravo, Liz. 2020. ‘Week 43: Apparel Exports to US’. *Favorites* (blog). 2020. https://www.makeovermonday.co.uk/week-43-2020/.

C2DH, University of Luxembourg. 2021. ‘Journal of Digital History - Article Fingerprint’. 2021. https://journalofdigitalhistory.org/en/issue/jdh001.

Camilla De Amicis. 2022. ‘5 Years’. 2022. https://www.informationisbeautifulawards.com/showcase/5808-5years.

Chuah, M.C., and S.G. Eick. 1998. ‘Information Rich Glyphs for Software Management Data’. *IEEE Computer Graphics and Applications* 18 (4): 24–29. https://doi.org/10.1109/38.689658.

Clark, Sarah Emery. 2022. ‘What’s Cookin’. *Winner of the 2022 Data Visualization Society ‘State of the Industry’ Award for Most Beautiful Design* (blog). 2022. https://www.sarahemeryclark.com/work/whats-cookin.

D’Ignazio, Catherine. 2017. ‘What Would Feminist Data Visualization Look Like?’ *Feminist Studies* 14 (3): 575. https://doi.org/10.2307/3178066.

D’Ignazio, Catherine, and Lauren F. Klein. 2020. *Data Feminism*. Strong Ideas Series. Cambridge, Massachusetts: The MIT Press.

Domestic Data Streamers. 2013. ‘The Mood Test’. 2013. https://domesticstreamers.com/projects/the-mood-test/.

Drucker, Johanna. 2011. ‘Humanities Approaches to Graphical Display’. *DHQ* 5 (1).

———. 2017. ‘Non-Representational Approaches to Modelling Interpretation in a Graphical Environment’. *Digital Scholarship in the Humanities* 33(2) (July). https://doi.org/10.1093/llc/fqx034.

———. 2020. *Visualization and Interpretation: Humanistic Approaches to Display*. Cambridge, Massachusetts: The MIT Press.

Drucker, Johanna, and Bethany Nowviskie. n.d. ‘Temporal Modelling: Conceptualization and Visualization of Temporal Relations for Humanities Scholarship’.

Ervin Vinzon. 2021. ‘14 Peaks’. 2021. https://public.tableau.com/app/profile/ervin.vinzon/viz/Eight-Thousanders_16386540239820/dash.

Ferry, Frederic. 2020. ‘1000 Runs’. 2020. https://public.tableau.com/app/profile/fredfery/viz/My1000Runs/1000Runs.

Flavio Gortana, Franziska von Tenspolde, Daniela Guhlman, and Marian Dörk. 2018. ‘Coins’. 2018. https://uclab.fh-potsdam.de/coins/.

Gabrielle Merite. 2020a. ‘Data Illustrations for WePresent’. 2020. https://www.behance.net/gallery/109802031/Data-Illustrations-WePresent-Ideas-Report-2020.

———. 2020b. ‘When the Going Gets Tough, the Tough Get Creative’. 2020. https://www.behance.net/gallery/109802031/Data-Illustrations-WePresent-Ideas-Report-2020.

Gambrell, Dorothy. 2018. ‘You Could Fill a Museum’. *Bloomberg Businessweek* (blog). 2018. https://www.informationisbeautifulawards.com/showcase/3266-you-could-fill-a-museum.

Giorgia Lupi. 2017. ‘Bruises. The Data We Don’t See.’ 2017. http://giorgialupi.com/bruises-the-data-we-dont-see.

Gitelman, Lisa, ed. 2013. *‘Raw Data’ Is an Oxymoron*. Infrastructures Series. Cambridge, Mass.: MIT Press.

Gray, Jonathan, Carolin Gerlitz, and Liliana Bounegru. 2018. ‘Data Infrastructure Literacy’. *Big Data & Society* 5 (2): 205395171878631. https://doi.org/10.1177/2053951718786316.

Hinrichs, Uta, Stefania Forlini, and Bridget Moynihan. 2019. ‘In Defense of Sandcastles: Research Thinking through Visualization in Digital Humanities’. *Digital Scholarship in the Humanities* 34 (Supplement_1): i80–99. https://doi.org/10.1093/llc/fqy051.

Inbal Rief. 2020. ‘Hair Me Out’. 2020. https://medium.com/nightingale/not-on-our-watch-17907b86b985.

Kadambari Komandur. 2022. ‘Data Selfi’. 2022. https://kadambarik.myportfolio.com/data-selfi.

Kate Wong, Liz Wahid, and Jan Willem Tulp. 2021. ‘What Scientists Have Learned from 100 Years of Bird Banding’. 2021. https://www.scientificamerican.com/article/what-scientists-have-learned-from-100-years-of-bird-banding/.

Kim Albrecht. 2016. ‘Quantifying the Evolution of Individual Scientic Impact’. 2016. http://sciencepaths.kimalbrecht.com/.

Kimly Scott. 2023. ‘#Moms Who Viz’. 2023. https://public.tableau.com/app/profile/kimly.scott/viz/MomsWhoVizWomensHistoryMonth/MomsWhoVizWomensHistoryMonth.

Kuijpers, Sonja. 2018. ‘Keuzestress. Searching for the “Correct” Mascara’. 2018. https://www.studioterp.nl/mascaras-a-datavisualization/.

Lauren Klein, Caroline Foster, Erica Pramer, Adam Hayward, and Shivani Negi. 2016. ‘The Shape of History’. 2016. http://shapeofhistory.net/.

Lazzaroni, Michaela. 2020. ‘Space Rocks’. *La Lettura #439 Corriere Della Sera* (blog). 2020. https://www.behance.net/gallery/96212805/Space-Rocks-La-Lettura-439-dataviz.

Lopez, Alberto Lucas, Ryan Williams, and Kaya Berne. 2019. ‘Migration Waves: This Is What 50 Years of Human Migration Looks Like’. *The National Geographic* (blog). 2019. https://www.nationalgeographic.com/magazine/graphics/graphic-shows-past-50-years-of-global-human-migration?sf215829698=1&sf217104276=1.

Loukissas, Yanni A. 2019. *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, Massachusetts: The MIT Press.

Lupi, Giorgia. 2017. ‘Data Humanism: The Revolutionary Future of Data Visualization’. *PrintMag*, 2017. https://www.printmag.com/article/data-humanism-future-of-data-visualization/.

Lupi, Giorgia, and Stefanie Posavec. 2016. *Dear Data*. First edition. New York: Princeton Architectural Press.

Maarten Lambrechts. 2023. ‘Xenographics’. 2023. https://xeno.graphics/.

Manovich, Lev. 2011. ‘What Is Visualisation?’ *Visual Studies* 26 (1): 36–49. https://doi.org/10.1080/1472586X.2011.548488.

———. 2020. *Cultural Analytics*. Cambridge, Massachusetts: The MIT Press.

Manuel Lima. 2014. ‘Visual Complexity’. 2014. http://www.visualcomplexity.com/vc/.

Martina Zunica. 2023. ‘You Make Me Bloom’. 2023. https://www.instagram.com/p/CrnsBoPt1Mw/?img_index=1.

MoMA. 1946. *Diagram, Average Day at the Museum: The Year’s Work: Annual Report to the Board of Trustees and Corporation Members of the Museum of Modern Art for the Year June 30, 1939–July 1, 1940*. MoMA Archives.

Nadieh Bremer. 2016. ‘My Life Vacations’. 2016. https://vacations.visualcinnamon.com/.

Nwosu, Chimdi. 2022. ‘Build Up Nepal for Viz for Social Good’. 2022. https://public.tableau.com/app/profile/chimdi.nwosu/viz/VizForSocialGood-BuildUpNepal/Dashboard1.

Pablo Robles, Adolfo Arranz and Dennis Wong. 2019. ‘How MMA’s Roots Are in Hong Kong’. 2019. https://multimedia.scmp.com/culture/article/SCMP-printed-graphics-memory/lonelyGraphics/201905A233.html.

Pablo Robles, Darren Long, Dennis Wong. 2019. ‘100 Days of Protests Rock Hong Kong’. 2019. https://multimedia.scmp.com/infographics/news/hong-kong/article/3027462/hong-kong-100-days-of-protests/index.html?src=arcade.

Panagiotidou, Georgia, Sinem Gorucu, and Andrew Vande Moere. 2020. ‘Data Badges: Making an Academic Profile Through a DIY Wearable Physicalization’. *IEEE*

Computer Graphics and Applications 40 (6): 51–60. https://doi.org/10.1109/MCG.2020.3025504.

Patil, DJ, and Hilary Mason. 2015. *Data Driven: Creating a Data Culture*.

Pentagram. 2020. ‘Covid-19 Charts’. 2021 2020. https://www.pentagram.com/work/covid-19-charts/story.

———. 2021a. ‘A Warming Planet’. *Bulleting of the Atomic Scientists* (blog). 2021. http://giorgialupi.com/bullettin-of-the-atomic-scientists.

———. 2021b. ‘Mindworks: The Science of Thinking’. 2021. https://www.pentagram.com/work/mindworks-the-science-of-thinking/story.

Samuel Parsons. 2020. ‘Viz for Social Good. Self Reflection 2017-2020’. 2020. https://public.tableau.com/app/profile/sparsonsdataviz/viz/VizForSocialGood-SelfReflection2017to2020/VizForSocialGood-SelfReflection2017to2020.

Sarah Stern. 2022. ‘Moon Mood’. 2022. https://www.informationisbeautifulawards.com/showcase/5525-moon-mood.

Scientific American. 2015. ‘Relativity’s Influence Is Still Going Strong on Its 100th Birthday’. 2015. https://www.scientificamerican.com/article/relativity-s-influence-is-still-going-strong-on-its-100th-birthday/.

Silvia Romanelli. 2022. ‘Endangered Languages’. 2022. https://www.behance.net/gallery/128273653/Endangered-Languages.

Sonja Kujpers. 2018. ‘20 Year History of National Team Appearances in the FIFA World Cup’. 2018. http://www.studioterp.nl/world-cup-national-team-appearances-datavisualization/.

———. 2019. ‘A View on Dispair’. 2019. https://www.studioterp.nl/a-view-on-despair-a-datavisualization-project-by-studio-terp/.

South China Morning Post. 2020. ‘Arrested Hong Kong Protesters’. 2020. https://multimedia.scmp.com/infographics/news/hong-kong/article/3088009/one-year-protest/index.html.

Stefanie Posavec. 2011. ‘Inch Time - The Floating World’. 2011. https://www.stefanieposavec.com/inch-time.

———. 2014. ‘Open Data Playground’. 2014. https://www.stefanieposavec.com/open-data-playground.

———. 2019. ‘Papworth Hospital, Inpatients Ward’. 2019. https://www.stefanieposavec.com/papworth.

———. 2023. ‘Data Murmurations: Points in Flight (People Like You) View Fullsize’. 2023. https://peoplelikeyou.ac.uk/system/.

Strasser, Bruno J., and Paul N. Edwards. 2017. ‘Big Data Is the Answer … But What Is the Question?’ *Osiris* 32 (1): 328–45. https://doi.org/10.1086/694223.

Suchman, Lucy. 1993. ‘Do Categories Have Politics? The Language/Action Perspective Reconsidered’. In *Proceedings of the Third European Conference on Computer-Supported Cooperative Work 13–17 September 1993, Milan, Italy ECSCW ’93*, edited by Giorgio de Michelis, Carla Simone, and Kjeld Schmidt, 1–14. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-011-2094-4_1.

Syennie Valeria. 2022. ‘Emotional Topographies. Cities.’ 2022. https://www.behance.net/gallery/140603619/Emotional-Topographies-Cities?tracking_source=search_projects%7Cinformation%20visualization%20data%20viz.

Tufte, Edward R. 1999. *The Visual Display of Quantitative Information*. 17. print. Cheshire, Conn: Graphics Press.

Valentina D’Efilippo and Lucia Kocincova. 2017. ‘Me Too’. 2017. http://metoomentum.com/about.html.

Viktoria Brüggemann, Mark-Jan Bludau, and Marian Dörk. 2020. ‘The Fold. Rethinking Interactivity in Data Visualization’. 2020. https://uclab.fh-potsdam.de/projects/the-fold/.

Yan Holtz. 2022. ‘Dataviz Inspiration’. 2022. https://www.dataviz-inspiration.com/.

Chapter 2: Historical exposomics: a manifesto

Contribution statement: All authors have contributed to the conceptualization, review and editing of the article. I have written the sections “Finding Significance in a Different Way” and “Data Visualization”, and I have created Figures 1, 2 and 4 based on the data collection performed by myself and Dagny Aurich. Dagny Aurich has written the Introduction, 'The LuxTIME Project as an Interdisciplinary Framework', 'Critique of the Current Approach', 'Botanical Samples', and parts of the 'Conclusion'. Laurent Pfister has writen the section on 'Freshwater Bivalves'. Simon Kreipl has writen the 'Historical Groundwater Reserves' section, and Emma L. Shymanski has written the abstract. Simon Kreipl, Dagny Aurich and Andreas Fickers have contributed to the sections on “Environmental history” and the “Anthropocene”. Andreas Fickers has written the section about “Definition of the Environment”, and the introduction to “Evidence from Social and Natural Archives”, “New historical narratives”, and contributed to the “Conclusion”. Andreas Fickers and Emma Schymanski were responsible for funding acquisition and supervision.

Submission statement*:* This article was published on August 18, 2023, in volume 3 of Exposome Journal.

Exposome, 2023, **3(1)**, osad007
https://doi.org/10.1093/exposome/osad007
Advance Access Publication Date: 18 August 2023
Commentary

Historical exposomics: a manifesto

Dagny Aurich [1,*,†], Aida Horaniet Ibanez [2,*,†], Christophe Hissler [3], Simon Kreipl [4], Laurent Pfister [3,5], Emma L. Schymanski [1], Andreas Fickers [2]

[1]Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Belvaux, Luxembourg
[2]Luxembourg Centre for Contemporary and Digital History (C2DH), University of Luxembourg, Esch-sur-Alzette, Luxembourg
[3]Catchment and Eco-hydrology Research Group (CAT), Luxembourg Institute of Science and Technology (LIST), Belvaux, Luxembourg
[4]Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands
[5]Faculty of Science, Technology and Medicine, University of Luxembourg, Belvaux, Luxembourg

*To whom correspondence should be addressed: Dagny Aurich: Email: dagny.aurich@uni.lu; Aida Horaniet Ibanez: Email: aida.horanietibanez@uni.lu

†These authors contributed equally to this work and share first authorship

Abstract

The exposome complements information captured in the genome by covering all external influences and internal (biological) responses of a human being from conception onwards. Such a paradigm goes beyond a single scientific discipline and instead requires a truly interdisciplinary approach. The concept of "historical exposomics" could help bridge the gap between "nature" and "nurture" using both natural and social archives to capture the influence of humans on earth (the Anthropocene) in an interdisciplinary manner. The LuxTIME project served as a test bed for an interdisciplinary exploration of the historical exposome, focusing on the Belval area located in the Minett region in southern Luxembourg. This area evolved from a source of mineral water to steel production through to the current campus for research and development. This article explores the various possibilities of natural and social archives that were considered in creating the historical exposome of Belval and reflects upon possibilities and limitations of the current approaches in assessing the exposome using purely a natural science approach. Issues surrounding significance, visualization, and availability of material suitable to form natural archives are discussed in a critical manner. The "Minett Stories" are presented as a way of creating new historical narratives to support exposome research. New research perspectives on the history of the Anthropocene were opened by investigating the causal relationships between factual evidence and narrative evidence stemming from historical sources. The concept of historical exposome presented here may thus offer a useful conceptual framework for studying the Anthropocene in a truly interdisciplinary fashion.

Keywords: exposome; environmental history; historical exposomics; digital history; social archives; natural archives

Introduction

The *nature versus nurture* debate is one of the oldest philosophical debates, dating back to ancient times (eg, Plato's Protagoras).[1] *Nature* is the genetic and biological "predetermination" of a human being, whereas *nurture* comprises all external factors influencing the individual from conception onwards. Today, this debate is—it seems—once more reframed as a scientific controversy opposing *genomics* to the concept of *exposomics*. With the highly funded Human Genome Project, next generation sequencing methods and genome-wide association studies (GWAS), genomics is both more mature and better financed than exposomics. However, interest in exposomics is increasing as many diseases cannot be traced back to genetics alone, but rather to the interplay of genetics and many other factors, which are captured under the exposome concept. The exposome covers all external influences and internal (biological) responses of a human being from conception onwards.[2] This paradigm[2] cannot be covered by just one scientific discipline; it requires a truly interdisciplinary approach. The "Luxemburg Time Machine" (LuxTIME) project is such an interdisciplinary setting—a "trading-zone" between different disciplines to find a common language between humanities and natural sciences. For this we propose the concept of "historical exposomics", aiming at interrogating the "trading zone"[3] between the exposome, history and data science to capture human influence on the natural world.

It is rare to find publications written from the perspective of both humanities and natural sciences due to the disciplinary specialization of each field. In general, both genome and exposome research is dominated by the natural sciences so far. Yet involving humanities in this research area, and more specifically adding a historical view, offers interesting perspectives for the exposome. Historical sources can help reconstruct the human exposome by revealing a multitude of historical data typically unexplored by the natural sciences. Exposomics evidence from the past is difficult to find in "natural archives" (the focus of natural sciences) present today. However, looking at "social archives"

Received: April 20, 2023. **Revised:** July 4, 2023. **Accepted:** July 21, 2023

opens new possibilities to detect historical evidence (see Evidence section with an explanation of natural and social archives). By combining "hard" and "soft" facts, a more complex understanding of past processes of environmental change is possible. This article explores this based on the example of the interdisciplinary LuxTIME project from the University of Luxembourg.[4]

The LuxTIME project as an interdisciplinary framework

The LuxTIME project aims at exploring new ways of analyzing and interpreting factual evidence of the past by building an interdisciplinary framework for the investigation of "big data" of the past. Building on the conceptual premises of the "European Time Machine" Flagship project,[5] that is to bring together High Performance Computing (HPC) facilities, the analytical capacities of Machine Learning and Artificial Intelligence, and "big" and complex historical data, LuxTIME includes information from three different fields and scientific perspectives, namely eco-hydrology, environmental chemistry and industrial history. LuxTIME uses a local case (the industrialization of Belval in the Minett region of Southern Luxembourg)[6] as a testbed for methodological and epistemological reflections on how to study the long term impact of environmental changes on the health of the local population. By mixing "contextual information" based on archival evidence with "scientific evidence" derived from chemical, biological, or medical investigations, the project explores new grounds for interpreting big data of the past in a truly interdisciplinary setting.

Environmental history and the anthropocene

LuxTIME is inspired by the rapidly growing field of environmental history, which aims at analyzing and describing the co-evolution of human society and the natural environment.[7] The interdependence of humans and other living and non-living systems on Earth is recognized, and the unfolding of history is studied in the framework of the natural world. Recently, the human influence on the development of the ecosystem is being discussed and debated under the label of the "Anthropocene"[8]—a new temporal regime and ecological period marked by the human imprint on the natural world, including the "Critical Zone", the Earth's skin which sustains nearly all terrestrial life including humanity.[9-12,13]

For most of the time since the origins of humankind, the population was too small, spatially scattered and technologically undeveloped to markedly influence the environment. With the Neolithic Revolution (~10 000 BCE) and the change from a hunting and gathering lifestyle to one of agriculture and settlement, the human species learnt new ways to modify the environment to serve their purposes. Deforestation, irrigation,[14] the development of systematic sowing methods and the use of fertilizers drastically increased food production, allowing for a surge in the human population size. In 10100 BCE the population rose above 5 million. With this came the development of trading networks and complex societies.[15] Ecosystems were brought out of their natural balance.[16] Soil erosion and salinization, the depletion of nutrients or a rise in infectious diseases were some of many consequences accompanying these technological advancements.[17,18] From the Roman Era to the Middle Age, new technological development of iron metallogeny and intense mining activities also enriched in heavy metals the local atmospheric deposition all over Europe.[19,20]

Since then, the scientific (~1540s) and industrial revolution (~1780s) have resulted in countless inventions, which have left their environmental traces on Earth. This process, ideologically framed as era of "modernization" and narrated as a history of "progress", has left a long-lasting toll on the planet.[21] It is estimated that humans have transformed 20 to 100% of all land surfaces.[22,23] The release of pesticides, fertilizers and other chemical contaminants is interfering with natural ecosystems, resulting in large-scale biodiversity loss and deterioration of soil and water quality.[24] Recent extinction rates are estimated to be 100 to 1000 times higher than the average rate on geological time scales.[25] Chlorofluorocarbons (CFCs) have spread to the upper layer of the atmosphere and depleted the ozone layer, which protects the Earth from ultraviolet radiation coming from the sun.[26] Microplastics have been detected in various environments, including the air, oceans and freshwater eco-systems.[27] The aftermath of anthropogenic activities affecting the natural environment is expected to persist for long time periods (>50 000 years).

The various developments arising from the wasteful and harmful treatment of the environment by the human species stimulated the rise of preservationist and conservationist philosophies[28,29] in the late 19th and early 20th century, acknowledging the boundaries of nature and advocating a wise and efficient use of resources where humans should see themselves as an integrative part of the natural environment, rather than the conquerors thereof. This has evolved to the global environmental movement of today, which includes a wide range of activist organizations, political parties, scientific organizations, and governmental policies focused on promoting environmental values, and combatting detrimental behavior, such as the emission of greenhouse gases, deforestation and pollution.[30]

The United Nations assumed a key role in debating climate change and global environmental issues in the political sphere. The 1972 United Nations Conference on the Environment in Stockholm was the first international conference connecting economic growth and human well-being to environmental aspects such as the pollution of air, water and the oceans.[31] The 1992 United Nations Conference on Environment and Development (UNCEP), held in Rio de Janeiro, composed a declaration of 27 principles aimed at guiding countries to a sustainable future, where "human beings are entitled to live healthy and productive lives in harmony with nature".[32] The Montreal Protocol of 1987 aimed at phasing out ozone depleting substances such as CFCs.[33] It is, to date, the only environmental treaty ratified by all 198 UN member states and is regarded as a benchmark in multilateral environmental regulation. The Kyoto Protocol of 1997 and the Paris Agreement of 2015 aim at reducing greenhouse gas emissions and limiting global warming to 2°C compared to pre-industrial levels.[34,35] In 2001 the Stockholm Convention on persistent organic pollutants (POPs) was signed, a global treaty to minimize the risk posed by POPs to the environment.[36] The initial list of 12 ("the dirty dozen") compounds prepared by the intergovernmental negotiating committee (INC) is extended regularly and to date covers 23 unique compounds (including the commercial mixture of decabromodiphenyl ether, c-decaBDE) and 8 compound classes (including compounds that are listed with their salts, isomers or esters) to be eliminated or restricted.

History and the exposome: a conceptual discussion

Critique of current approach

In his famous 2005 paper, the epidemiologist Christopher Wild was one of the first to give the complementary concept to the genome a name: the exposome. According to Wild, the exposome includes "life-course environmental exposures (including lifestyle factors) from the prenatal period onwards".[37] Gary Miller and Dean Jones extended the concept of the exposome in 2014, shifting the focus of the exposome beyond solely the exposures.[38] They expanded the concept adding the aspects of diet, behavior and endogenous processes, with a particular focus on the biological responses to exposures. Even genetic and genomic alterations serve as evidence of past exposures.[39] The resulting shift in exposomics research from exposure-focused (introduced by Wild)[37] to more metabolomics focused approaches[38] to capture biological endpoints has been accompanied by several changes. Of course, the study of internal metabolic changes and the search for biomarkers in exposomics research is highly necessary, but not the sole goal of exposomics. The last years showed a trend to using increasingly expensive resources and sophisticated techniques, rather than focusing on the initial research question. Research is driven by metabolomics approaches using databases that contain mostly "known" knowledge, necessary to ensure workflow efficiency, yet at the same time trying to find new discoveries and improve the understanding of metabolic pathways. However, with the rapid pace of innovations in the age of the Anthropocene, new synthetic compounds (and their related transformation products) are appearing at an alarming rate with up to 20 million new registrations a year,[40] which can cause new problems and influence the human metabolism very differently. Although it is possible to create dynamic workflows and databases to profit from "new" knowledge generated via high throughput exposomics (eg, PubChemLite for Exposomics[41]), this requires concerted efforts at FAIR and Open data exchange,[42,43] for which resources and infrastructure are still under development. The exchange of information will be vital to support reinterpretation of older data (eg, retrospective screening[44,45]) or large community initiatives for large scale discovery such as the Global Natural Products Social Molecular Networking (GNPS) ecosystem.[46] However, while modern analytics and monitoring techniques can help capture relatively recent perturbations in the exposome, even using natural archives to an extent to investigate past pollution, there is a limit to the availability of suitable samples. The historical exposome concept can give additional perspectives to this technologically driven bio-exposome. Looking at environmental history as such, especially focusing on developments in the Anthropocene, can show many interconnections of environment and health, where expensive technological studies using biological samples of today may fail to find these connections. Historical archives can, for instance, reveal if currently perceived extreme events such as drought or flooding (relative to written records of ∼100 years) were in fact observed centuries earlier. Closer to the exposome, anecdotal, written documents (letters, newspaper articles) can also reveal the catastrophic side-effects of industrial pollution in factories, which is only partially captured in scientific literature (and if covered, often in foreign languages).[47]

Definitions of the environment

One way to approach the exposome from a more holistic perspective is to embed it into the wider framework of what has been called "the environment" since the early post-war years. As demonstrated by the three environmental historians Paul Warde, Libby Robbin, and Sverker Sörlin in their book "The Environment. A History of an Idea",[30] the term gained prominence as a political concept as a result of the success of popular science writers such as William Vogt ("Road to Survival", 1948)[48] and Rachel Carlson's best-seller "Silent Spring" (1962).[29] Although the term had been around for more than a century ("milieu" in French, "Umwelt" in German), a transformation of meaning from "a world where man was molded by environment to him being able to alter the nature of his world"[30 (p. 8)] only occurred with the "environmental revolution" in the 1970s.[49] Sparked by new academic interventions such as the "Limits to Growth" report to the Club of Rome in 1972[50] and the first "oil crisis" in 1973, ecological thinking became a mainstream concern for both scientists, environmental or ecological movements, and world politics. The problem of "the environment" became both scaled (ie, local problems of pollution or waste were interpreted as a subset of a planetary issue) and the result of a complex interplay of past, present, and future temporalities.[21]

To handle the growing complexity of issues at stake, new concepts such as the "ecosystem" were put forward to highlight the holistic nature of environmental changes. Inspired by the post-war boom of "systems theory" and "cybernetics", and long before climate science would emerge as a new interdisciplinary field, even Eugene and Howard Odum introduced the metaphor of balance between the living (organic) and nonliving (abiotic) environment in their textbook "Fundamentals of Ecology" (1953).[51] Ecosystem science became an important framework for setting up large-scale international research projects, shaping the thinking of the environment as a dynamic interplay between "nature" and "nurture" in a system called "earth" at a planetary scale.[30 (pp. 154–158)] Since the introduction of the concept of the "Anthropocene" by Paul Crutzen and Eugene Stoermer in 2000,[8] the human factor has entered the equation: according to recent scholarship, both the living and the nonliving environment has been radically affected by human activity on earth, crossing planetary boundaries.[52-54] With the human being acknowledged as an ecological factor in history (as had been suggested by historian long before),[55] the concept of environment as a tripartite dynamic system has emerged, combining the "natural" environment with the "human built" or socio-technical environment and the subjective environment of the individual.

This "holy trinity" of natural, individual, and social factors that form "the environment" suggests that the physical surrounding in which life takes place, which is comprised of the atmosphere, the biosphere, the hydrosphere and the lithosphere,[56 (p. 14)] is fundamentally interconnected with human activities on earth. We cannot grasp the full complexity of one element without considering the impact of the others. The interconnected, dynamic relationships render "the environment" an incredibly complex system that seems hard to grasp—even from a truly interdisciplinary perspective. Yet it is this "maelstrom" of temporalities, scales, and uncertain causalities that challenges mono-disciplinary approaches and interdict mono-causal scientific explanations.[55] What is needed are new questions, concepts, and narratives aiming at forging an understanding of past and present environmental challenges, starting with a new search for evidence in and significance of existing data from both natural and social archives. Combining evidence from historical and natural sciences and thereby bridging the "two scientific cultures" is a key condition for performing historical exposomics research.

Finding significance in a different way

The search for significance is a common factor across all disciplines. Experts from the fields of humanities, natural and social sciences aim to make significant contributions in their areas of research. In a more general sense, significant contributions have the quality of being important, of being worthy of attention because they answer relevant questions and have meaningful results. In practice, the mindset, practices, and evaluation methods of assessing significance are often specific to each discipline, with a diverse range of views also present within each discipline.

One reason may be the nature of the information sources in each discipline. Focusing on the fields of history and natural sciences, a historian works more frequently with primary sources such as historical manuscripts or photographs, while a researcher in natural sciences often works with numerical data resulting from measurements. The inherent difference in the nature of the sources, as well as the *a priori* impossibility of making measurements in the past or repeating an experiment, lead to different approaches to the assessment of significance. Additionally, an apparent widespread dualism can be perceived, where natural sciences more often try to prove a theory that can be generalized for a larger population, whereas history is more frequently concerned with specific knowledge about a particular fact or occasion, with the consequent differences in the researcher's approach to significance.

Statistical hypothesis testing is an inference method widely used in natural and social sciences to determine a possible conclusion between two hypotheses. A null hypothesis is defined against an alternative hypothesis and the P-value is defined as the probability, calculated under the null hypothesis, that a test statistic is as extreme or more than its observed value. If the P-value is less than the selected significance level, then the null hypothesis is rejected. Traditionally, a $P\text{-value} < 0.05$ has been considered a reasonable significance level, but in some fields such as genomic studies, more stringent levels are adopted. The criticism of this method is extensive across fields of research,[57-60] including technical limitations such as the difference between the characteristics of the scientific data as opposed to the assumptions upon which the significance tests are defined, sampling issues regarding size and randomness, the arbitrary level of significance, the dichotomous reject/not-reject, the misinterpretation of P-values and the lack of reproducibility. However not only technical issues have been raised, but the actual scientific value of the tests has repeatedly been questioned, pointing out weaknesses in the use of isolated tests, the difficulty to generalize results and to integrate previous knowledge, causing erroneous scientific reasoning.

Numerous options have been proposed over decades of research, from the use of stricter P-values to the total abandonment of statistical significance, as well as a wide range of alternatives,[61-63] but the statistical significance in hypothesis testing is still a widely used method due to its intuitive interpretation, ease of calculation with existing tools, and facility given to the choice of the research path with yes/no questions. Bayesian approaches, older than frequentist statistics, gained popularity with the advances in computational methods. Bayesian methods depend on a prior and on the probability of the observed data, allowing the sensitivity of the experiment result to be measured for different priors.[64,65] In Bayesian inference, the uncertainty or "degree of belief" with respect to the parameters is quantified by probability distributions.

Controversy about statistics at a more general level includes the opposition of two attitudes to the question of reality, one realistic or objectivist (ie, pre-existing); and the other, relativistic or historicist (ie, constructed),[66] very present when bringing together history and natural sciences. In *Trust in Numbers*, Theodore M. Porter challenges the ubiquity of quantification in the sciences of nature, highlighting that only a small proportion of the numbers and quantitative expressions "make any pretense of embodying laws of nature, or event providing complete and accurate descriptions of the external world".[67]

There is an increasing interest in qualitative research methods across disciplines, such as research techniques that rely on non-statistical or numerical methods of data collection, analysis and evidence production. Qualitative research allows to retain complexity and nuance through data collection methods that adapt to the context and make it possible to explore emergent issues; and they present a reflexive approach that acknowledges the perspective of the researcher in the process.[68] They are used in the social sciences and the humanities, but they can also complement quantitative approaches in the natural sciences.[69] Qualitative research can be used independently to uncover topics that are not amenable to quantitative research. Furthermore, it can be used as the preliminary of quantitative research, that is to uncover ambiguities and misunderstandings regarding terms and definitions, to help understanding the reasons behind certain results, or to validate quantitative research by providing a different perspective on the same phenomena, sometimes forcing major reinterpretations of quantitative data.[70] Despite the many potential applications, qualitative methods are often evaluated according to quantitative measures of rigor and dismissed as unscientific and unreliable.[71]

Criteria of significance in history includes two main questions: what is important to learn about the past and how do historians know what they know. The first refers to deciding which events or people resulted in a change that had deep consequences. The second is about the historical interpretation based on inference made from sources: what questions turn a source into evidence, who created it, when and for what purpose, what is the historical context and how the inferences can be corroborated.[72] Partington[73] identifies three criteria of significance. First, the importance to people in the past "if we use the egalitarian principle of counting heads to establish priorities here we are in difficulty, because we have unequal access to the opinions and judgements of different social and ethnic groups [...]."[73] Second, objective criteria that includes profundity, as whether an event profoundly changed people's lives; quantity, as the number of lives impacted; and du-rability of the event in time. Thirdly, criterion is relevance, or how an event contributes to an increased understanding of the present.

An interdisciplinary approach to historical exposomics blurs the boundaries between discipline-specific approaches to significance and makes it possible to study the past from new perspectives. Historical sources like pollution measurements included in official reports, complaint letters from the citizens about the dust from the factories and contaminated water, newspapers publishing about new labor laws, innovative industrial techniques, or disease outbreaks, can be used to test hypotheses about pollution concentration or the impact on health in the past, while assessing at the same time for historical significance. The historical archives are filled with scientific and non-scientific, numerical, and non-numerical sources, that together with current scientific knowledge and measurements about water systems or use of chemicals, allow to simulate and validate hypothesis about the past. The dualism “specific” versus “universal” between science and history is replaced by a combination of close and distant reading from a historical and scientific point of view adapted to the source and specific analysis, which allows access to much more information, to analyze it from different points of view, and to use multiple approaches to assess significance.

Evidence from social and natural archives

A historical exposomics approach requires the combination of a great variety of historical and current data. As shown here using data collected for the research project “LuxTIME” as an example, the information gathered comes from both “natural” and “social” archives. Natural archives are physical objects which have been collected, processed and deposited in the environment under natural circumstances (ie, without the interference of the human species), and preserve information about the characteristics of the surroundings from the time and place where deposition took place.[74] Through various analysis techniques, this stored information can be extracted from the archive, enabling the reconstruction of environmental conditions far beyond the period where direct measurements have taken place. This helps determine the baseline conditions at the time, as well as spatial and temporal changes, allowing researchers to relate these to natural or anthropogenic influences.[75]

Social archives

Social archives refer to archives that are the result of a conscious and directed collection of past evidence, be it for cultural (religious, artistic), political (administrative, governmental, legislative), or economic (companies, trade) purposes. Information stored in such archives (be it textual, audio-visual, or material remains/objects) provide documentary evidence of past activities that have shaped the environment, be it through new production and consumption processes, regulatory or legal procedures, or political and cultural discourses on scientific or technological progress, nature protection or market and trade regulation. Based on methods of historical source criticism, social archival records are important to contextualize the development of science, technology, industry production, agriculture, energy consumption and all other human activities that have shaped our contemporary environment.[76] The information or data that can be extracted from social archives contains different kind of evidence, such as narrative data (eg, weather reports on newspapers), long series of quantitative proxies (eg, by procedures related to census or fiscal data), illustrative sources (photographs, drawings, maps), and instrumental measurements (eg, research diaries).[76] Such sources have been used extensively for the production of a rich historiography on Luxembourgish steel industry and the industrialization of the Minett region more generally.[77-79] Comparative studies on the emergence of coal and steel industries in various European regions have highlighted the specificities of the Luxembourgish development, both in terms of availability of natural resources and its dependencies on technology and capital transfers (especially from the Ruhr region), economic alliances (such as the adherence to the German “Zollverein” [customs union] until the end of the Great War), and political constellations (such as the creation of the European Community of Coal and Steel [ECCS] in 1951).[80-84]

The LuxTIME research efforts included consultation of many different archives in Luxembourg, including national and local archives, libraries, museums, and deposits of public administrations (see Figure 1).[85] The sources from these archives were categorized and indexed, specifying the temporal and spatial coverage as well as the exposome category that the sources are referring to (eg, physical/chemical; ecosystem; lifestyle, etc.). The analysis of this data inventory of the social archives has been turned into a treemap showing the number of datasets covering exposome categories and subcategories, see Figure 2.

By means of example (without any claim for comprehensiveness), the following sections describe several natural archives that have been studied or were considered in the framework of the LuxTIME project.

Historical groundwater reserves

Groundwater serves as an archive for historical climatic and hydrological conditions. When aquifer recharge takes place, the ambient environmental conditions are stored in the precipitated water as chemical and isotopic signals, which preserve sequential changes as the groundwater moves away from the point of recharge.[86] Extracted groundwater samples, once related to a proper chronology, for example through radiocarbon dating (^{14}C), have been found to provide indications about the temperature, air-mass circulation and rainfall intensity of the past.[87-89] In its simplest sense, the mere presence of dated groundwater can be an indicator for prolonged wet periods, while absence thereof is a sign of no recharge, possibly due to a period of drought or ice cover.[86,90] The concentration of dissolved noble gases in groundwaters is a function of the temperature and salinity at the time of recharge (Henry’s Law). The concentration decreases with increasing temperature.[91] Since these concentrations are largely unaffected in groundwater bodies, this relationship can be utilized to establish historical temperature records.[87,89,92] The stable isotopes of water (H and O) are separated into heavier and lighter compounds as they move through the hydrological cycle. This creates an isotopic ratio characteristic to the air temperature and humidity at the time of groundwater recharge, as well as environmental processes such as evaporation and condensation.[93] These isotopic ratios have been utilized to reconstruct air-mass circulation and temperatures of the past.[89,92] Age dating uncertainties and mixing of different waters in the aquifer over time creates a smoothing of the signal fluctuations. Groundwater thus serves as an archive which stores an average of the conditions at glacial/interglacial time scales.[89]

Freshwater bivalves

Stable isotopes of oxygen and hydrogen in precipitation and stream water are used for conceptualizing and modeling hydrological, ecological, biogeochemical, and atmospheric processes.[94-97] However, their full potential—for example in climate change impact studies[98] or climate and earth system modeling[99]—cannot be leveraged due to short and truncated time series.[100,101] A promising way forward for reconstructing the history of flowing waters is the conceptualization of rivers as living entities.[102,103] Past changes in eco-hydrological catchment functions and their impacts on freshwater habitats may eventually be recorded in natural archives (and *vice versa*). This feature qualifies trees and freshwater bivalves as (biotic) sensors for reconstructing chronologies of past climate and flowing waters into pre-instrumental times. During the growth process, bivalves (or mollusks) record environmental data in their shells in the form of variable geochemical and microstructural properties, as well as variable increment widths, like tree rings. Controlled by biological clocks, the shell growth process occurs periodically and eventually results in the formation of growth patterns (representing periods of fast and slow growth) that can then be placed in a precise temporal context. By archiving in-stream environmental

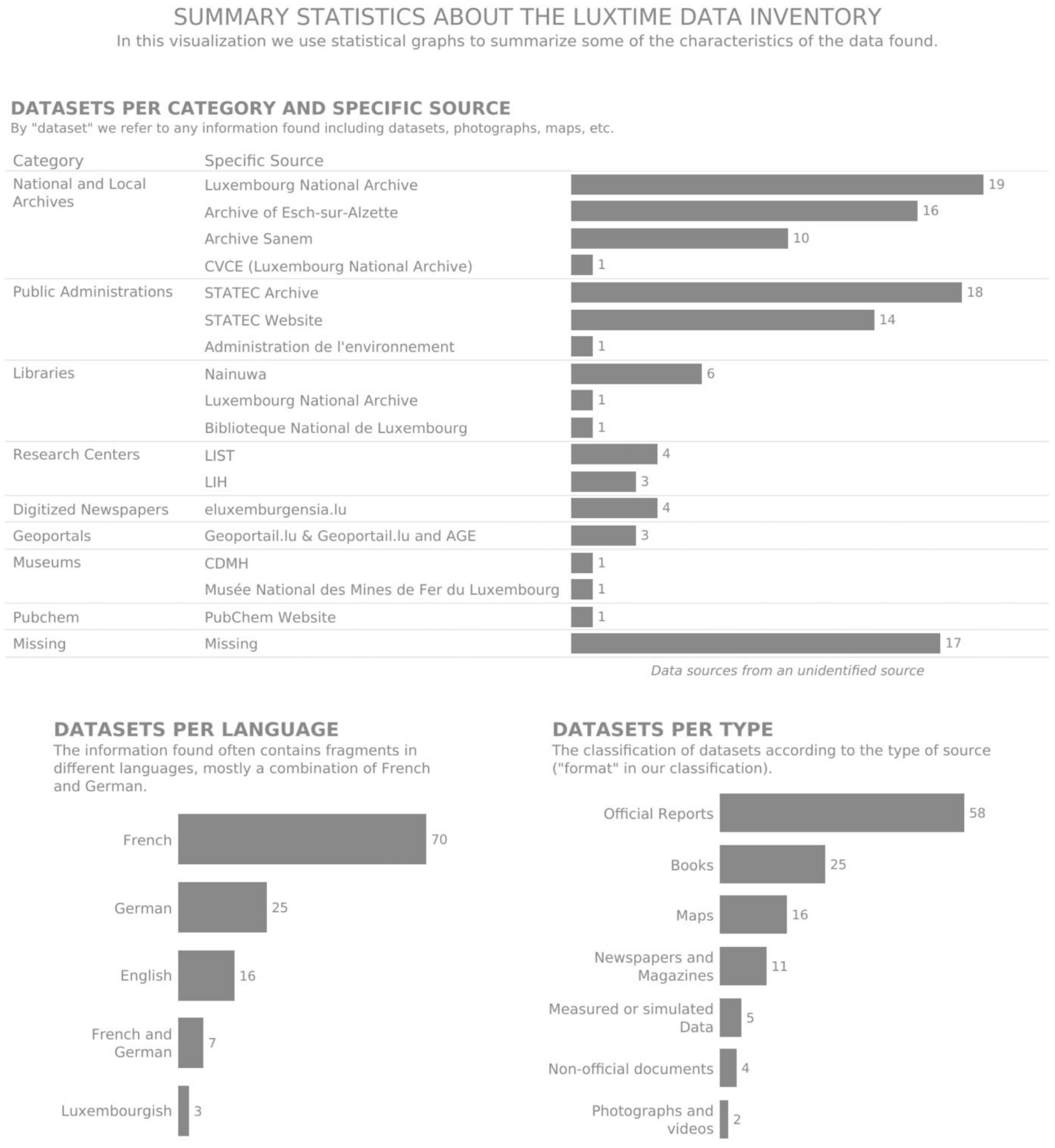


Figure 1. Summary Statistics Data Inventory for the LuxTIME project.

conditions,[104,105] freshwater bivalves thus show considerable potential for complementing stream water isotope records. By forming their shells near equilibrium with the oxygen isotope value of ambient water, changes in isotope values in the shell (represented as δ^{18}O values) can serve two different purposes. They may help in reconstructions of both stream water temperature (if water δ^{18}O values are known) and δ^{18}O values (provided water temperature during shell formation is known or can be reconstructed by other means). While stream water temperature reconstructions have been extensively used,[104] hydroclimate reconstructions have received considerably less attention.[105] Recent proof-of-concept work has demonstrated the potential for freshwater mollusks to solve the problem of limited δ^{18}O isotope records in stream water.[106,107] New analytical protocols based on Secondary Ion Mass Spectrometry have revealed the previously untapped variability in stream water δ^{18}O signatures over nearly 200 years and their connection with changes in atmospheric circulation patterns.

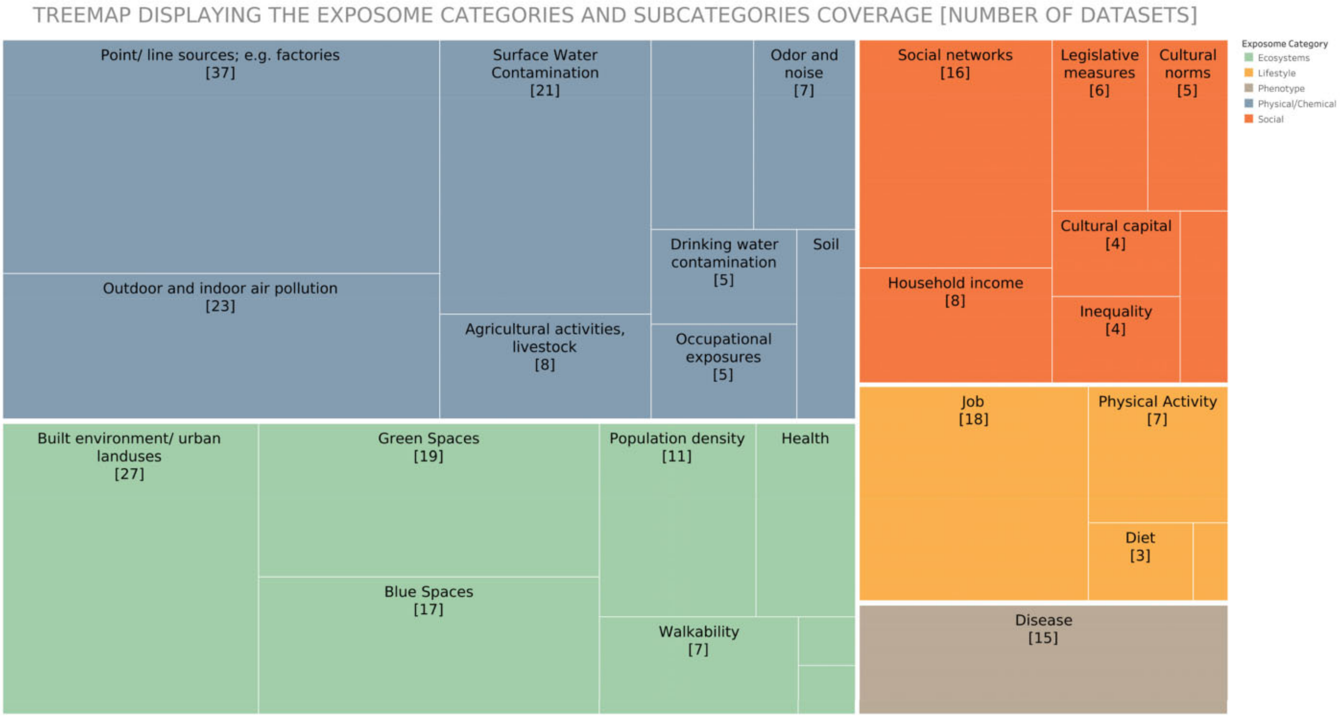


Figure 2. Treemap visualization of exposome categories and subcategories covered by the data inventory.

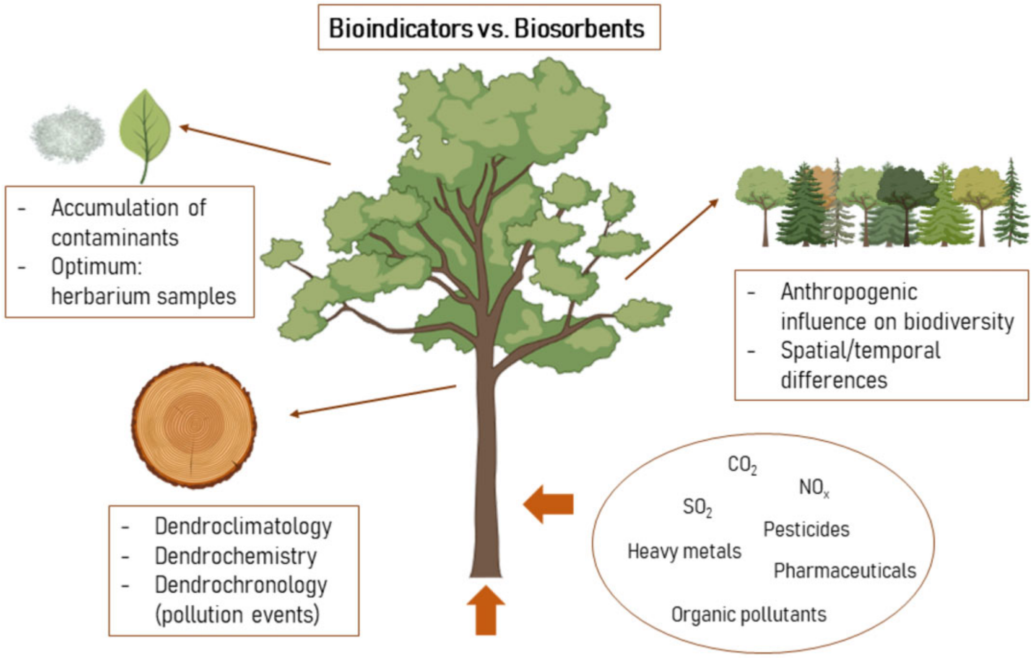


Figure 3. The function of trees as bioindicators versus biosorbents.

Botanical samples

Many studies show that botanical samples can serve as evidence for the impact of man in the age of the Anthropocene.[108-113] Plants, as sensitive organisms, respond strongly to anthropogenic influences such as air pollution. The susceptibility of plants to air pollution can be evaluated by their Air Pollution Tolerance Index (APTI), which is determined by the parameters ascorbic acid, chlorophyll, relative water content, and leaf-extract pH.[114] The APTI is often used studying the potential of plants as biosorbents (remediation activities)[115,116] or bioindicators (biomonitors)[113,117] of pollution (Figure 3).

Plants, such as trees, can be considered as evidence for historical events and thus as natural archives. Trees can provide accurate information about climatic conditions in the past,[118] using dendrochronology information (dendroclimatology). Additionally, those measurements can help to differentiate between natural and anthropogenic causes (increase of CO$_2$ and other gases) of changing environmental conditions.[119]

The population of specific species in certain areas is often connected to human activities. For example, tree species such as birch and alder grow well at post-mining sites, even in highly “contaminated” mine soils, and can even change the microbial properties of the soil.[120,121] As the biodiversity of an area depends directly on factors such as soil, air or water contamination, any changes in plant species may correspond directly or indirectly to an (anthropogenic) change in the direct ecosystem. Analyzing the age of those species could even provide information about the date of these environmental changes, for example, the time of reforestation in post-industrial areas. While looking at historical documents about the influence of air pollution on plants, geographical mapping of species can be found, even if no measurement data is available.[122] If one examines plant species for environmental contamination, one can—under favorable circumstances—make statements about the health of an ecosystem (even in the past). This can be done, for example by analyzing lichens,[123-125] which act as a chemical filter for air pollution, accumulating contaminants over many years. The same applies for tree bark and leaves,[126] however, the time point of pollution is not analyzed in most studies. In the optimal case, one can resort to archived plant samples from herbaria, for example, for the analysis of per- and polyfluoroalkyl substances (PFAS) in pine needles,[127] although samples need to be stored suitably to allow for this analysis. Moreover, using state of the art techniques like matrix-assisted laser desorption ionization—imaging mass spectrometry (MALDI-IMS) allows the monitoring and localization of the contaminant uptake and distribution pathways through the plant.[128,129] There are some studies using tree rings as archives of atmospheric pollution, which can be analyzed in a temporal and spatial manner (dendrochemistry).[108,112] However, there are several limitations and uncertainty factors, such as soil acidity or growth rate to be considered.[130,131] In general, looking at botanical samples always has some uncertainty factors due to plant metabolic processes altering concentrations, as well as external influences (eg, sample handling) influencing all research outcomes—but could nonetheless reveal interesting information about the past that may not otherwise be accessible.

Data visualization

Both natural sciences and humanities use data in their research, whether they are historical or modern texts, measurements from observations or experiments, elements of a photograph, time series, geographical data or many more. Data is the connecting element in interdisciplinary research. Various disciplines often share data analysis methods, but also share the same challenges, such as how to discover what is relevant for their specific research question when dealing with a large volume of data, how to validate the data, or how to communicate the results to readers according to their specific needs. They also face fundamental differences in the definition and use of data and the required communication objectives. Data visualization supports multiple scenarios including discovery and communication. It helps to navigate large volumes of data, whether to perform an initial exploratory data analysis to define the line of research, or to make discoveries that form part of the project results. It is also essential for communicating results, which allows creation of the desired user experience in each case.

Especially in the context of interdisciplinary projects, it is important to understand the different approaches to data visualization since a combination of these might be required for the analysis of data from the point of view of different disciplines. In the case of the natural sciences, data visualization is generally used as a tool to simplify a complex topic, to make data-driven decisions about the best path of research, to build upon previous scientific knowledge by integrating references and benchmarks, and to communicate a clear message to an audience. Scientific data visualization is largely based on statistical graphics which are characterized by the principles of abstraction, reduction, standardization, representation, and legibility. The user is expected to be familiar with the visual elements—the statistical charts, for example, bar charts, line charts—and to ask predefined questions to get formatted answers, to apply filters, and to drill down to predefined levels of aggregation. In the case of the humanities, the use of data visualization often aims at showing the complexity of the subject of research, to engage the user in the interpretation and to present multiple narratives. The data visualization is defined by its granularity, specificity, and full coverage.

To study historical exposomics, data visualization techniques such as concept maps can be used to understand the specific contributions of each field and the joint area of work, see Figure 4.[85]

During the exploration of the data, the use of data visualization allows identification of the types of data available—time series, geographical, texts, images—the geographical area and the period covered, the variety of sources—historical archives, sampling campaigns, web archives—the domains—air, water and soil pollution, demographics, health, industry, urbanism, green and blue areas—and the percentage of sources digitized, among others. This facilitates the identification of gaps, areas of interest and future challenges. Lastly, one of the most frequent use of data visualization is to share the research output. A combined approach to data visualization, scientific and humanistic, allows integration of representative analyses to understand the general impact of the main environmental exposures on health and exploration of the different narratives by integrating individual cases, when relevant.[85]

New historical narratives

In studying historical evidence of human traces on the environment by applying both scientific and hermeneutic methods of interpretation, LuxTIME aims at contributing to a new form of data-driven scholarship on the Anthropocene by experimenting with digital forms of historical storytelling, such as animations of dust pollution in Esch-sur-Alzette over nearly one century (1911–1996) shown in the website “minett-stories” and in Figure 5.

The animation that was produced within the virtual exhibition “Minett Stories”[132]—an interdisciplinary public history project in the framework of the “Esch22/Capital of Europe” initiative,[133] was the result of a collaboration between historians, metallurgy experts and chemists. Based on a data set including information about the production output of the Belval steel plant throughout the 20th century, the type of production process (eg, LD-AC or Thomas for steelmaking), the typical usage of filtering systems, and the prevailing wind directions in Luxembourg over time, the animation offers a plausible approximation of how the dust emissions of the Belval iron and steel complex might have affected local air quality over the decades. Even if the animated map can only offer a hypothetical approximation rather than an exact reproduction of past reality, such temporal visualizations can trigger our historical imagination and open new paths for further historical investigation.

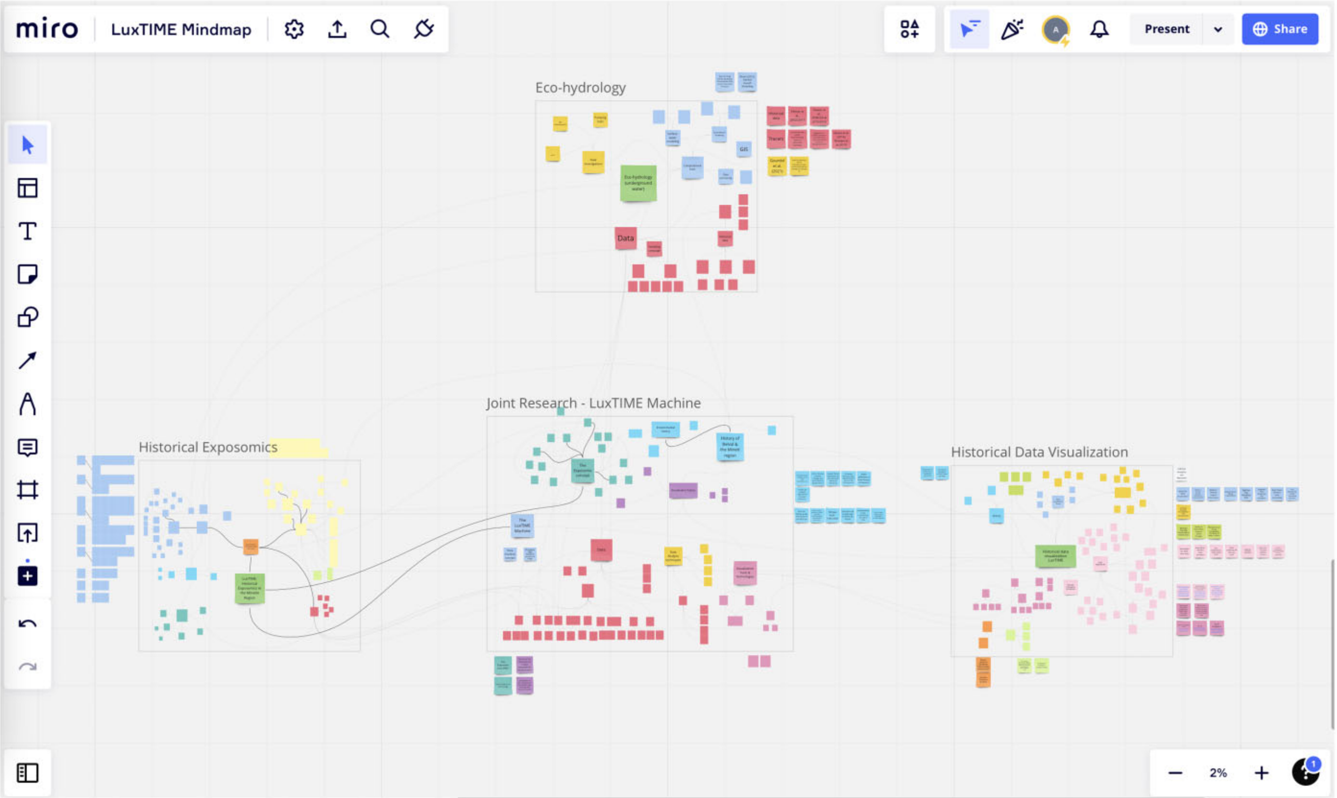


Figure 4. LuxTIME Concept Map including History, Data Visualization, Environmental Cheminformatics and Eco-hydrology.

Figure 5. Screenshots from the Minett Stories Website (https://minett-stories.lu/en/story/air-pollution-visualized).

Conclusions

Bridging the gap between “nature” and “nurture” by bringing a great variety of different historical sources and datasets together is an exciting and challenging exercise in interdisciplinary collaboration. While the LuxTIME project was conceived as a testbed for data-driven scholarship in the field of environmental history with a focus on the Minett region, the interdisciplinary collaboration in the team has opened new research questions and approaches beyond the field of environmental humanities. In exploring the concept of “historical exposomics”, this article hopes to make a convincing argument for enlarging the scope of current research on the exposome by enriching it by historical evidence from “social archives”. In contextualizing the data from both natural and social archives and investigating the causal relationships between factual evidence from the sciences and the narrative evidence stemming from historical sources, new research perspectives on the history of the Anthropocene can be opened, challenging classical forms and formats of historical storytelling and interpretation.[134] The concept of historical exposome, such is our hypothesis, might offer a useful conceptual framework for studying the Anthropocene in a truly interdisciplinary fashion.

Future efforts in this field require research and training in “digital hermeneutics”, looking critically at the use of digital tools and (big) data of all kinds (as covered by the Digital History &

Hermeneutics Doctoral Training Unit DTU of the C^{2}DH).[135] Interdisciplinary collaboration is required for the testing of hypotheses and scientific interpretation of evidence studying the historical exposome. Finding a common language to discuss past, present and future exposomics-related developments is required to tackle this challenging topic.

The historical exposome is presented here as an interdisciplinary approach to address the exposome. Adding value from both perspectives, natural sciences and humanities to study this highly complex paradigm opens new possibilities for present and future research. Instead of solely performing expensive cohort studies looking at the present state, looking at past digital data already present can provide valuable insights. If scientists wish to learn from the past, they must dig into the past, as physical evidence often is not present to be analyzed today. Social archives provide a wealth of soft and hard data sources of information. It is worth investing time to look at historical documents, perform simulations to estimate the past state and even go one step further and simulate future developments. In terms of environmental pollution, looking at past evidence might often prove to be more useful to find for example connections to present day diseases, than analyzing present evidence (eg, looking at past exposure to pesticides studying present day Parkinson's disease[136]). Combining it with today's state of knowledge and tools can help to improve the understanding of the exposome and may help prevent further environmental contamination and disease.

Acknowledgments

We would like to thank all the LuxTIME team members who took part in the discussion.

Author Contributions

Dagny Aurich (Conceptualization [equal], Data curation [equal], Visualization [equal], Writing—original draft [equal], Writing—review and editing [lead]), Aida Horaniet Ibanez (Conceptualization [equal], Data curation [equal], Visualization [equal], Writing—original draft [equal], Writing—review and editing [equal]), Christophe Hissler (Conceptualization [equal], Writing—original draft [supporting], Writing—review and editing [equal]), Simon Kreipl (Conceptualization [supporting], Writing—original draft [equal], Writing—review and editing [equal]), Laurent Pfister (Conceptualization [equal], Writing—original draft [equal], Writing—review and editing [equal]), Emma Schymanski (Conceptualization [equal], Funding acquisition [equal], Supervision [equal], Writing—original draft [supporting], Writing—review and editing [lead]), and Andreas Fickers (Conceptualization [equal], Funding acquisition [equal], Supervision [equal], Writing—original draft [equal], Writing—review and editing [equal])

Funding

D.A. and A.H.I. are funded by the University of Luxembourg Institute for Advanced Studies Audacity Program and S.K. has been funded by the Luxembourg Institute of Science and Technology (LIST) for the Luxembourg Time Machine (LuxTIME). E.L.S. acknowledges funding from the Luxembourg National Research Fund (FNR) for project A18/BM/12341006.

Conflict of interest statement

None declared.

Data availability

Not applicable.

References

1. Misawa K. Nature and nurture. In: Smeyers P, ed. International Handbook of Philosophy of Education. Springer International Handbooks of Education. Springer International Publishing; 2018:905–919. https://doi.org/10.1007/978-3-319-72761-5_64.
2. Canali S. What is new about the exposome? Exploring scientific change in contemporary epidemiology. Int J Environ Res Public Health. 2020;17(8):2879. https://doi.org/10.3390/ijerph17082879.
3. Galison P. Computer simulations and the trading zone. In: Galison P, Stump DJ, eds. The Disunity of Science: Boundaries, Contexts, and Power. Stanford University Press; 1996:118–157.
4. University of Luxembourg. Luxembourg Time Machine. Published 2021. Accessed March 1, 2022. https://luxtimemachine.uni.lu/.
5. Time Machine Europe. Time Machine Europe. Accessed December 8, 2022. https://www.timemachine.eu/.
6. Banken R. Early start and late breakthrough: the industrialization of the Minette region in Lorraine and Luxembourg. In: Czierpka J, Oerters K, Thorade N, eds. Regions, Industries, and Heritage. Palgrave Macmillan, UK; 2015:54–72. https://doi.org/10.1057/9781137333414_5.
7. Hughes JD. An Environmental History of the World: Humankind's Changing Role in the Community of Life. 2nd ed. Routledge; 2009.
8. Crutzen PJ, Stoermer EF. The 'Anthropocene' (2000). In: Robin L, Sörlin S, Warde P, eds. The Future of Nature. Yale University Press; 2017:479–490. https://doi.org/10.12987/9780300188479-041.
9. Waldron P. Critical zone science comes of age. Eos. 2020;101. https://doi.org/10.1029/2020EO148734.
10. Latour B. Some advantages of the notion of "critical zone" for geopolitics. Procedia Earth Planet Sci. 2014;10:3–6. https://doi.org/10.1016/j.proeps.2014.08.002.
11. Jones J. 'Critical Zones' reimagines humanity's relationship with the earth. *Stir World*. Published 2020. Accessed January 12, 2023. https://www.stirworld.com/inspire-people-critical-zones-reimagines-humanitys-relationship-with-the-earth.
12. Solnick S, Latour B, Peter W, eds. Critical zones. *e-flux criticism*. Published 2020. Accessed March 20, 2022. https://www.e-flux.com/criticism/367892/bruno-latour-and-peter-weibel-eds-critical-zones.
13. Chakrabarty D. Anthropocene time. Hist Theory. 2018;57(1):5–32. https://doi.org/10.1111/hith.12044.
14. Sojka RE, Bjorneberg DL, Entry JA. Irrigation: historical perspective. In: Encyclopedia of Soil Science, 3rd ed. CRC Press; 2017:1264–1268. https://doi.org/10.1081/E-ESS3-120042707.
15. Turchin P, Whitehouse H, Gavrilets S, *et al.* Disentangling the evolutionary drivers of social complexity: a comprehensive test of hypotheses. Sci Adv. 2022;8(25):eabn3517. https://doi.org/10.1126/sciadv.abn3517.
16. Headrick DR. Humans versus Nature: A Global Environmental History. Oxford University Press; 2020.
17. Nentwig W. 19 Human Environmental impact in the Paleolithic and Neolithic. In: Henke W, Tattersall I. eds. Handbook of

Paleoanthropology. Springer; 2007:1881–1900. https://doi.org/10.1007/978-3-540-33761-4_62.

18. Uekötter F. Im Strudel: Eine Umweltgeschichte Der Modernen Welt. Campus; 2020.
19. Cloy JM, Farmer JG, Graham MC, MacKenzie AB, Cook GT. Historical records of atmospheric Pb deposition in four Scottish ombrotrophic peat bogs: an isotopic comparison with other records from western Europe and Greenland. Global Biogeochem Cycles. 2008;22(2):1–16. https://doi.org/10.1029/2007GB003059.
20. Stille P, Pourcelot L, Granet M, *et al.* Deposition and migration of atmospheric Pb in soils from a forested silicate catchment today and in the past (Strengbach case): evidence from 210Pb activities and Pb isotope ratios. Chem Geol. 2011;289(1–2):140–153. https://doi.org/10.1016/j.chemgeo.2011.07.021.
21. Brooke JL. Climate Change and the Course of Global History: A Rough Journey. Cambridge University Press; 2014.
22. Watson RT, Rodhe H, Oescher H, *et al.* Intergovernmental Panel on Climate Change. In: Houghton JT, Jenkins GJ, Ephraums JJ, eds. Climate Change: The IPCC Scientific Assessment. Cambridge University Press; 1990.
23. Hooke RLeB, Martín-Duque JF. Land transformation by humans: a review. GSA Today. 2012;12(12):4–10. https://doi.org/10.1130/GSAT151A.1.
24. Brühl CA, Zaller JG. Biodiversity decline as a consequence of an inappropriate environmental risk assessment of pesticides. Front Environ Sci. 2019;7:177. https://doi.org/10.3389/fenvs.2019.00177.
25. Pimm SL, Russell GJ, Gittleman JL, Brooks TM. The future of biodiversity. Science. 1995;269(5222):347–350. https://doi.org/10.1126/science.269.5222.347.
26. Molina MJ, Rowland FS. Stratospheric sink for chlorofluoromethanes: chlorine atom-catalysed destruction of ozone. Nature. 1974;249(5460):810–812. https://doi.org/10.1038/249810a0.
27. Barnes DKA, Galgani F, Thompson RC, Barlaz M. Accumulation and fragmentation of plastic debris in global environments. Philos Trans R Soc Lond B Biol Sci. 2009;364(1526):1985–1998. https://doi.org/10.1098/rstb.2008.0205.
28. Leopold A. A Sand County Almanac and Sketches Here and There. Special commemorative ed. Oxford University Press; 1989.
29. Carson R, Darling L, Silent Spring. Houghton Mifflin; Riverside Press; 1962.
30. Warde P, Robin L, Sörlin S. *The Environment: A History of the Idea*. John Hopkins University Press. 2018.
31. UN. *Report of the United Nations Conference on the Human Environment, Stockholm, June 5–16, 1972*. 1973. Accessed March 31, 2022. https://digitallibrary.un.org/record/523249.
32. UN. *Report of the United Nations Conference on Environment and Development*. 1992. Accessed March 31, 2022. https://www.un.org/esa/dsd/agenda21/Agenda%2021.pdf.
33. UN. *The Montreal Protocol on Substances That Deplete the Ozone Layer*. Vol. 1522. 1989. Accessed November 20, 2022. https://treaties.un.org/doc/publication/unts/volume%201522/volume-1522-i-26369-english.pdf.
34. UN. *Kyoto Protocol to the United Nations Framework Convention on Climate Change*. United Nations Publication; 1998. Accessed 20 November 2022. https://unfccc.int/resource/docs/convkp/kpeng.pdf.
35. UN. *Paris Agreement*. 2015. Accessed March 31, 2022. https://unfccc.int/sites/default/files/english_paris_agreement.pdf.
36. Stockholm Convention. Stockholm Convention-Home page. Accessed March 13, 2023. http://www.pops.int/.
37. Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. Cancer Epidemiol Biomarkers Prev. 2005;14(8):1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456.
38. Miller GW, Jones DP. The nature of nurture: refining the definition of the exposome. Toxicol Sci. 2014;137(1):1–2. https://doi.org/10.1093/toxsci/kft251.
39. Li S, Hursting SD, Davis BJ, McLachlan JA, Barrett JC. Environmental exposure, DNA methylation, and gene regulation. Ann N Y Acad Sci. 2003;983(1):161–169. https://doi.org/10.1111/j.1749-6632.2003.tb05971.x.
40. Arp HPH, Aurich D, Schymanski EL, Sims K, Hale SE. Avoiding the next silent spring: our chemical past, present, and future. Environ Sci Technol. 2023;57(16):6355–6359. https://doi.org/10.1021/acs.est.3c01735.
41. Schymanski EL, Kondić T, Neumann S, Thiessen PA, Zhang J, Bolton EE. Empowering large chemical knowledge bases for exposomics: PubChemLite meets MetFrag. J Cheminform. 2021;13(1):19. https://doi.org/10.1186/s13321-021-00489-0.
42. Schymanski EL, Bolton EE. FAIRifying the exposome journal: templates for chemical structures and transformations. Exposome. 2022;2(1)osab006. https://doi.org/10.1093/exposome/osab006.
43. Schymanski EL, Schymanski SJ. Water science must be Open Science. Nat Water. 2023;1(1):4–6. https://doi.org/10.1038/s44221-022-00014-z.
44. Alygizakis NA, Samanipour S, Hollender J, *et al.* Exploring the potential of a global emerging contaminant early warning network through the use of retrospective suspect screening with high-resolution mass spectrometry. Environ Sci Technol. 2018;52(9):5135–5144. https://doi.org/10.1021/acs.est.8b00365.
45. Alygizakis NA, Oswald P, Thomaidis NS, *et al.* NORMAN digital sample freezing platform: A European virtual platform to exchange liquid chromatography high resolution-mass spectrometry data and screen suspects in "digitally frozen" environmental samples. TrAC Trends Anal Chem. 2019;115:129–137. https://doi.org/10.1016/j.trac.2019.04.008.
46. Wang M, Carver JJ, Phelan VV, *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat Biotechnol. 2016;34(8):828–837. https://doi.org/10.1038/nbt.3597.
47. Blanc PD. Fake Silk: The Lethal History of Viscose Rayon. 1st ed. Yale University Press; 2016. https://doi.org/10.12987/yale/9780300204667.001.0001.
48. Vogt W, Baruch BM, Freeman SI. Road to Survival. W. Sloane Associates; 1948.
49. Chien JP. Umwelt, milieu(x), and environment: a survey of cross-cultural concept mutations. Semiotica. 2007;2007(167):65–89. https://doi.org/10.1515/SEM.2007.071.
50. Meadows DH, Meadows DL, Randers J, Behrens WW. The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind. Universe Books; 1972.
51. Odum EP, Odum HT. Fundamentals of Ecology. Saunders; 1953. Accessed November 25, 2022. https://books.google.de/books?id=EawlAAAAMAAJ.
52. Persson L, Carney Almroth BM, Collins CD, *et al.* Outside the safe operating space of the planetary boundary for novel entities. Environ Sci Technol. 2022;56(3):1510–1521. https://doi.org/10.1021/acs.est.1c04158.
53. Rockström J, Steffen W, Noone K, *et al.* Planetary boundaries: exploring the safe operating space for humanity. E&S. 2009;14(2). https://doi.org/10.5751/ES-03180-140232.
54. Steffen W, Richardson K, Rockström J, *et al.* Planetary boundaries: guiding human development on a changing planet.

Science. 2015;347(6223):1259855. https://doi.org/10.1126/science.1259855.

55. White L. The historical roots of our ecological crisis. Science. 1967;155(3767):1203–1207. https://doi.org/10.1126/science.155.3767.120.
56. Grotzinger JP, Jordan TH. Understanding Earth. 7th ed. W.H. Freeman and Company, a Macmillan Higher Education Company; 2014.
57. Halsey LG. The reign of the p-value is over: what alternative analyses could we employ to fill the power vacuum? Biol Lett. 2019;15(5):20190174. https://doi.org/10.1098/rsbl.2019.0174.
58. Dahiru T. P-value, a true test of statistical significance? A cautionary note. Ann Ib Postgrad Med. 2008;6(1):21–26. https://doi.org/10.4314/aipm.v6i1.64038.
59. Goodman S. A dirty dozen: twelve p-value misconceptions. Semin Hematol. 2008;45(3):135–140. https://doi.org/10.1053/j.seminhematol.2008.04.003.
60. Wagenmakers EJ. A practical solution to the pervasive problems of p values. Psychon Bull Rev. 2007;14(5):779–804. https://doi.org/10.3758/BF03194105.
61. Benjamin DJ, Berger JO. Three recommendations for improving the use of *p*-values. Am Stat. 2019;73(Supp 1):186–191. https://doi.org/10.1080/00031305.2018.1543135.
62. Ioannidis JPA. The proposal to lower *p* value thresholds to.005. JAMA. 2018;319(14):1429–1430. https://doi.org/10.1001/jama.2018.1536.
63. Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. Nature 2019;567(7748):305–307. https://doi.org/10.1038/d41586-019-00857-9.
64. Bayarri MJ, Berger JO. The interplay of Bayesian and frequentist analysis. Statist Sci. 2004;19(1). https://doi.org/10.1214/088342304000000116.
65. Hackenberger, BK. Bayes or not Bayes, is this the question? Croat Med J. 2020;61 (6):564–568. https://doi.org/10.3325/cmj.2019.60.50.
66. Desrosières A. La Politique Des Grands Nombres: Histoire de La Raison Statistique. La Découverte/Poche 2010.
67. Porter TM. Trust in Numbers: The Pursuit of Objectivity in Science and Public Life. Princeton University Press; 1996.
68. Ritchie J, Lewis J, McNaughton Nicholls C, Ormston R. Qualitative Research Practice: A Guide for Social Science Students and Researchers. 2nd ed. SAGE; 2014.
69. Bhangu S, Provost F, Caduff C. Introduction to qualitative research methods–Part I. Perspect Clin Res. 2023;14(1):39–42. https://doi.org/10.4103/picr.picr_253_22.
70. Pope C, Mays N. Qualitative methods in health research. In: Pope C, Mays N, eds. Qualitative Research in Health Care. Blackwell Publishing Ltd; 2006:1–11. https://doi.org/10.1002/9780470750841.ch1.
71. McAleese S, Kilty J. Stories matter: reaffirming the value of qualitative research. Qual Rep. Published online April 17, 2019. Accessed November 25, 2022. https://doi.org/10.46743/2160-3715/2019.3713.
72. Seixas P, Morton T. The Big Six: Historical Thinking Concepts. Nelson Education; 2013.
73. Partington G. What history should we teach? Oxf Rev Educ. 1980;6(2):157–176. https://doi.org/10.1080/0305498800060204.
74. Nowak Z. Theorising the natural archive. Environ Hist Camb. 2022;28(1):105–127. https://doi.org/10.3197/096734019X15740974883852.
75. Blais JM, Rosen MR, Smol JP, eds. Environmental Contaminants: Using Natural Archives to Track Sources and Long-Term Trends of Pollution. 1st ed. Imprint: Springer; 2015. https://doi.org/10.1007/978-94-017-9541-8.
76. Pfister C, Wanner H. Climate and Society in Europe: The Last Thousand Years. Haupt Verlag; 2021.
77. Trausch G. Luxemburg, ein Land vom Stahl geprägt. Banque de Luxembourg; 1995:28.
78. Barthel C, Kirps J, Feltes P. Terres rouges, histoire de la sidérurgie luxembourgeoise. Archives nationales du Luxembourg Centre d'études et de recherches européennes Robert Schuman; 2011.
79. Knebeler C, Scuto D. Belval: Passé, présent et avenir d'un site luxembourgeois exceptionnel (1911-2011). Le Phare; 2010.
80. Leboutte R, Lehners JP. Formation et muations des bassins industriels en Europe. Passé et Avenir Des Bassins Industriel En Europe. Centre Universitaire de Luxembourg; 1995.
81. Leboutte R. Vie et Mort Des Bassins Industriels En Europe. L'Harmattan; 2013.
82. Leboutte R, Puissant J, Scuto D. Un Siècle D'histoire Industrielle (1873 - 1973): Belgique, Luxembourg, Pays-Bas; Industrialisation et Société. SEDES; 1998.
83. Raphael L. Jenseits Von Kohle Und Stahl: Eine Gesellschaftsgeschichte Westeuropas Nach Dem Boom: Frankfurter Adorno-Vorlesungen 2018. Erste Auflage, Suhrkamp; 2019.
84. Brüggemeier FJ. Grubengold: Das Zeitalter Der Kohle Von 1750 Bis Heute. C.H.Beck; 2018.
85. Aurich D, Horaniet Ibanez A. How can data visualization support interdisciplinary research? LuxTIME: Studying historical exposomics in Belval. (forthcoming).
86. Edmunds WM. Groundwater as an archive of climatic and environmental change. In: Aggarwal PK, Gat J, Froehlich KFO, eds. Isotopes in the Water Cycle: Past, Present and Future of a Developing Science. Springer; 2005:341–352.
87. Andrews JN, Lee DJ. Inert gases in groundwater from the Bunter Sandstone of England as indicators of age and palaeoclimatic trends. J Hydrol. 1979;41(3–4):233–252. https://doi.org/10.1016/0022-1694(79)90064-7.
88. Gasse F. Hydrological changes in the African tropics since the Last Glacial Maximum. Quat Sci Rev. 2000;19(1–5):189–211. https://doi.org/10.1016/S0277-3791(99)00061-X.
89. Stute M, Talma AS. Glacial temperatures and moisture transport regimes reconstructed from noble gases and δ18O, Stampriet aquifer, Namibia. In: Isotope Techniques in the Study of Environmental Change: Proceedings of an International Symposium on Isotope Techniques in the Study of past and Current Environmental Changes in the Hydrosphere and the Atmosphere. Proceedings series. International Atomic Energy Agency; 1998:307–318.
90. Sonntag C, Muennich KO, Junghans C, *et al.* Palaeoclimatic information from deuterium and oxygen-18 in carbon-14-dated north Saharian groundwaters. Groundwater formation in the past. In: Agence Internationale De L'Energie Atomique. [Vienne.] Conférence. [1978, 19–23 juin. Neuherberg.], Organisation Des Nations Unies Pour L'Education, La Science Et La Culture. [Paris.] Conférence. [1978, 19–23 juin. Neuherberg.], eds. *Isotope Hydrology 1978: Proceedings of an International Symposium on Isotope Hydrology Jointly*. Proceedings series. I.A.E.A; 1979:341–352.
91. Benson BB, Krause D. Empirical laws for dilute aqueous solutions of nonpolar gases. J Chem Phys. 1976;64(2):689–709. https://doi.org/10.1063/1.432215.
92. Stute M, Deak J. Environmental Isotope Study (14-C, 13-C, 18-O, D, Noble Gases) on Deep Groundwater Circulation Systems in Hungary with Reference to Paleoclimate. Radiocarbon 1989; 31(03):902–918.

93. Geyh M. Environmental Isotopes in the Hydrological Cycle: Principles and Applications-Vol. IV. International Hydrological Programm; 2000:200.
94. McGuire KJ, McDonnell JJ. A review and evaluation of catchment transit time modeling. J Hydrol. 2006;330(3–4):543–563. https://doi.org/10.1016/j.jhydrol.2006.04.020.
95. Rinaldo A, Benettin P, Harman CJ, *et al.* Storage selection functions: a coherent framework for quantifying how catchments store and release water and solutes: on storage selection functions. Water Resour Res. 2015;51(6):4840–4847. https://doi.org/10.1002/2015WR017273.
96. Kirchner JW. Aggregation in environmental systems - Part 2: Catchment mean transit times and young water fractions under hydrologic nonstationarity. Hydrol Earth Syst Sci. 2016; 20(1):299–328. https://doi.org/10.5194/hess-20-299-2016.
97. Rodriguez NB, Klaus J. Catchment travel times from composite storage selection functions representing the superposition of streamflow generation processes. Water Resour Res. 2019; 55(11):9292–9314. https://doi.org/10.1029/2019WR024973.
98. Rank D, Wyhlidal S, Schott K, Weigand S, Oblin A. Temporal and spatial distribution of isotopes in river water in Central Europe: 50 years experience with the Austrian network of isotopes in rivers. Isotopes Environ Health Stud. 2018;54(2): 115–136. https://doi.org/10.1080/10256016.2017.1383906.
99. Risi C, Ogee J, Bony S, *et al.* The Water Isotopic Version of the Land-Surface Model ORCHIDEE: implementation, evaluation, sensitivity to hydrological parameters. Hydrol Curr Res. 2016; 07(04):1–24. https://doi.org/10.4172/2157-7587.1000258.
100. Halder J, Terzer S, Wassenaar LI, Araguás-Araguás LJ, Aggarwal PK. The Global Network of Isotopes in Rivers (GNIR): integration of water isotopes in watershed observation and riverine research. Hydrol Earth Syst Sci. 2015;19(8):3419–3431. https://doi.org/10.5194/hess-19-3419-2015.
101. Vitvar T, Aggarwal PK. Monitoring Isotopes in Rivers: Creation of the Global Network of Isotopes in Rivers (GNIR): Results of a Coordinated Research Project 2002-2006. International Atomic Energy Agency; 2012.
102. Dooge JCI. Looking for hydrologic laws. Water Resour Res. 1986; 22(9S):46S–58S. https://doi.org/10.1029/WR022i09Sp0046S.
103. Zehe E, Ehret U, Pfister L, *et al.* HESS Opinions: from response units to functional units: a thermodynamic reinterpretation of the HRU concept to link spatial organization and functioning of intermediate scale catchments. Hydrol Earth Syst Sci. 2014; 18(11):4635–4655. https://doi.org/10.5194/hess-18-4635-2014.
104. Versteegh EAA, Vonhof HB, Troelstra SR, Kaandorp RJG, Kroon D. Seasonally resolved growth of freshwater bivalves determined by oxygen and carbon isotope shell chemistry: FRESHWATER BIVALVE INTRA-ANNUAL GROWTH. Geochem Geophys Geosyst. 2010;11(8):1–16. https://doi.org/10.1029/2009GC002961.
105. Kelemen Z, Gillikin DP, Graniero LE, *et al.* Calibration of hydroclimate proxies in freshwater bivalve shells from Central and West Africa. Geochim Cosmochim Acta. 2017;208:41–62. https://doi.org/10.1016/j.gca.2017.03.025.
106. Pfister L, Grave C, Beisel JN, McDonnell JJ. A global assessment of freshwater mollusk shell oxygen isotope signatures and their relation to precipitation and stream water. Sci Rep. 2019; 9(1):4312. https://doi.org/10.1038/s41598-019-40369-0.
107. Schöne BR, Meret AE, Baier SM, *et al.* Freshwater pearl mussels from northern Sweden serve as long-term, high-resolution stream water isotope recorders. Hydrol Earth Syst Sci. 2020; 24(2):673–696. https://doi.org/10.5194/hess-24-673-2020.
108. Alterio E, Cocozza C, Chirici G, Rizzi A, Sitzia T. Preserving air pollution forest archives accessible through dendrochemistry. J Environ Manage. 2020;264:110462. https://doi.org/10.1016/j.jenvman.2020.110462.
109. Barroso PJ, Martín J, Santos JL, Aparicio I, Alonso E. Analytical method for the evaluation of the outdoor air contamination by emerging pollutants using tree leaves as bioindicators. Anal Bioanal Chem. 2018;410(2):417–428. https://doi.org/10.1007/s00216-017-0733-8.
110. Dzierzanowski K, Popek R, Gawrońska H, Saebø A, Gawroński SW. Deposition of particulate matter of different size fractions on leaf surfaces and in waxes of urban forest species. Int J Phytoremediation. 2011;13(10):1037–1046. https://doi.org/10.1080/15226514.2011.552929.
111. Krizkova S, Ryant P, Krystofova O, *et al.* Multi-instrumental analysis of tissues of sunflower plants treated with silver(I) ions - plants as bioindicators of environmental pollution. Sensors (Basel). 2008;8(1):445–463. https://doi.org/10.3390/s8010445.
112. Perone A, Cocozza C, Cherubini P, *et al.* Oak tree-rings record spatial-temporal pollution trends from different sources in Terni (Central Italy). Environmental Pollution. 2018;233: 278–289. https://doi.org/10.1016/j.envpol.2017.10.062.
113. Forbes PBC, van der Wat L, Kroukamp EM. Biomonitors. In: Comprehensive Analytical Chemistry. Vol 70. Elsevier; 2015: 53–108. https://doi.org/10.1016/bs.coac.2015.09.003.
114. Singh SK, Rao DN, Agrawal M, Pandey J, Naryan D. Air pollution tolerance index of plants. J Environ Manage. 1991;32(1):45–55. https://doi.org/10.1016/S0301-4797(05)80080-5.
115. Kvesitadze G, Khatisashvili G, Sadunishvili T, Kvesitadze E. Plants for remediation: uptake, translocation and transformation of organic pollutants. In: Öztürk M, Ashraf M, Aksoy A, Ahmad MSA, Hakeem KR, eds. Plants, Pollutants and Remediation. Springer Netherlands; 2015:241–308. https://doi.org/10.1007/978-94-017-7194-8_12.
116. Adeniyi AG, Ighalo JO. Biosorption of pollutants by plant leaves: an empirical review. J Environ Chem Eng. 2019;7(3): 103100. https://doi.org/10.1016/j.jece.2019.103100.
117. Sahu C, Basti S, Sahu SK. Air pollution tolerance index (APTI) and expected performance index (EPI) of trees in Sambalpur town of India. SN Appl Sci. 2020;2(8):1327. https://doi.org/10.1007/s42452-020-3120-6.
118. Briffa KR, Schweingruber FH, Jones PD, *et al.* Trees tell of past climates: but are they speaking less clearly today? Philos Trans R Soc Lond B Biol Sci. 1998;353(1365):65–73. https://doi.org/10.1098/rstb.1998.0191.
119. Jacoby GC, D'Arrigo RD. Tree rings, carbon dioxide, and climatic change. Proc Natl Acad Sci U S A. 1997;94(16):8350–8353. https://doi.org/10.1073/pnas.94.16.8350.
120. Pietrzykowski M. Tree species selection and reaction to mine soil reconstructed at reforested post-mine sites: Central and eastern European experiences. Ecol Eng. 2019;142:100012. https://doi.org/10.1016/j.ecoena.2019.100012.
121. Chodak M, Sroka K, Woś B, Pietrzykowski M. Chemical and microbial properties of post-mining and post-fire soils afforested with different tree species. Appl Soil Ecol. 2022;171:104321. https://doi.org/10.1016/j.apsoil.2021.104321.
122. Diederich P. Cartographie de la pollution atmosphérique au grand-duché de Luxembourg à l'aide de bioindicateurs. Société des naturalistes luxembourgeois association sans but lucratif SNL; 1979.
123. Hissler C, Stille P, Krein A, *et al.* Identifying the origins of local atmospheric deposition in the steel industry basin of Luxembourg using the chemical and isotopic composition of

the lichen Xanthoria parietina. Sci Total Environ. 2008; 405(1–3):338–344. https://doi.org/10.1016/j.scitotenv.2008.05.029.

124. Will-Wolf S, Jovan S, Amacher MC. Lichen elements as pollution indicators: evaluation of methods for large monitoring programmes. The Lichenologist. 2017;49(4):415–424. https://doi.org/10.1017/S0024282917000299.
125. Forbes P, Wat L, Strumpher J. Comparative Perspectives on Extraction Methods for Organic Metabolites and Pollutants from Lichens. In: Yusuf M, ed. Lichen-Derived Products. 1st ed. Wiley; 2020:27–73. https://doi.org/10.1002/9781119593249.ch2.
126. El-Khatib AA, Barakat NA, Youssef NA, Samir NA. Bioaccumulation of heavy metals air pollutants by urban trees. Int J Phytoremediation. 2020;22(2):210–222. https://doi.org/10.1080/15226514.2019.1652883.
127. Kirkwood KI, Fleming J, Nguyen H, Reif DM, Baker ES, Belcher SM. Utilizing pine needles to temporally and spatially profile per- and polyfluoroalkyl substances. Pharmacol Toxicol. 2021. https://doi.org/10.1101/2021.08.24.457570.
128. Lagarrigue M, Caprioli RM, Pineau C. Potential of MALDI imaging for the toxicological evaluation of environmental pollutants. J Proteomics. 2016;144:133–139. https://doi.org/10.1016/j.jprot.2016.05.008.
129. Annangudi SP, Myung K, Avila Adame C, Gilbert JR. MALDI-MS Imaging Analysis of Fungicide Residue Distributions on Wheat Leaf Surfaces. Environ Sci Technol. 2015;49(9):5579–5583. https://doi.org/10.1021/es506334y.
130. Ferretti M, Innes JL, Jalkanen R, *et al.* Air pollution and environmental chemistry – what role for tree-ring studies? Dendrochronologia 2002;20(1–2):159–174. https://doi.org/10.1078/1125-7865-00014.
131. Ruppert H, Wischow D. Tree rings—A questionable indicator of heavy metal pollution in air or soil. Geochim Cosmochim Acta. 2006;70(18):A546. https://doi.org/10.1016/j.gca.2006.06.1007.
132. Van de Maele J. *Minett Stories - Air Pollution Visualized. Minett Stories*. Published 2022. Accessed March 31, 2023. https://minett-stories.lu/en/story/air-pollution-visualized.
133. Esch2022. *Esch2022*. Published 2022. Accessed March 13, 2023. https://esch2022.lu/de/aboutesch/.
134. Chakrabarty D, Latour B. The Climate of History in a Planetary Age. The University of Chicago Press; 2021.
135. Digital History & Hermeneutics Doctoral Training Unit (DTU). C2DH | Luxembourg Centre for Contemporary and Digital History. Published 2017. Accessed December 9, 2022. https://www.c2dh.uni.lu/de/projects/digital-history-hermeneutics-doctoral-training-unit-dtu.
136. Nandipati S, Litvan I. Environmental Exposures and Parkinson's Disease. Int J Environ Res Public Health. 2016;13(9):881. https://doi.org/10.3390/ijerph13090881.

Chapter 3: Simulating and visualizing data in environmental history: Airbone dust concentration from the Belval plant in Luxembourg (1911-1997)

***Contribution statement*:** All authors have contributed to the conceptualization, review and editing of this article. I have written section 2 “Modelling and simulation” and 7 “data visualization”, and I have contributed to writing the description of the data processing and visualization in section 6 “The simulation process”. I have also created the data visualization video in Python and added the screenshots of the process to the article. Dagny Aurich has written section 3, 5, 6, 8 and the conclusions. Jens van de Maele has written the abstract, the introduction, and section 4.

***Submission statement**:* This article has been submitted to the Journal of Digital History, as a Jupyter notebook using Python. It has passed the technical review and is currently under peer review.

Simulating and visualising data in environmental history: Airborne dust concentration from the Belval plant in Luxembourg (1911-1997)

Dagny Aurich[1]*, Aida Horaniet Ibanez[2]*, Jens van de Maele[2]

[1] Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 6 avenue du Swing, 4367, Belvaux, Luxembourg.

[2] Luxembourg Centre for Contemporary and Digital History (C²DH), University of Luxembourg, 11, Porte des Sciences, 4366 Esch-sur-Alzette, Luxembourg.

*Corresponding authors: dagny.aurich@uni.lu, aida.horanietibanez@uni.lu

To be published at: https://journalofdigitalhistory.org/en

Keywords: Dust Pollution, Environment, Data visualization, Simulation, Steel Industry, History of Luxembourg

Abstract

Luxembourg's industrial history can be reconstructed using primary sources containing quantitative and qualitative industrial and environmental data, which can then be analysed and visualised digitally. This research combines current scientific knowledge and historical data to model the airborne dust concentration generated by the Belval steelworks in Esch-sur-Alzette between 1911 and 1997. The calculations are based on a model of atmospheric dispersion, using parameters such as production volumes, the presence of filter systems and meteorological data. Visualisation of the simulated data offers insights into the extent and variability of dust concentrations over time, along with the underlying reasons (e.g. wars or technical innovations). An interdisciplinary approach allowed the integration of chemical and health perspectives, which were then embedded into a historical context. Simulations, a form of data generation widely used in engineering, open new opportunities for the re*creation* of historical data, which can be compared with scientific research. These historical data can also be re-analysed and visualised on the basis of new and/or current perspectives.

This research is a collaboration between the REMIX project (an interdisciplinary project focusing on the history of the Minett region in connection with Esch-sur-Alzette – European Capital of Culture 2022) and the interdisciplinary LuxTIME project (which attempts to build and visualise historical datasets with the aim of studying the impact of environmental changes on Belval's population). The aim is to explore new ways of generating, validating, analysing and visualising past data through interdisciplinary collaboration, rather than to validate a model for the study of historical dust pollution. This work discusses (*a*) the model, (*b*) the parameters used, (*c*) the results of the dust simulation, (*d*) the impact on health, (*e*) the historical sources, and (*f*) the limitations of the model and the challenges involved in visualising and simulating historical data.

1. Introduction

Air pollution has a history. In today's society, this historicity is most commonly perceived in the context of global warming, where the 'weight' of billions of tonnes of carbon dioxide released in the past has become a factor of major importance for our global ecosystem. The mechanism behind this has been elegantly described by Malm, who observed that our climate has essentially become a 'product of past emissions', while the 'emissions produced by cars [today will generate] their greatest impact on generations not yet born'. For Malm, present-day emissions are effectively 'invisible missiles aimed at the future' (Malm 2016). The adjective 'invisible' is rather important here. After all, the peculiar interrelationship between past, present and future is not the only paradoxical characteristic of air pollution history (on this interrelationship, see also (Uekötter 2020)). Another paradox is the fact that the ongoing global warming is caused by a gas that, when encountered under everyday circumstances, is both invisible and odourless. This lack of perceptibility does not imply, however, that pollution problems went completely unnoticed by previous generations. Prior to the scientific and societal problematisation of the greenhouse effect, other manifestations of air pollution – such as dust, dark smoke, soot particles and stench – were particularly palpable in cities and industrial regions, and as such were a regular source of indignation (for two classic studies on this indignation, see (Brüggemeier and Rommelspacher 1992; Mosley 2001)).

The southwest of the Grand Duchy of Luxembourg, an important iron and steelmaking region from the 1870s onward, was one of those places where people were constantly faced with air pollution that was as omnipresent as the industrial plants themselves.[1] In the Minett, as Luxembourg's metallurgical region is called, the soot and smoke of the steelworks were part of a 'sensescape' that could be seen and smelled on a near-permanent basis (on the notion of sensescape, see (Pritchard and Zimring 2020)). To borrow the words of an elderly inhabitant of Dudelange, who recently reminisced about his life in Luxembourg's industrial region: day in, day out, 'the dust and dirt fell from the skies' (Back-Hoffmann 2021). In order to objectify the severity of this problem, several scientific studies were carried out in the Minett towns of Differdange and Esch-sur-Alzette between the late 1930s and the early 1970s, with the specific aim of measuring the typical quantity of airborne dust particles at or near industrial sites. Yet these measurements – which we discuss in section 8.2 – only lasted for a limited period (e.g. three months), and thus only provide a snapshot in time.

In this contribution, we attempt to create a *long-term* picture of past dust pollution levels by using the Belval iron and steel complex in Esch-sur-Alzette – the informal capital of the Minett (hereafter: Esch) – as a test case (Figure 1). For this purpose, historians, metallurgy experts, chemists and data visualisation specialists have collaborated to analyse a major metallurgical pollutant: dust. We present a simulation and visualisation of the diachronic evolution of local airborne dust concentrations in the vicinity of the Belval plant, using various parameters as input for our model. These include the production numbers of the plant throughout the 20th century, the production process (e.g. LD-AC or Thomas for steelmaking), the typical usage of filtering systems, and the prevailing wind directions over time. Based on this data set, we offer a plausible approximation of how the dust emissions of the Belval iron and steel complex might

---

[1] For the economic and social history of Luxembourg's metallurgical industry, see the six-volume series *Terres Rouges: Histoire de la sidérurgie luxembourgeoise,* published from 2009 to 2018 by the National Archives of Luxembourg.

have affected the local air quality between 1911 (the first year of operation) and 1997 (when the last blast furnace was shut down).

Figure 1: Smoking chimneys of the Belval steel plant. Slide from around 1960. Source: Archives de la Ville d'Esch-sur-Alzette.

An important precedent for modelling in environmental history can be found in the work of Tarr, who has investigated emissions of man-made components such as pesticides, herbicides and heavy metals in the estuary of the Hudson and Raritan rivers between 1700 and 1980. To achieve this aim, Tarr worked together with physicists and environmental engineers ((Tarr 1996); see also (Tarr 2021)). Likewise, in their 2021 study on climate change and its effects on people, Pfister and Wanner attempted to reconcile the scientific cultures of the natural sciences and the humanities: 'Scientists explain how natural systems work, while (environmental) historians tell stories of people who grapple with the effects of weather and climate' (Pfister and Wanner 2021). As such, the research by Pfister and Wanner is in line with the approach outlined by Guldi and Armitage, who have discussed the role of history in the process of understanding 'the multiple pasts which gave rise to our conflicted present'. Regarding climate change, Guldi and Armitage have argued that 'history, with its rich, material understanding of human experience and institutions and its apprehension of multiple causality, is re-entering the arena of long-term discussions of time where evolutionary biologists, archaeologists, climate scientists, and economists have long been the only protagonists' (Guldi and Armitage 2014).

Similar to the way in which environmental history and the history of technology have become more entangled in the last decade (resulting in the development of a crossover field known as 'envirotech', e.g. (Pritchard and Zimring 2020)), interdisciplinary endeavours between the human and natural sciences are indeed essential if one wants to adequately evaluate environmental changes that have occurred in the past. This has been made explicit by Massard-Guilbaud:

> How can we write a history of climate without contacting climatologists? A history of pollution without touching on chemistry? A history of sanitation systems or energy resources without taking an interest in technology? Let us note, moreover, that it is not only historians who are discovering this need for interdisciplinarity; specialists in other disciplines […] are also approaching historians [...] on environmental issues because [...] they feel the need to do so. (Massard-Guilbaud 2007)[2]

The research questions that motivate our own interdisciplinary research are as follows. Throughout the 20th century, how were people living near the Belval iron and steel complex affected by dust pollution? Can this pollution from the past be quantified and displayed on a map? And how did it evolve over time? Using modelling and simulation with historical data as input, we generate new data allowing us to analyse such questions, thus uniting the points of view of exact sciences (e.g. the health impact of various concentration levels) and history (e.g. the industrial development of the region and the societal narratives about pollution). We hypothesise that the pollution generated by the Belval plant affected many (if not all) of the inhabited quarters of both Esch and the surrounding municipalities (including some on French territory, which would make Belval a case of cross-border air pollution).

It must be emphasised that our contribution offers only an *estimation* of pollution values, due to the limitations inherent in any simulation. In the current study, such limitations include the lack of adequate corporate sources on the presence and efficiency of filtering systems, forcing a reliance on assumptions (based on data in non-Luxembourgish secondary literature on ideal typical metallurgical filtering systems). Another important limitation of our model is the omission of other pollution sources, such as iron and steel complexes other than Belval (including two more plants in Esch), non-metallurgical industrial complexes, traffic and domestic heating. As such, we agree with a warning recently posited by Joy Parr concerning interdisciplinary work. While Parr endorsed collaborations between 'researchers in environmental history and the history of technology' and scientists from fields like 'metallurgy [and] bio-chemistry […]', she pointed at possible incompatibilities between historical science and the natural sciences, due to the latter's tendency to 'simplify' for the purpose of experimentation, as well as to 'derogate knowledge and reasoning that is not readily represented in symbols and signs' (Parr 2010). In our interdisciplinary contribution, we take a rather modest view on this matter, by explicitly pointing at the potential of our research as well as the lacunae and possibilities for future improvement.

The structure of our article can be summarised as follows. In section 2, we discuss the terms 'modelling' and 'simulation', their history and application across disciplines, and their largely untapped potential in the humanities. Section 3 explains the general impact of dust pollution on health, in order to provide a basis for the subsequent analysis of our results. Section 4 establishes the historical context of the metallurgical industry in Luxembourg, the Minett region and Esch, while also briefly exploring the societal narratives about pollution. In section 5, we introduce steelmaking production techniques as well as their evolution and impact on dust pollution. In section 6, we explain the actual simulation process by focusing on the data generated by the dispersion model. Here, we discuss in detail the data basis and the model specifications, to explore the hypothesis presented (regarding the impact of pollution on the inhabitants of Esch and the surrounding area). In section 7, we explore the data visualisation process. Section 8 contains an analysis of the results including the validation of the model used, a basic Gaussian dispersion

---

[2] All translations throughout have been done by the authors.

model that has already been proven through its use in similar scenarios (in various contexts) (Leelőssy et al. 2014); this section is focussing on our way of applying the model. Further, a comparison of the model results against available measurements is presented in section 8.2, and the limitations and potential ways to improve the model in future research in 8.3 and 8.4. For future research on historical dust pollution, the model could be used with alternative data sources for new simulations, or new models could be tested adding new variables. Our conclusions are offered in the final section (9).

2. Modelling and simulation

Since modelling and simulation are key elements in our research, a brief introduction of terms and usage across disciplines is appropriate. Banks defined simulation as 'the imitation of the operation of a real-world process or system over time' (such as the generation and dispersion of dust pollution of a steel plant) that can model both existing and conceptual systems (Banks 1999). As such, simulation is used for a variety of purposes including prediction, performance, training, entertainment, proof and discovery. While *induction* can be used to find patterns in data and *deduction* to find consequences of assumptions, *simulation* generates data that can be analysed inductively (Carson II 2005).

Initiated during the Second World War 'to address problems too complex for theory and too remote from laboratory materials for experiment' (Galison 1996), simulation spread through the social sciences by the end of the 1960s, where it was (and still is) often used for modelling artificial populations and investigating human behaviour (Lebherz et al. 2018). However, 'it has remained almost unknown to the humanities' (McCarty 2016), notwithstanding a few exceptions. Apart from the environmental historical research mentioned in the introduction, Lebherz et al. have presented a case of text mining by scientific workflows and computer simulation, with the aim of investigating potential influences on the author's literary productivity. In this context, they have identified four prerequisites: a solid data basis, the specification and construction of a model, the definition of hypotheses to answer research questions, and the reuse of proven models in similar scenarios (Lebherz et al. 2018). Simulation is also frequently used in certain fields of archaeological research, such as evolutionary archaeology and the study of human evolution (Lake 2014).

Champion has discussed the distinction between model ('a physical or digital representation of a product or process') and simulation ('the re-configurative use of a model to reveal new and potential aspects') (Champion 2016). In the context of the humanities, McCarty has referred to modelling as 'the analogical bridge between computing and the interpretative disciplines' that 'keeps the digital construct separate, informing humanistic research both by what it discovers and especially by what it cannot'. The defining moment of simulation occurs 'when it becomes the only way to know something or to form a coherent picture from fragmentary knowledge' (McCarty 2016).

3. Dust pollution from a chemical and health perspective

Pollution caused by dust (atmospheric particulate matter) is a global problem, both historically and in the present (Butte and Heinzow 2002; Han et al. 2021). Dust coming from various anthropogenic sources such as industry, households and traffic is a major source of chemical contamination of entire regions and has a negative impact on health. Dust particles can also come from natural sources, like pollen, microorganisms and the natural erosion of soils. The size, origin and chemical composition of dust particles play a major role when looking at health effects. Airborne dust comes in all shapes and sizes, which are either visible or invisible to the naked eye (aerodynamic diameter 1 to 400 µm) (Kumar and Kumar 2018).

Large dust particles settle quickly (> 100 µm) or are trapped in the nose, mouth and larynx region when inhaled (*inhalable* fraction ≤ 100 µm), as shown in Figure 2 (Wippich et al. 2020). Smaller particles stay for longer in the air, while very small particles can penetrate the respiratory system. The *respirable* dust fraction that can enter the alveolar region is defined to be below an aerodynamic diameter of 10 µm, independent of the particle's length (WHO n.d.). Small particles that remain in the gas-exchange region of the lungs can cause allergic reactions, cancer and other serious diseases or disorders (Wippich et al. 2020; WHO n.d.; Bala and Tabaku 2010).

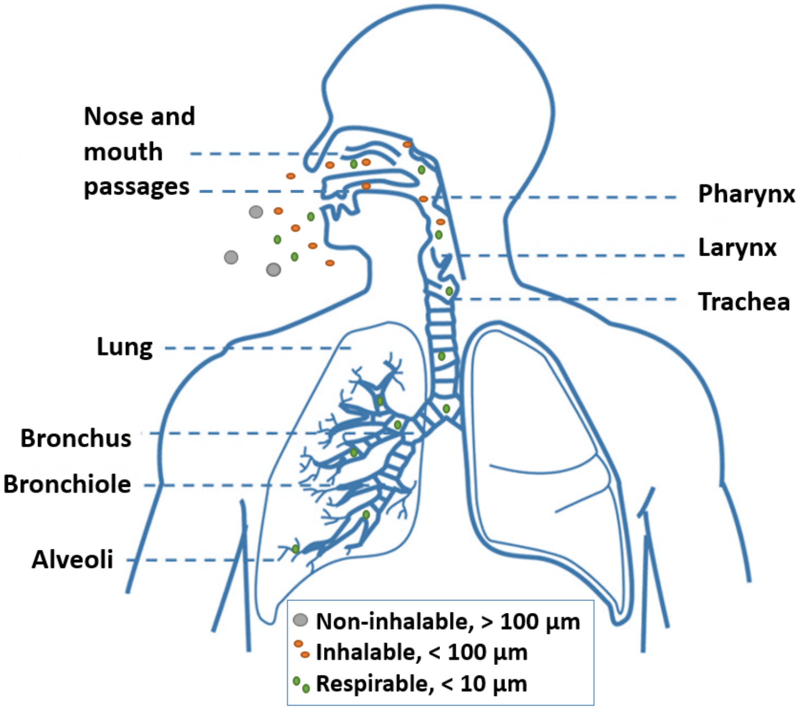


Figure 2: Respiratory system with different dust fractions and their aerodynamic diameter

Dust sources can be either point or area sources, from where particles (having long lifetimes) can spread over many kilometres (Leelőssy et al. 2014). Their chemical composition is essentially linked to their origin: during the extraction and processing of minerals, for instance, dust containing silica is often released in the air, which can result in permanent health damage (e.g. diseases such as silicosis) (Mlika, Adigun, and Bhutta 2022). Likewise, metallic dusts (e.g. lead and cadmium) from smelters cause harm to fauna and flora, as do chemical dusts created in agriculture (e.g. pesticides), vegetable dusts (e.g. wood, cotton), moulds and spores (WHO n.d.). Other air pollutants are created by the reaction of sunlight with atmospheric compounds, leading to photochemical smog (sunlight reacting with volatile organic compounds and nitrogen oxides).

The iron and steel industry is known to be a major contributor to air pollution. In the past, burning coal for steel production produced vast amounts of black coal dust, resulting in fine particles covering entire cities and regions. Such particles often contain silica and various metals; those with a diameter below 2.5

µm can easily enter the alveoli (i.e. the gas-exchange regions of the lungs) (Figure 2) and cause serious harm (Su, Ding, and Zhuang 2020). Early types of filtering devices based on cyclone techniques only removed larger particles, thus leaving smaller fractions uncaptured (‘The Steel City and a Brief History of Dust Collection’ 2019). In the Western world, with the rise of environmental regulation and technical advances from the mid-20th century onward, it became possible to significantly decreased dust emissions as long as steel companies were prepared to make the necessary investments. The decrease was mainly enabled by improved filtering techniques, including shaker bag collectors, reverse air or pulse jet baghouses, and (later) cartridge collectors (‘The Steel City and a Brief History of Dust Collection’ 2019). However, even today many iron and steel plants continue to rely on coal fuel (the ‘dirtiest’ but also cheapest option available), with little concern for the environmental effects and the availability of cleaner techniques (‘Do We Really Need Coal to Make Steel?’ 2020). As well as producing greenhouse gases like nitrous oxide (N_2O), steelmaking generates other airborne contaminants (Koponen et al. 1980). Metals like cobalt, lead and chromium are released into the atmosphere (Nurul, Shamsul, and Noor Hassim 2016) when not filtered appropriately, and so are concentrations of polycyclic aromatic hydrocarbons (PAHs), iron oxides and sulphur dioxide (SO_2) (Bala and Tabaku 2010). Historically as well as in the present, respirable particle emissions from the iron and steel industry pose risks to human health (Valenti et al. 2016).

4. Esch-sur-Alzette and the Belval metallurgical complex: a brief historical overview

Measured per capita, Luxembourg was the world’s foremost iron and steel producer throughout the 20th century (for global production trends, see (Hemmer 1953)). While all of the Minett’s towns were severely affected by air pollution, Esch had the specific characteristic of being surrounded by iron and steel complexes. The local historians Assa and Pagliarini have rightfully noted that ‘Esch and its factories formed one unity; their evolution occurred along parallel lines’ (Assa and Pagliarini 1998). From the moment the first blast furnace in Esch started operating, the urban infrastructure of this ‘mushroom town’ (Knebeler and Scuto 2010) indeed developed almost exclusively to serve the metallurgical industry. An engraving from the 1920s shows that iron and steel complexes could be found to the southwest of Esch (the *Terres Rouges* plant, established in 1872), to the east (the Esch-Schifflange plant, 1871) and to the west (the Belval plant, 1911). Other polluting factories, such as a cement factory, a brewery and a coal-fired power plant (which supplied electricity to the steelworks) were also present in the town throughout much of industrial era (for a general overview, see (Scuto 1993)). A scale model of the Minett, made for the 1937 World’s Fair in Paris, clearly illustrates how Esch was ‘embraced’ by factories and, consequently, by industrial smoke and dust (on Luxembourg’s participation at the 1937 World’s Fair, see (Millim 2014)).

Figure 3: Engraving by G. Peltier showing the three iron and steel complexes surrounding Esch: Esch-Schifflange (left), Terres Rouges (centre) and Belval (below right). Source: *(Knebeler and Scuto 2010)*.

Figure 4: Scale model of Esch exhibited at the 1937 World's Fair in Paris, showing the three iron and steel complexes surrounding Esch: Esch-Schifflange (bottom), Terres Rouges (top left) and Belval (top right). Source: Archives Nationales de Luxembourg, Fonds ARBED, File AES-U1-54.

Contrary to present-day popular belief (e.g. (Moia 1998; Logelin-Simon 2006; Back-Hoffmann 2021)), the air pollution caused by the metallurgical complexes of the Minett was not uncontested by the local population. From the early 20th century onwards, inhabitants voiced their criticism about dusty factory smoke through articles in newspapers and magazines, readers' letters and interventions in parliament and municipal council meetings. This paragraph will shortly explore a few criticisms concerning the Belval complex from the period when the plant was established. A more in-depth analysis of complaints from inhabitants about air pollution in the entire Minett region is offered in a separate article, which focuses on the period of the *Trente Glorieuses* (c. 1945-1965) and conceptualises the role of newspapers, politicians and scientists as 'mediators' in the public debate on air pollution (Van de Maele forthcoming).

From around 1900, about three decades after the dawn of industrialisation in Luxembourg, the country's press began to report on the so-called 'plagues' of dust and smoke (in German: *Staubplage, Rauchplage*). These 'plagues' were reported to soil clothes and buildings, irritate lungs and eyes, and affect animal and plant life. As early as 1899, for example, the newspaper *Luxemburger Wort* published a remarkably lucid analysis of air pollution in industrial regions (for the history of this newspaper, see (*150 Jahre Und Kein Bisschen Alt: 150 Joër Wort, 1848-1998* 1998)). Despite its conservative, pro-business profile, the newspaper offered a remarkably modern-sounding, holistic approach towards both human and non-human nature, thus offering a rather bleak picture of the industrial era:

> No more than fifteen years ago, everyone took for granted that a chimney should smoke, and every stranger who came to a factory town looked in admiration at the myriad of smoking chimneys […]. Even those who were directly harmed by the smoke considered this to be so inevitable that they hardly ever complained about it, and when they did, they were told – with a shrug – that nothing could be changed. And so for decades, the vegetation in industrial areas withered, entire forests had to be cut down prematurely […], and soot particles and poisonous gases inhaled by human lungs slowly but surely planted the germ of many a deadly disease.[3]

Already around this time, awareness about the detrimental side effects of metallurgical dust also existed on a governmental level. This becomes evident in a 1910 investigation undertaken at the instigation of the Luxembourg Prime Minister in response to the construction request for the Belval plant submitted by the German corporation Gelsenkirchener Bergwerks AG. At that time, the corporation had recently acquired a plot of forested land just to the west of Esch. Conceived as a vertically integrated plant, the Belval complex was to include two blast furnaces (later supplemented with a third), steel converters, a sintering plant and rolling mills. In a letter to three engineers (who were most likely state-employed), the Prime Minister expressed reservations about the addition of yet another metallurgical factory in the Minett: '[…] [It] is necessary to investigate all [technical] methods to curtail noise […] and dust […].'[4] Consistent with the legal framework of the time (Parmentier 2008), the Prime Minister's reservations were mainly driven by a desire to minimise damage to the private property of adjacent landowners; a true 'environmental' awareness (as can be seen embryonically in the very *avant-garde* 1899 newspaper article mentioned above) was not yet part of the politicians' intellectual *habitus*. In response, the experts offered a highly ambiguous evaluation of the prospective plant. Although they acknowledged the existence of a pollution burden, this burden was considered to be a 'normal' phenomenon in industrial regions:

> The inconveniences these kinds of establishment can have for certain [neighbouring] owners in terms oftheir enjoyment of the tranquillity and pure air of the countryside are not to be denied. These inconveniences stem from the vapours and dust that detract from the purity of the air, as well as the noise caused by the machines and the movement of workers […]. Such inconveniences are common in industrial

---

[3] 'Die Rauchplage und ihre Beseitigung', *Luxemburger Wort für Wahrheit und Recht*, 13 May 1899, p. 1.

[4] Archives Nationales de Luxembourg, *Ministère de la Justice* Collection, File J-090-00804, Letter from the Prime Minister, March 1910.

> centres; everyone experiences them in Esch, Differdange, Rodange, Dudelange, etc. Yet they cannot be invoked as a reason to refuse authorisation.[5]

This discourse was typical of the professional ethos of late-19th and early-20th-century engineers, who typically saw themselves as expert organisers, both within and beyond their technical sphere of competence (on the role of engineers in Luxembourg society, see (Glesener and Wilhelm 2009)). As such, the engineers also gave advice about the *moral* economy in which the steelmaking was to take place: in their view, the imperative of economic development outweighed the time-honoured right of neighbouring inhabitants to 'pure' air. The engineers' opinion proved to be decisive: shortly after the governmental enquiry, a green light was given to the Gelsenkirchener Bergwerks AG, and in 1911 the two blast furnaces of the iron and steel complex (which would later become known as Belval) were fired up. Soon enough, newspapers would again voice concerns about the increasing dust problem in Esch. In April 1914, just three months before the start of the First World War, another conservative newspaper, the *Obermosel-Zeitung*, for instance voiced concerns about the worsening 'smoke plague':

> If any town far and wide suffers greatly from the smoke plague, it is Esch. No fewer than 30 to 40 blast furnace chimneys in the town's immediate vicinity spew their contents on the Minett's capital day and night. If a resourceful person were to come up with an effective invention against the smoke plague, he should certainly file a patent in Esch. It is only a small consolation for the people of Esch to hear that the smoke plague is also known to be a great evil in other places, and that in many cases desperate efforts are being made to combat it […].[6]

Still in 1914, the Esch-based architect Paul Flesch declared in a municipal council meeting that Esch should keep a corridor without industry in the north, to allow the town's 30 000 inhabitants 'to breathe some clean air for at least a few days a year'. With his call for the incorporation of public hygiene principles in urban planning policies, Flesch sought to pair economic development with a more decent quality of life for the town's inhabitants (Scuto 2005). As it turned out, the north of Esch would effectively remain industry-free – even though the 'right to produce' (and, consequently, the right to pollute) of the existing iron and steel complexes would never be called into question by local and national authorities. Using a term from German historiography (e.g. (Geissler 2016)), it can be said that these authorities consistently affirmed the *Ortsüblichkeit* ('geographical appropriateness') of industry. This principle was further underpinned by the powerful political and economic position of the ARBED company (*Aciéries réunies de Burbach-Eich-Dudelange*), which took over ownership of the Belval complex in 1919 (Knebeler and Scuto 2010). In this constellation, the plant would thrive: between 1911 and 1997, a total of almost 80 million tonnes of iron was produced at the Belval site – or about one third of Luxembourg's total production during that period. Belval's corresponding steel production amounted to approximately 75 million tonnes – a number that is continuing to grow to this day, as the steelworks are still in operation even after the closure of the last blast furnace in 1997 (for production numbers, see (Knebeler and Scuto 2010)).

---

[5] Archives Nationales de Luxembourg, *Ministère de la Justice* Collection, File J-090-00804, Letter from Noppeney, Haardt and Jaans, 6 May 1910.

[6] 'Chronik aus dem Erzbassin', *Obermosel-Zeitung,* 10 April 1914, p. 2.

5. Steelmaking techniques in Luxembourg

As steelmaking techniques developed, the degree of dust pollution resulting from the steel industry in the Minett varied over time. Initially, at the onset of industrialisation in the 1870s, Luxembourg's plants used the Bessemer process, the first industrial method to produce steel from molten pig iron, which was invented in the 1850s and complemented in 1864 by the Siemens-Martin process (Knebeler 2011; Metz 1972). Impurities were removed by oxidation, i.e. by blowing air through the molten iron. However, as the iron ore found in the Minett is high in phosphorous, only low-quality steel could be manufactured (Knebeler 2011). This changed with the modification of the Bessemer process by Sidney Gilchrist Thomas, who used dolomite, which is basic, as the converter lining (instead of packed sand, which is acidic), thereby removing the phosphorous from the steel into the slag (Knebeler 2011; Metz 1972). In Luxembourg, the first Thomas steel was produced in 1886; the process was complemented by electric arc furnace (EAF) steelmaking after 1900. From the 1950s onwards, pure oxygen processes like the LD-AC process (*Linz-Donawitz – ARBED – Centre national de recherches métallurgiques de Liège*) were predominantly used to produce higher quality steel with fewer impurities compared to Thomas steel (Metz 1972; Knebeler 2011).[7]

The LD-AC method was reported to generate less dust ('Compte-rendu Chambre des Députés, volume 1 - 1975-1976' 1976). In a 1972 article, Metz signalled that the steelworks in the Minett used two principal techniques to perform dedusting: electro-static purification with negatively charged dust particles collected on a positive collection plate, and wet processes which retained dust with water (Metz 1972). The efficiency of these processes seems to have been limited, however. According to Hoffmann (1974), who based her findings on an undisclosed source, dust pollution levels in Esch regularly exceeded 1.8 g/m^3/day, which was well beyond the West German environmental limits of the time (0.42 g/m^2/day for urban zones and 0.85 g/m^2/day for industrial zones) (Hoffmann 1974). Yet even though she was a critic of the steel industry, Hoffmann stressed that 90% of the air pollution in Luxembourg resulted from households (oil heating) and traffic, and not from the iron and steel works – a position in line with the steel industry itself (Metz 1972). A number of measurements undertaken between the mid-1950s and the early 1970s, which we discuss below (section 8.2), indicate that this estimation was most likely too flattering for the metallurgical industry, and that the airborne iron oxide load caused by steel processes was probably much higher.

By the time Metz and Hoffmann published their articles, governmental awareness about the environmental impact of the iron and steel industry was on the rise. This rise was spurred by multiple factors, including decades of public complaints, the international emergence of the 'modern' environmental movement in the second half of the 1960s, and the ongoing scientific research on metallurgical pollution conducted in the transnational framework of the European Coal and Steel Community (Van de Maele forthcoming). As a result, in the mid-1970s, the demand for the 'polluter pays' principle became stronger, eventually requiring Luxembourg's industry to minimise its environmental impact ((Hoffmann 1974); on the worldwide rise of environmental policy-making during the 1970s, see for instance (Jarrige and Le Roux 2017; Buelens 2022; Uekötter 2020)). The iron oxide content in the Minett's atmosphere was addressed in a 1976 parliamentary debate, when an MP stated that the 'red clouds' of iron oxide were predominantly a result of Thomas steel production, while the LD-AC process (which was usually operated with pre-installed filters) was reported to generate less dust ('Compte-rendu

---

[7] Other techniques replacing Thomas steel were the Lance Bubbling Equilibrium (LBE), ARBED Ladle Treatment (ALT) or LD Kaldo, and Rotovert. (Metz, Knebeler)

Chambre des Députés, volume 1 - 1975-1976' 1976). In Luxembourg, the last Thomas steel was produced in 1977 (Metz 1972; Hoffmann 1974).

6. The simulation process

The calculation of our dust simulation model – estimating dust concentrations from 1911 to 1997 – was performed at our request by Inspyro, a technical consultancy company assisting metallurgical enterprises.[8] The various sources (and their limitations) used as input data for the Gaussian model are explained below.

There are several ways to create the atmospheric dispersion models needed to understand and predict air pollution. Gaussian, Lagrangian, Eulerian and computational fluid dynamics models all offer possibilities,[9] differing in their mathematical complexity and field of application. All models require the input of parameters like meteorological data, emission parameters and terrain information. To develop such a model, an interdisciplinary approach including disciplines like meteorology, chemistry and physics is often the best option. The output data is usually plotted on maps indicating areas of higher/lower air pollution concentrations and therefore higher/lower health risk.

In our case, a basic Gaussian dispersion model, as shown by Leelőssy et al. – known for its fast response time in the application of long-term average loads for distances between 1 and 100 km – was used (Leelőssy et al. 2014). The equation used (1), based on the Gaussian plume model for atmospheric dispersion modelling, is shown below.

$$\bar{c}\,(x, y, z) = \frac{Q}{2\pi\sigma_y\sigma_z\bar{u}} \exp\left(\frac{-y^2}{2\sigma_y^2}\right)\left(\exp\left(\frac{-(z-h)^2}{2\sigma_z^2}\right) + \exp\left(\frac{-(z+h)^2}{2\sigma_z^2}\right)\right) \qquad (1)$$

A detailed formula derivation can be found in Stockie's 'Mathematics of atmospheric dispersion modelling'. Stockie defines dispersion as the 'combination of turbulent diffusion and advection by the wind' (Stockie 2011). Atmospheric contaminant concentrations can therefore be explained using the advection-diffusion equation. Using the simplified model of a Gaussian plume, it is assumed that air contaminants come unidirectionally (given the wind direction) from one point source – here one chimney of the Belval steel plant –, as outlined in Figure 5.

---

[8] The metallurgical experts (civil engineers) who performed the calculation were Sander Arnout, Yannick Cryns and Cem Tekin.

[9] Gaussian models assume a normal statistical distribution (parabolic behaviour near the origin of the coordinates, "bell curve") and are typically used for buoyant air pollution plumes. This approach historically dominated dispersion models with a number of simplifications to be taken in the advection-diffusion equations. Lagrangian models use trajectories – calculated using ordinary differential equations – of air pollutants determined by e.g. wind field, buoyancy and turbulence (Leelőssy et al. 2014). Pollutant particles are followed in time and space along their trajectories. Eulerian models use a fixed coordinate frame providing a spatiotemporal evolution of pollutant concentration at each time step and point in the grid. For environmental and health protection measures, computational fluid dynamics models are often used in complex geometry where a fine grid resolution is required to calculate turbulence effects (Lagrangian and Eulerian: coarse grid). This model does account for flow velocities and turbulence in a complex 3D terrain, unlike the Gaussian approach (Tripathi et al. 2018). For more information and a detailed comparison see Leelőssy (Leelőssy et al. 2014).

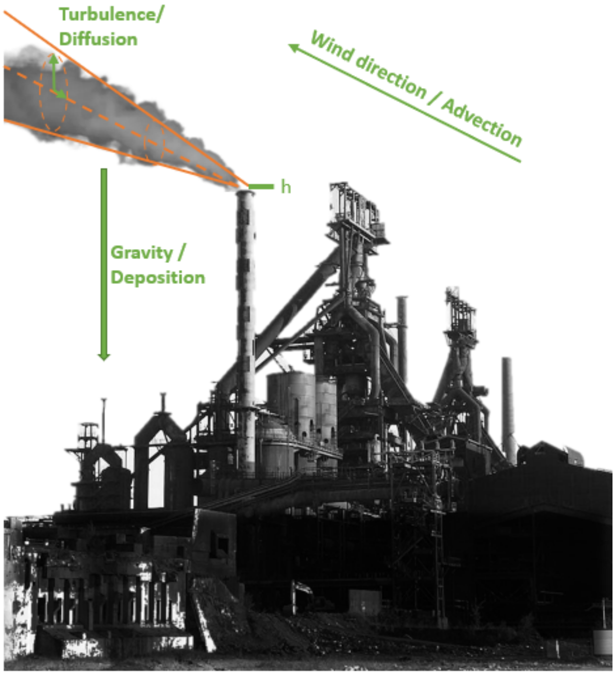
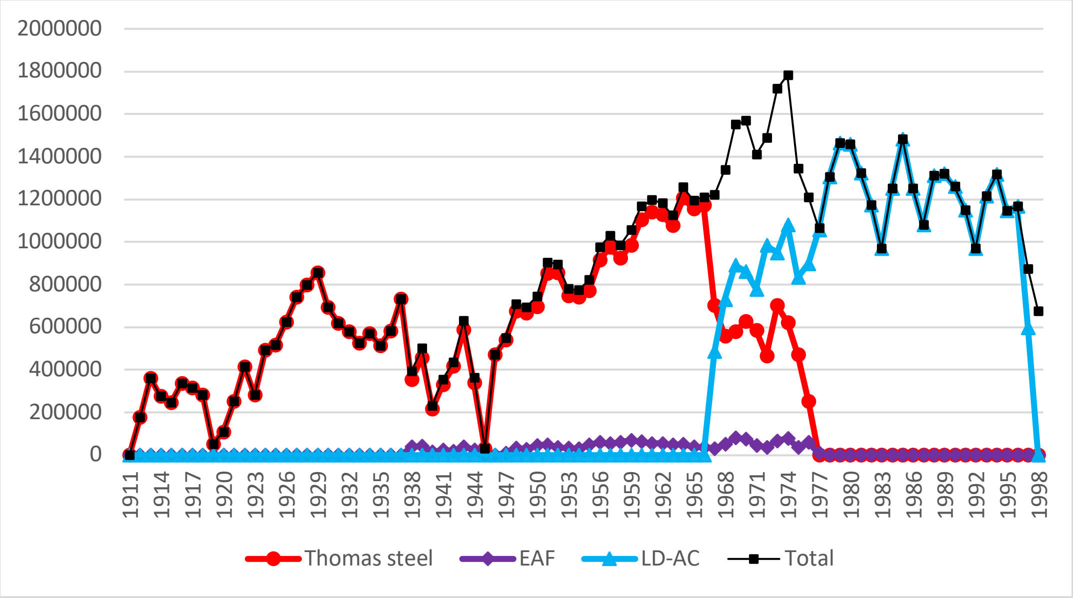


Figure 5: Schematic demonstration of the three main contributors to dust transportation: advection, diffusion and deposition at the Belval steel plant (photo taken in 2003) indicating the release height h. Photo modified by the authors. Source of the photo: www.agora.lu/

In formula (1) $\bar{c}$ is defined as the time-averaged concentration at a given position, Q represents the constant emission rate [kg/s] and (x,y,z) stand for the wind directions (downwind, crosswind, vertical direction) with the standard deviations σ_y and σ_z . $\bar{u}$ is the time-averaged wind speed at contaminant release height h (see Figure 5). Moreover, a homogeneous steady-state flow at the point source (0;0;h) is assumed (Leelőssy et al. 2014).

For the Belval case only one chimney of the steel plant was assumed to be the point source. To calculate the amount of dust produced per month, the various steelmaking processes were categorised; in relation to ironmaking, sintering and crushing of iron ore were the main processes contributing to dust production. Consequently, the average amounts of dust produced by the three steelmaking processes were used to calculate the amount of dust generated during the production of raw iron. For electric steel production (using an electric arc furnace (EAF)), the average amount of dust was multiplied by the amount of steel produced. The same approach was applied to Thomas steel and LD-AC steel. Annual steel production data at the Belval plant was retrieved from a table reproduced in a study by Knebeler and Scuto (Knebeler and Scuto 2010). This table includes the production data of steel in tonnes per steelmaking technique (Thomas, LD-AC, EAF, total steel) per year, shown in Graph 1. As the simulation output shows, the dust concentrations are calculated as averages per month. (The annual production values turn out to be a limitation, since monthly data would generate a more exact output of the model.) The hypothetical average dust values generated per process (blast furnace, sintering furnace, EAF and LD-AC) were taken from Schueneman, as summarised in Table 1 (Schueneman 1963). Schueneman offered typical values for steel plants in the US in 1963; as such, our hypothetical average dust values offer only an estimation of dust-producing steel processes in Luxembourg from 1911 to 1997.



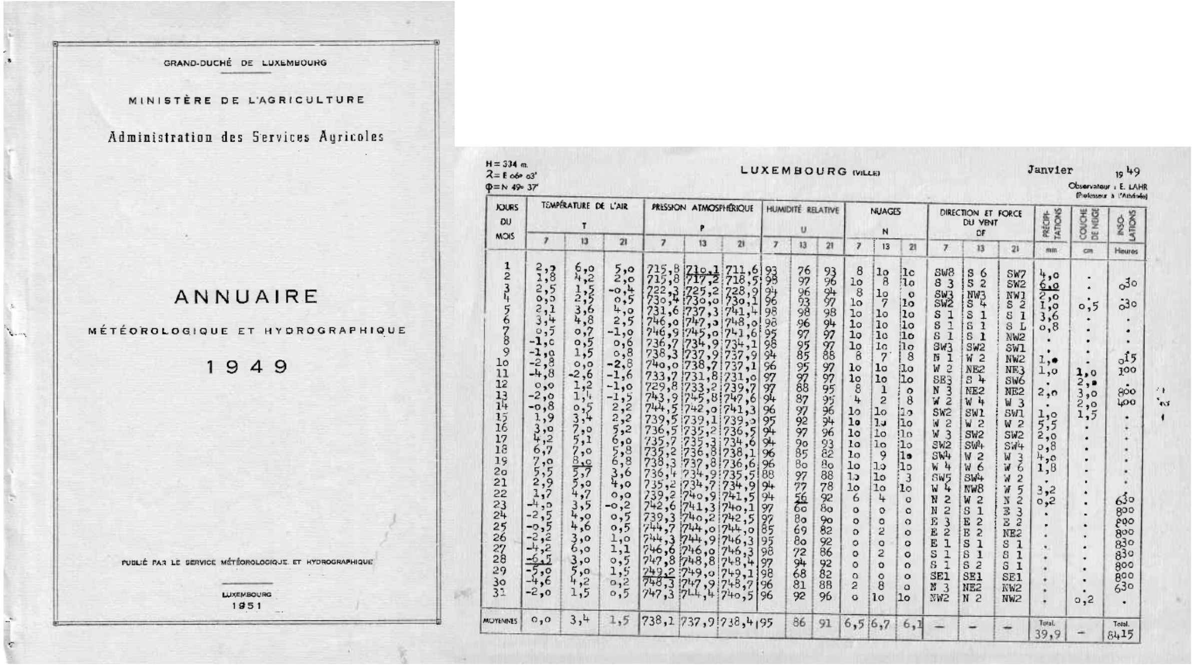
Graph 1: Steel production at the Belval plant from 1911 to 1998, in t per technique

The hypothetical installation date of filters (with varying efficiencies) was likewise taken from Schueneman (Schueneman 1963) (Table 1). Based on this hypothesis, steel production techniques in Belval after 1963 were assumed to have state-of-the-art filters installed. Again, the installation dates are typical estimates for US steel plants until 1963, which will not correspond 1:1 with our case study on Luxembourg. We used these assumptions because of the lack of literature on the historical presence and efficiency of filter systems in the European iron and steel industry. Table 1 summarises the filters installed per technique (based on Schueneman); the amount of dust caught by a filter was subtracted from the total amount produced. Afterwards the calculated amount of dust in t/year was converted into g/sec.

Technique	Average dust production [g/t]	Installation year	Filter	Efficiency [%]
Blast furnace	90718.5	1911	Preliminary	97
		1948	High efficiency	99.50
		1976	State of the art	99.90
Sintering furnace	9071.85	1931	Centrifugal separators	90
EAF	48080.79	1938	No filters	-
LD-AC	18143.7	1967	State of the art	99.50

Table 1: Average dust produced and filters installed per technique and year, adapted from Schueneman (Schueneman 1963)

To determine the wind directions and velocities per year and month, weather data from the yearly meteorological reports for Luxembourg City (about 20 km from Esch) ('Annuaire météorologique et hydrographique' 1949) were used and converted from the Beaufort scale to m/s. Since weather data for the years 1911-1949 were not available in digitised form, an average wind profile consisting of speed and direction – taken from the 10th and 20th day of each month – was created and used for all years, based on the data for the 1949-1997 period (see Figure 6).



GRAND-DUCHÉ DE LUXEMBOURG

MINISTÈRE DE L'AGRICULTURE

Administration des Services Agricoles

ANNUAIRE

MÉTÉOROLOGIQUE ET HYDROGRAPHIQUE

1949

PUBLIÉ PAR LE SERVICE MÉTÉOROLOGIQUE ET HYDROGRAPHIQUE

LUXEMBOURG
1951

H = 334 m.
λ = E 06° 03'
φ = N 49° 37'

LUXEMBOURG (VILLE)

Janvier 1949

Observateur : E. LAHR
(Professeur à l'Athénée)

JOURS DU MOIS	TEMPÉRATURE DE L'AIR T			PRESSION ATMOSPHÉRIQUE P			HUMIDITÉ RELATIVE U			NUAGES N			DIRECTION ET FORCE DU VENT DF			PRÉCIPITATIONS	COUCHE DE NEIGE	INSOLATIONS
	7	13	21	7	13	21	7	13	21	7	13	21	7	13	21	mm	cm	Heures
1	2,0	6,0	5,0	715,8	710,1	711,6	93	76	93	8	10	10	SW8	S 6	SW7	4,0	.	
2	1,8	4,2	2,0	715,8	717,2	718,5	98	97	96	10	8	10	S 3	S 2	SW2	6,0	.	0^{30}
3	2,5	1,5	-0,4	722,3	725,2	728,9	94	96	94	8	10	0	SW3	NW3	NW1	2,0	.	
4	0,0	2,5	0,5	730,4	730,0	730,1	96	93	97	10	7	10	SW2	S 4	S 2	1,0	0,5	0^{30}
5	2,1	3,6	4,0	731,6	737,3	741,4	98	98	98	10	10	10	S 1	S 1	S 1	3,6	.	.
6	3,4	4,8	2,5	746,0	747,0	748,0	98	96	94	10	10	10	S 1	S 1	S L	0,8	.	.
7	0,5	0,7	-1,0	746,9	745,0	741,6	95	97	97	10	10	10	S 1	S 1	NW2	.	.	.
8	-1,0	0,5	0,6	736,7	734,9	734,1	98	95	97	10	10	10	SW3	SW2	SW1	.	.	.
9	-1,0	1,5	0,8	738,3	737,9	737,9	94	85	88	8	7	8	N 1	W 2	NW2	1,●	.	0^{15}
10	-2,8	0,0	-2,8	740,0	738,7	737,1	96	95	97	10	10	10	W 2	NE2	NE3	1,0	1,0	1^{00}
11	-4,8	-2,6	-1,6	733,7	731,8	731,0	97	97	97	10	10	10	SE3	S 4	SW6	.	2,●	.
12	0,0	1,2	-1,0	729,8	733,2	739,7	97	88	95	8	1	0	N 3	NE2	NE2	2,0	3,0	8^{00}
13	-2,0	1,4	-1,5	743,9	745,8	747,6	94	87	95	4	2	8	W 2	W 4	W 3	.	2,0	4^{00}
14	-0,8	0,5	2,2	744,5	742,0	741,3	96	97	96	10	10	10	SW2	SW1	SW1	1,0	1,5	.
15	1,9	3,4	2,2	739,5	739,1	739,0	95	92	94	10	10	10	W 2	W 2	W 2	5,5	.	.
16	3,0	7,0	5,2	736,5	735,2	736,5	94	97	96	10	10	10	W 3	SW2	SW2	2,0	.	.
17	4,2	5,1	6,0	735,7	735,3	734,6	94	90	93	10	10	10	SW2	SW4	SW4	0,8	.	.
18	6,7	7,0	5,8	735,2	736,8	738,1	96	85	82	10	9	1●	SW4	W 2	W 3	4,0	.	.
19	7,0	8,0	6,8	738,3	737,8	736,6	96	80	80	10	10	10	W 4	W 6	W 6	1,8	.	.
20	5,5	5,7	3,6	736,4	734,9	735,5	88	97	88	10	10	3	SW5	SW4	W 2	.	.	.
21	2,9	5,0	4,0	735,2	734,7	734,9	94	77	78	10	10	10	W 4	NW8	W 5	3,2	.	.
22	1,7	4,7	0,0	739,2	740,9	741,5	94	56	92	6	4	0	N 2	W 2	N 2	0,2	.	6^{30}
23	-4,0	3,5	-0,2	742,6	741,3	740,1	97	60	80	0	0	0	N 2	S 1	E 3	.	.	8^{00}
24	-2,5	4,0	0,5	739,3	740,2	742,5	97	80	90	0	0	0	E 3	E 2	E 2	.	.	6^{00}
25	-0,5	4,6	0,5	744,7	744,0	744,0	85	69	82	0	2	0	E 2	E 2	NE2	.	.	8^{00}
26	-2,2	3,0	1,0	744,3	744,9	746,3	95	80	92	0	0	0	E 1	S 1	S 1	.	.	8^{30}
27	-4,2	6,0	1,1	746,6	746,0	746,3	98	72	86	0	2	0	S 1	S 1	S 1	.	.	8^{30}
28	-6,5	3,0	0,5	747,8	748,8	748,4	97	94	92	0	0	0	S 1	S 2	S 1	.	.	8^{00}
29	-5,0	5,0	1,5	749,2	749,0	749,1	98	68	82	0	0	0	SE1	SE1	SE1	.	.	8^{00}
30	-4,6	4,2	0,2	748,3	747,9	748,7	96	81	88	2	8	0	N 3	NE2	NW2	.	.	6^{30}
31	-2,0	1,5	0,5	747,3	744,4	740,5	96	92	96	0	10	10	NW2	N 2	NW2	.	0,2	.
MOYENNES	0,0	3,4	1,5	738,1	737,9	738,4	95	86	91	6,5	6,7	6,1	–	–	–	Total. 39,9	–	Total. 84^{15}

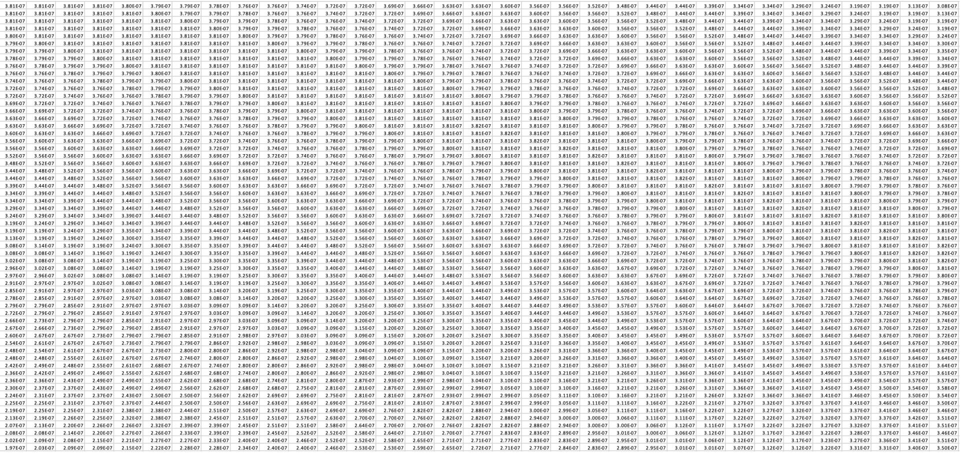
Figures 6 a and b from left to right: (a) Cover of and (b) excerpt from the yearly meteorological report for Luxembourg (in this case showing the data for Luxembourg City, including the average wind speeds and directions in January 1949). Source: (*Annuaire Météorologique et Hydrographique 1949* 1951)

Using this input data, the dust concentration (g/m^3) of all 15m x 15m x 1m cells – in a 2.5 km radius around the Belval plant – was calculated, using a MATLAB ('MATLAB - MathWorks' 2022) script based on equation (1). The code was generated by metallurgical experts[10] and is therefore not included here. Nevertheless, we show the input, output (csv files) and further processing (data visualization) of the MATLAB script. MATLAB is a computing environment and a programming language used by engineers and scientists, aiding with data analytics, linear algebra, signal processing, etc. MATLAB language is a matrix-based language that includes many specialized libraries to solve engineering and scientific problems. Although it also includes plotting capabilities, we have decided to use Python for the data visualization, to have more flexibility.

[Explore data included in Jupyter notebook, matrix below serves as example for illustration – not an actual image]

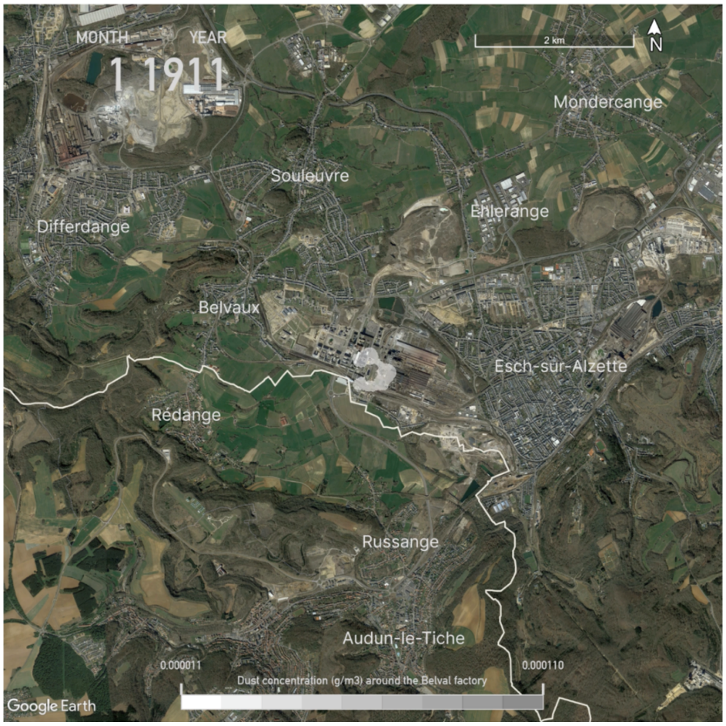
---

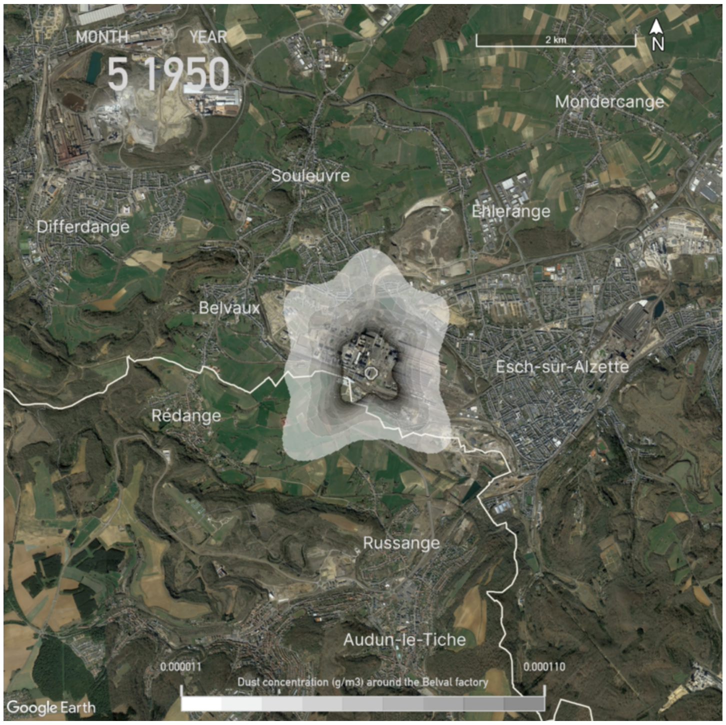
[10] The metallurgical experts (civil engineers) who performed the calculation were Sander Arnout, Yannick Cryns and Cem Tekin.



3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07	3.44E-07	3.44E-07	3.39E-07	3.34E-07	3.34E-07	3.29E-07	3.24E-07	3.19E-07	3.19E-07	3.13E-07	3.08E-07
3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07	3.44E-07	3.44E-07	3.39E-07	3.34E-07	3.34E-07	3.29E-07	3.24E-07	3.19E-07	3.19E-07	3.13E-07
3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07	3.44E-07	3.44E-07	3.39E-07	3.34E-07	3.34E-07	3.29E-07	3.24E-07	3.19E-07	3.19E-07
3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07	3.44E-07	3.44E-07	3.39E-07	3.34E-07	3.34E-07	3.29E-07	3.24E-07	3.19E-07
3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07	3.44E-07	3.44E-07	3.39E-07	3.34E-07	3.34E-07	3.29E-07	3.24E-07
3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07	3.44E-07	3.44E-07	3.39E-07	3.34E-07	3.34E-07	3.30E-07
3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07	3.44E-07	3.44E-07	3.39E-07	3.34E-07	3.35E-07
3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07	3.44E-07	3.44E-07	3.39E-07	3.34E-07
3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07	3.44E-07	3.44E-07	3.39E-07
3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07	3.44E-07	3.44E-07
3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07	3.44E-07
3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07	3.48E-07
3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07	3.52E-07
3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07	3.56E-07
3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07	3.56E-07
3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07	3.60E-07
3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07	3.63E-07
3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07	3.63E-07
3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07	3.66E-07
3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07	3.69E-07
3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07	3.72E-07
3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07	3.72E-07
3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07	3.74E-07
3.44E-07	3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07	3.76E-07
3.39E-07	3.44E-07	3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07	3.76E-07
3.34E-07	3.39E-07	3.44E-07	3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07	3.78E-07
3.34E-07	3.34E-07	3.39E-07	3.44E-07	3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07	3.79E-07
3.29E-07	3.34E-07	3.34E-07	3.39E-07	3.44E-07	3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07	3.79E-07
3.24E-07	3.29E-07	3.34E-07	3.34E-07	3.39E-07	3.44E-07	3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07	3.80E-07
3.19E-07	3.24E-07	3.29E-07	3.34E-07	3.34E-07	3.39E-07	3.44E-07	3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07	3.81E-07
3.19E-07	3.19E-07	3.24E-07	3.29E-07	3.35E-07	3.34E-07	3.39E-07	3.44E-07	3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07	3.81E-07
3.13E-07	3.19E-07	3.19E-07	3.24E-07	3.30E-07	3.35E-07	3.35E-07	3.39E-07	3.44E-07	3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07	3.81E-07
3.08E-07	3.14E-07	3.19E-07	3.19E-07	3.24E-07	3.30E-07	3.35E-07	3.35E-07	3.39E-07	3.44E-07	3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.81E-07	3.81E-07	3.82E-07
3.08E-07	3.08E-07	3.14E-07	3.19E-07	3.19E-07	3.24E-07	3.30E-07	3.35E-07	3.35E-07	3.39E-07	3.44E-07	3.44E-07	3.48E-07	3.52E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.82E-07	3.82E-07
3.02E-07	3.08E-07	3.08E-07	3.14E-07	3.19E-07	3.19E-07	3.25E-07	3.30E-07	3.35E-07	3.35E-07	3.39E-07	3.44E-07	3.44E-07	3.48E-07	3.53E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.66E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07	3.82E-07
2.96E-07	3.02E-07	3.08E-07	3.08E-07	3.14E-07	3.19E-07	3.19E-07	3.25E-07	3.30E-07	3.35E-07	3.35E-07	3.40E-07	3.44E-07	3.44E-07	3.48E-07	3.53E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.67E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07	3.81E-07
2.97E-07	2.96E-07	3.02E-07	3.08E-07	3.08E-07	3.14E-07	3.19E-07	3.19E-07	3.25E-07	3.30E-07	3.35E-07	3.35E-07	3.40E-07	3.44E-07	3.44E-07	3.48E-07	3.53E-07	3.56E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.67E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07	3.80E-07
2.91E-07	2.97E-07	2.97E-07	3.02E-07	3.08E-07	3.08E-07	3.14E-07	3.19E-07	3.19E-07	3.25E-07	3.30E-07	3.35E-07	3.35E-07	3.40E-07	3.44E-07	3.44E-07	3.49E-07	3.53E-07	3.57E-07	3.56E-07	3.60E-07	3.63E-07	3.63E-07	3.67E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07	3.79E-07
2.85E-07	2.91E-07	2.97E-07	2.97E-07	3.03E-07	3.08E-07	3.08E-07	3.14E-07	3.20E-07	3.19E-07	3.25E-07	3.30E-07	3.35E-07	3.35E-07	3.40E-07	3.44E-07	3.44E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.60E-07	3.64E-07	3.63E-07	3.67E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07	3.79E-07
2.78E-07	2.85E-07	2.91E-07	2.97E-07	2.97E-07	3.03E-07	3.08E-07	3.08E-07	3.14E-07	3.20E-07	3.20E-07	3.25E-07	3.30E-07	3.35E-07	3.35E-07	3.40E-07	3.44E-07	3.44E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.60E-07	3.64E-07	3.64E-07	3.67E-07	3.69E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07	3.78E-07
2.79E-07	2.79E-07	2.85E-07	2.91E-07	2.97E-07	2.97E-07	3.03E-07	3.09E-07	3.09E-07	3.14E-07	3.20E-07	3.20E-07	3.25E-07	3.30E-07	3.35E-07	3.35E-07	3.40E-07	3.44E-07	3.44E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.60E-07	3.64E-07	3.64E-07	3.67E-07	3.70E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07	3.76E-07
2.72E-07	2.79E-07	2.79E-07	2.85E-07	2.91E-07	2.97E-07	2.97E-07	3.03E-07	3.09E-07	3.09E-07	3.14E-07	3.20E-07	3.20E-07	3.25E-07	3.30E-07	3.35E-07	3.35E-07	3.40E-07	3.44E-07	3.44E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.60E-07	3.64E-07	3.64E-07	3.67E-07	3.70E-07	3.72E-07	3.72E-07	3.74E-07	3.76E-07
2.66E-07	2.73E-07	2.79E-07	2.79E-07	2.85E-07	2.91E-07	2.97E-07	2.97E-07	3.03E-07	3.09E-07	3.09E-07	3.14E-07	3.20E-07	3.20E-07	3.25E-07	3.30E-07	3.35E-07	3.35E-07	3.40E-07	3.45E-07	3.44E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.60E-07	3.64E-07	3.64E-07	3.67E-07	3.70E-07	3.72E-07	3.72E-07	3.74E-07
2.67E-07	2.66E-07	2.73E-07	2.79E-07	2.79E-07	2.85E-07	2.91E-07	2.97E-07	2.97E-07	3.03E-07	3.09E-07	3.09E-07	3.15E-07	3.20E-07	3.20E-07	3.25E-07	3.30E-07	3.35E-07	3.35E-07	3.40E-07	3.45E-07	3.45E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.60E-07	3.64E-07	3.64E-07	3.67E-07	3.70E-07	3.72E-07	3.72E-07
2.60E-07	2.67E-07	2.67E-07	2.73E-07	2.79E-07	2.79E-07	2.85E-07	2.91E-07	2.98E-07	2.97E-07	3.03E-07	3.09E-07	3.09E-07	3.15E-07	3.20E-07	3.20E-07	3.25E-07	3.30E-07	3.35E-07	3.35E-07	3.40E-07	3.45E-07	3.45E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.60E-07	3.64E-07	3.64E-07	3.67E-07	3.70E-07	3.72E-07
2.54E-07	2.61E-07	2.67E-07	2.67E-07	2.73E-07	2.79E-07	2.79E-07	2.86E-07	2.92E-07	2.98E-07	2.98E-07	3.03E-07	3.09E-07	3.09E-07	3.15E-07	3.20E-07	3.20E-07	3.25E-07	3.31E-07	3.36E-07	3.35E-07	3.40E-07	3.45E-07	3.45E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.61E-07	3.64E-07	3.64E-07	3.67E-07	3.70E-07
2.48E-07	2.54E-07	2.61E-07	2.67E-07	2.67E-07	2.73E-07	2.80E-07	2.80E-07	2.86E-07	2.92E-07	2.98E-07	2.98E-07	3.04E-07	3.09E-07	3.09E-07	3.15E-07	3.20E-07	3.20E-07	3.26E-07	3.31E-07	3.36E-07	3.36E-07	3.40E-07	3.45E-07	3.45E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.61E-07	3.64E-07	3.64E-07	3.67E-07
2.48E-07	2.48E-07	2.55E-07	2.61E-07	2.67E-07	2.67E-07	2.74E-07	2.80E-07	2.80E-07	2.86E-07	2.92E-07	2.98E-07	2.98E-07	3.04E-07	3.10E-07	3.09E-07	3.15E-07	3.21E-07	3.20E-07	3.26E-07	3.31E-07	3.36E-07	3.36E-07	3.40E-07	3.45E-07	3.45E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.61E-07	3.64E-07	3.64E-07
2.42E-07	2.49E-07	2.48E-07	2.55E-07	2.61E-07	2.68E-07	2.67E-07	2.74E-07	2.80E-07	2.80E-07	2.86E-07	2.92E-07	2.98E-07	2.98E-07	3.04E-07	3.10E-07	3.10E-07	3.15E-07	3.21E-07	3.21E-07	3.26E-07	3.31E-07	3.36E-07	3.36E-07	3.41E-07	3.45E-07	3.45E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.61E-07	3.64E-07
2.36E-07	2.42E-07	2.49E-07	2.49E-07	2.55E-07	2.62E-07	2.68E-07	2.68E-07	2.74E-07	2.80E-07	2.80E-07	2.86E-07	2.92E-07	2.98E-07	2.98E-07	3.04E-07	3.10E-07	3.10E-07	3.15E-07	3.21E-07	3.21E-07	3.26E-07	3.31E-07	3.36E-07	3.36E-07	3.41E-07	3.45E-07	3.45E-07	3.49E-07	3.53E-07	3.57E-07	3.57E-07	3.61E-07
2.36E-07	2.36E-07	2.43E-07	2.49E-07	2.49E-07	2.55E-07	2.62E-07	2.68E-07	2.68E-07	2.74E-07	2.81E-07	2.80E-07	2.87E-07	2.93E-07	2.99E-07	2.98E-07	3.04E-07	3.10E-07	3.10E-07	3.16E-07	3.21E-07	3.21E-07	3.26E-07	3.31E-07	3.36E-07	3.36E-07	3.41E-07	3.45E-07	3.45E-07	3.49E-07	3.54E-07	3.57E-07	3.57E-07
2.30E-07	2.37E-07	2.37E-07	2.43E-07	2.49E-07	2.49E-07	2.56E-07	2.62E-07	2.68E-07	2.68E-07	2.75E-07	2.81E-07	2.81E-07	2.87E-07	2.93E-07	2.99E-07	2.99E-07	3.05E-07	3.10E-07	3.10E-07	3.16E-07	3.21E-07	3.21E-07	3.26E-07	3.31E-07	3.36E-07	3.36E-07	3.41E-07	3.45E-07	3.45E-07	3.50E-07	3.54E-07	3.58E-07
2.24E-07	2.31E-07	2.37E-07	2.37E-07	2.43E-07	2.50E-07	2.50E-07	2.56E-07	2.62E-07	2.69E-07	2.69E-07	2.75E-07	2.81E-07	2.81E-07	2.87E-07	2.93E-07	2.99E-07	2.99E-07	3.05E-07	3.11E-07	3.10E-07	3.16E-07	3.21E-07	3.21E-07	3.26E-07	3.32E-07	3.36E-07	3.36E-07	3.41E-07	3.46E-07	3.45E-07	3.50E-07	3.54E-07
2.25E-07	2.25E-07	2.31E-07	2.37E-07	2.37E-07	2.44E-07	2.50E-07	2.50E-07	2.56E-07	2.63E-07	2.69E-07	2.69E-07	2.75E-07	2.81E-07	2.81E-07	2.87E-07	2.93E-07	2.99E-07	2.99E-07	3.05E-07	3.11E-07	3.11E-07	3.16E-07	3.22E-07	3.21E-07	3.27E-07	3.32E-07	3.37E-07	3.37E-07	3.41E-07	3.46E-07	3.46E-07	3.50E-07
2.19E-07	2.25E-07	2.25E-07	2.31E-07	2.38E-07	2.38E-07	2.44E-07	2.51E-07	2.50E-07	2.57E-07	2.63E-07	2.69E-07	2.69E-07	2.76E-07	2.82E-07	2.82E-07	2.88E-07	2.94E-07	3.00E-07	2.99E-07	3.05E-07	3.11E-07	3.11E-07	3.16E-07	3.22E-07	3.22E-07	3.27E-07	3.32E-07	3.37E-07	3.37E-07	3.41E-07	3.46E-07	3.46E-07
2.13E-07	2.19E-07	2.26E-07	2.25E-07	2.32E-07	2.38E-07	2.38E-07	2.45E-07	2.51E-07	2.51E-07	2.57E-07	2.63E-07	2.70E-07	2.70E-07	2.76E-07	2.82E-07	2.82E-07	2.88E-07	2.94E-07	3.00E-07	3.00E-07	3.06E-07	3.11E-07	3.11E-07	3.17E-07	3.22E-07	3.22E-07	3.27E-07	3.32E-07	3.37E-07	3.37E-07	3.42E-07	3.51E-07
2.07E-07	2.13E-07	2.20E-07	2.26E-07	2.26E-07	2.32E-07	2.39E-07	2.39E-07	2.45E-07	2.51E-07	2.51E-07	2.58E-07	2.64E-07	2.70E-07	2.70E-07	2.76E-07	2.82E-07	2.82E-07	2.88E-07	2.94E-07	3.00E-07	3.00E-07	3.06E-07	3.12E-07	3.11E-07	3.17E-07	3.22E-07	3.22E-07	3.27E-07	3.32E-07	3.37E-07	3.41E-07	3.51E-07
2.08E-07	2.08E-07	2.14E-07	2.20E-07	2.27E-07	2.26E-07	2.33E-07	2.39E-07	2.39E-07	2.45E-07	2.52E-07	2.52E-07	2.58E-07	2.64E-07	2.71E-07	2.70E-07	2.77E-07	2.83E-07	2.83E-07	2.89E-07	2.95E-07	3.01E-07	3.00E-07	3.06E-07	3.12E-07	3.12E-07	3.17E-07	3.23E-07	3.22E-07	3.28E-07	3.37E-07	3.46E-07	3.46E-07
2.02E-07	2.09E-07	2.08E-07	2.15E-07	2.21E-07	2.27E-07	2.27E-07	2.33E-07	2.40E-07	2.40E-07	2.46E-07	2.52E-07	2.52E-07	2.58E-07	2.65E-07	2.71E-07	2.71E-07	2.77E-07	2.83E-07	2.83E-07	2.89E-07	2.95E-07	3.01E-07	3.01E-07	3.06E-07	3.12E-07	3.12E-07	3.17E-07	3.23E-07	3.27E-07	3.36E-07	3.41E-07	3.51E-07
1.97E-07	2.03E-07	2.09E-07	2.09E-07	2.15E-07	2.22E-07	2.28E-07	2.28E-07	2.34E-07	2.40E-07	2.40E-07	2.46E-07	2.53E-07	2.53E-07	2.59E-07	2.65E-07	2.72E-07	2.71E-07	2.77E-07	2.84E-07	2.83E-07	2.89E-07	2.95E-07	3.01E-07	3.01E-07	3.07E-07	3.12E-07	3.12E-07	3.22E-07	3.31E-07	3.31E-07	3.40E-07	3.50E-07

The visualisation of the model was generated using a Python in Jupyter notebook, from 1044 generated matrices of 302x302 values, saved in separate csv files. Each file corresponds to a month and a year between January 1911 and December 1997. The data is not geocoded; instead, the reference point (or source point) is one hypothetical chimney of the steelworks, located at the centre of the matrix, from which concentration values are calculated. The visualisation process of the data is explained in section 7.





[Code to generate the visualizations included in the Jupyter notebook, images above are just for illustration]

It can be said in advance for our study that the calculated concentration values are most likely to be underestimations, since only one emission source in southern Luxembourg was considered (thus eliminating many others). As such, our analysis serves as an experiment and can be taken as a basis for further research (e.g. calculating the dust concentrations from all industrial plants in the entire Minett region). Other limitations and improvements of our model are discussed in sections 8.3 and 8.4..

7. Data visualisation

As introduced above, the data visualisation was generated using a Jupyter notebook with Python for a multi-layered approach, which allows us to share not only the narrative layer but also the data and the code. The data was not geocoded, since explicitly assigning coordinates to each concentration point in each dataset would have increased the volume of data considerably. Instead, the data was kept in the form of multiple matrices and positioned on the map using Google Earth to calculate the distance covered by the density matrix (see Figure 8).

Measure the circumference or area of a circle on the ground

Radius: 2.26 Kilometers

Area: 15,991,797.00 Square Meters

Circumference: 14.20 Kilometers

☑ Mouse Navigation

Figure 8: Screenshot of Google Earth[11] demonstrating how the distance from the Belval plant is measured.

Initially, the option of using a historical map (see Figure 9) was considered. This was discarded in favour of facilitating more accurate calculation of distances and generating a view with the steelworks at the centre.

---

[11] https://www.google.com/intl/de/earth/

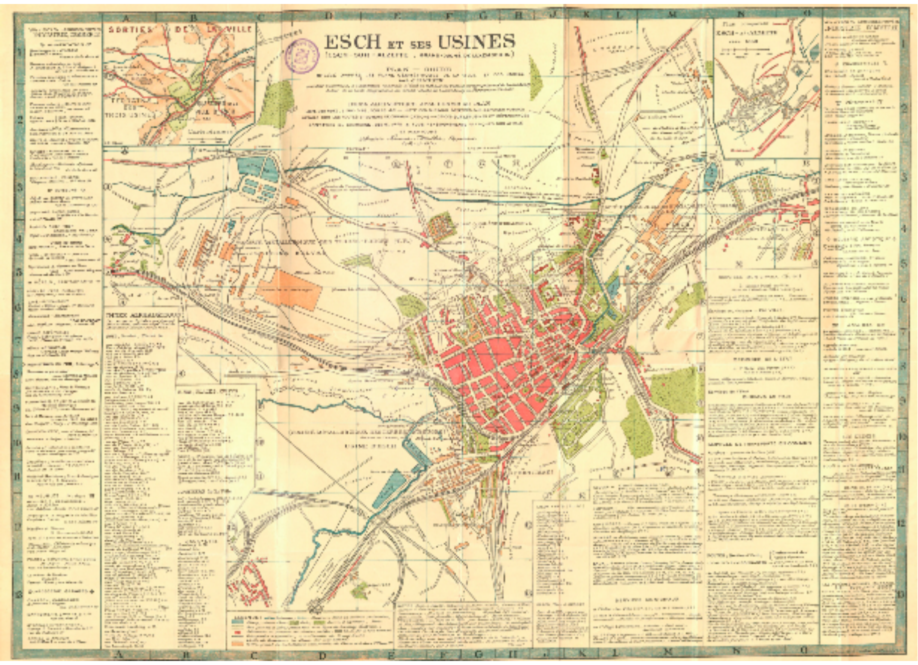


Figure 9: 1923 map of Esch. Source: (Hansen 1923)

The dust concentrations around the Belval plant (in g/m^3) were visualised using a contour map with ten predefined levels. We used the extent parameter to plot the concentrations on the map, and a customised sequential colour palette to indicate the presence of dust and emphasise a higher concentration near the plant.

[Code to generate the contour map using the extent parameter and colour palette included in the Jupyter notebook]

Contour maps were generated for every month and year between January 1911 and December 1997 in English, French and German. For this purpose, the names of the main localities were translated; using Figma, they were added to the base maps extracted from Google Earth (see Figures 10 a and b).

Figures 10 a and b from left to right: (a) English and French version of the map, (b) German version of the map.

The visualisations were subsequently animated to show evolution over time.

[Code to generate the animation included in the Jupyter notebook]

To visualise the impact of pollution in the inhabited quarters of Esch, we repeated the process with a zoomed-in map, thus highlighting the immediate vicinity of the Belval plant.

[Code to calculate the contour plot, the extent, and the video in the Jupyter notebook]

The videos were integrated into a dedicated page of the *Minett Stories* virtual exhibition, which was launched in May 2022. Produced by the Centre for Contemporary and Digital History (University of Luxembourg) in connection with *Esch-sur-Alzette – European Capital of Culture 2022*, the exhibition offers twenty-two stories, outlining various transformations of landscapes, places and people, from the region's industrialisation in the late 19th century to the crisis of the 1970s and the subsequent deindustrialisation. Instead of focusing on the often-told stories about the growth and success of the steel industry, *Minett Stories* explores everyday life in the Minett, with attention to questions such as environmental pollution.

Our visualisation helps us to answer the initially formulated questions about the geographical distribution of dust pollution. By using various historical maps of Esch (see Figures 11 a,b,c) and comparing them to the data visualisation for the same year, we can analyse which neighbourhoods were most affected over time. July has been chosen, as this is the month when the pollution typically reached the farthest in the town of Esch.

Figure 11 a,b,c: from left to right: (a) 1956 map of Esch, (b) simulation of dust pollution around the Belval plant in July 1956 showing the entire region, and (c) simulation of dust pollution for the same month and year with a focus on the centre of Esch. Source: (Buchler et al. 2020)

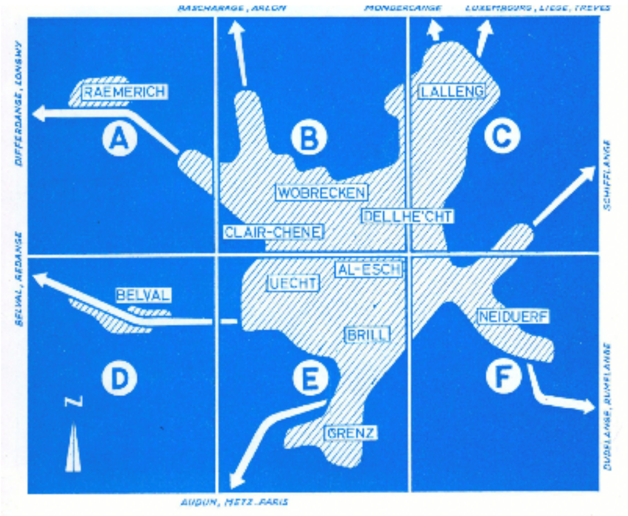


Figure 12: 1958 map of Esch, showing the location of various urban quarters. Source: ('Esch-Sur-Alzette Nouveau Plan 1958. Répertoire Des Rues et Places' 1958)

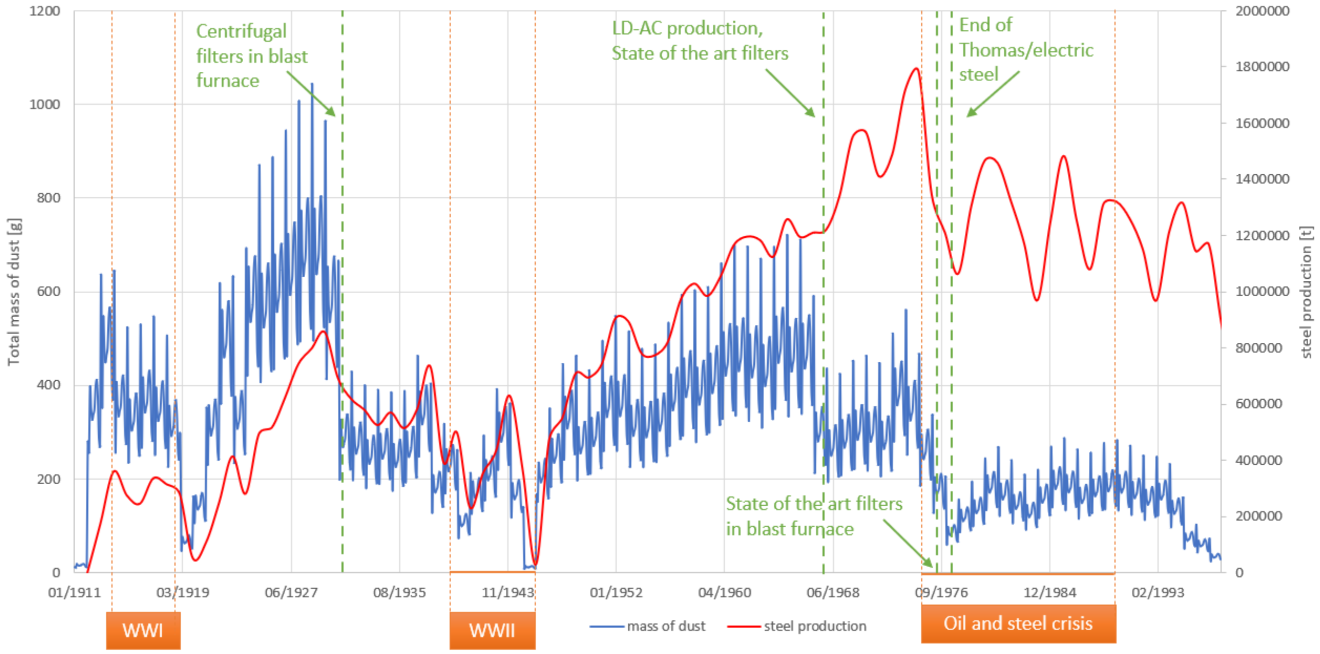
It is clear, for example, that high dust levels were often prevalent in the centre of Esch; unsurprisingly, the western part of town was the most severely affected. Using a 1958 map of Esch as a reference*,* the affected areas primarily included the neighbourhoods of Uecht, Clair-Chêne, Al-Esch and Wobrecken. Moreover, from the outset, Belval's dust pollution was a cross-border phenomenon, with the French villages of Rédange, Russange and Audun-le-Tiche (the latter itself the site of an ironworks from the 1880s to the 1960s) sharing the pollution burden with small Luxembourgish locations such as Soleuvre, Ehlerange and Belvaux.

8. Final analysis

To conclude, we perform a critical evaluation and interpretation of the dust simulation. We explore the plausibility of the calculated values, comparing them to historical and current air pollution legislation, as well as to historical measurements. Limitations and possible ways to optimise the model are also discussed.

8.1 Validation of the model

Graph 2 shows the summarised (total) amount of dust calculated by year *versus* steel production in tonnes, with dust in blue and production data in red. Important events like the (hypothetical) installation of filters are indicated. The graph further indicates a number of historical dates (such as the installation of filters, changing techniques, World War I and II, and the steel crisis from the second half of the 1970s onward), which all correspond to visible changes in the production and/or dust curves. This effect is for instance highly visible after 1965, with steel production increasing and the dust curve decreasing, owing to more intensive use of oxygen-based steelmaking and the hypothetical installation of 'state-of-the-art' filters (see section 6 and Table 1).



Graph 2: Total amount of dust calculated in blue; steel production numbers in red; important events indicated

As well as checking the plausibility of the data, it is necessary to look at the actual meaning of the calculated dust values (with the highest calculated value being about 3 mg/m^3 in October 1929). To understand the effects of such values on human health, we will compare them to present-day regulatory limit values.

Directives and measures regarding air quality were first introduced by the European Economic Community in the 1970s. In subsequent decades, air pollution resulting from industry decreased significantly (European Environment Agency EEA 2013). In Luxembourg, the first legislation regarding air pollution was the 'Loi du 21 juin 1976 relative à la lutte contre la pollution de l'atmosphère'. No limit values were laid down in this Act. In the more recent context of EU law, Directive 2008/50/EC (European Union 2008) postulates limit values for different atmospheric pollutants to protect human health, including values for PM$_{2.5}$ (25 µg/m^3) and metals like Cd (5 ng/m^3) (European Parliament. Directorate General for Parliamentary Research Services. 2021). Such fine-grained values not only indicate recent improvements in measuring methods; they are also indicative of an increase in awareness among legislators in recent decades.

Our model, however, does not distinguish between the various dust fractions or different pollutants. When comparing the simulation model with the current EU legislation values, the calculated maximum (3 mg/m^3) clearly exceeds the limit related to particulate matter concentration (25 µm/m^3 for PM$_{2.5}$). Furthermore, the EU directive from 2008 is under revision following the recent proposal of more stringent values by the World Health Organization (WHO). In 2021, the WHO indeed put forward a maximum value of 5 µg/m^3 for long-term PM$_{2.5}$ exposure in order to protect human health (World Health Organization 2021).

Specific legislative values have also been established for industrial producers. Relevant EU laws include the 2001 'Directive on the limitation of emissions of certain pollutants into the air from large combustion plants (2001/80/EC)', followed by the 2010 'Directive on industry emissions (2010/75/EU)', which focuses particularly on industrial pollution. Summarising the various sector-specific directives, a dust limit concentration of 30 mg/Nm^3 for coal and other solid fuels (lower values apply for gaseous fuels and higher thermal nominal power) is valid in all EU Member States. The N in the concentration unit emphasises an important point to consider. Since dust values depend on pressure and temperature, a comparison with other values requires standardised conditions (N: 25°C, 1 atm). This is another limitation of our model, as temperature and pressure data were not included. Compared to the simulated dust concentrations, the limit value is much higher.

Decision (EU) 2017/1757 of the Council of the European Union defines separate dust limit values for the iron and steel industry: 50 mg/m^3 for sintering, 15-20 mg/m^3 for pelletising plants, 10 mg/m^3 for hot blast stoves in blast furnaces (> 2.5t/h), 30 mg/m^3 for oxygen steelmaking and casting (> 2.5 t/h), and 5 (new ovens) to -15 (existing ovens) mg/m^3 for steel production and casting using the EAF process (> 2.5 t/h) (European Union 2017). In comparison, our calculated dust amount is below these limits. These values, however, define steelmaking processes of today, not of the past; moreover, they are process specific.

Regarding workplace exposure limits measured in time-weighted averages, pollutant-specific levels can be found in legislation; in Luxembourg they are regulated by the 'Réglement Grand-Ducal du 14 novembre 2016 concernant la protection de la sécurité et de la santé des salariés contre les risques liés à des agents chimiques sur le lieu de travail' (Le Gouvernement du Grand-Duché de Luxembourg 2016).

In general, it can be said that it is difficult to compare the simulated dust values with legislative limit values. First, our values do not include information on external conditions like temperature and pressure that is necessary to perform an exact comparison. Second, and even more important, the calculated dust concentrations reveal neither the composition of dust and the type of pollutants, nor the dust generating process in steelmaking. Therefore, limit values for specific dust fractions or processes cannot be applied. Moreover, when looking at dust concentrations in legislation, it clearly makes a difference whether one looks for general or sector-specific directives, as illustrated above.

Nevertheless, there is evidence for the health impact of air pollution coming from the iron and steel industry in Luxembourg and other countries. Of course, other factors such as climate, noise, vibrations and other types of pollution (e.g. soil, water) had an influence, too. Regarding health effects analysed in Luxembourg, Molitor and Mosinger published a research study on the effects of atmospheric dusts on human health in 1960 (Molitor and Mosinger 1960). They found evidence for the connection of industry dust with pulmonary diseases like silicosis and siderosis. In 1975 there was a medical symposium (industrial medicine) taking place in Luxembourg, collecting several studies on the effects of industry emissions on respiratory diseases and proving the impact of dust emission on health (Comission of the European Communities ECSC 1976). A special focus of researchers was on the occupational health of steel workers, which were exposed to even higher amounts of pollution than the citizens around the steel plant (occupational medicine (Gochfeld 2005)). Other studies demonstrating the impact of atmospheric steel pollution on human health were conducted e.g. in England, studying mortality effects (Beach and Hanlon 2018), in China, looking at heavy metals - representing a carcinogenic risk above certain limits - (Qing 2015) or in Albania (Bala and Tabaku 2010) and Brazil (Valenti et al. 2016) studying chronic obstructive pulmonary disease.

In combination with time witness reports ('Compte-rendu Chambre des Députés, volume 1 - 1975-1976' 1976) and dust measurements conducted in other countries (see above) this information supports the assumption that the dust pollution in the area around Luxembourgish steel plants exceeded most legislative limit values introduced until today.

8.2. Comparison of the model results with air pollution measurements from the late 1930s to the early 1970s

To compare the simulated dust levels with real conditions, it is important to scrutinise historical documents on dust measurements in the Minett region. We were able to retrieve data on four measurements performed between the late 1930s and the early 1970s: 1939 (Schifflange, undertaken by the ARBED steel company) ('Étude sur le dépoussiérage des gaz des hauts-fourneaux à l'usine d'Arbed-Esch' 1939), 1956-1957 (Differdange, undertaken by the government) (Molitor and Barthel 1958), 1966-1969 (Esch, undertaken by the municipality) ('Retombées de poussières: Rapport sur les mesures de retombées en poussière effectuées de 1966 à 1969 à Esch-sur-Alzette' 1972), and 1970-1971 (Belval, undertaken in an academic context) ('Les Composes du Fluor dans la pollution atmospherique' 1971). In all cases, dust was collected for a specified period through passive samplers (known as Bergerhoff devices). Both the motivation for the measurements and the methodological framework for the subsequent interpretation differed over time, as we will outline below.

In 1939, ARBED undertook dust measurements on the premises of its Schifflange iron and steel complex, located directly to the east of Esch (about 2.5 km from the Belval plant, which was owned by the same company) ('Étude sur le dépoussiérage des gaz des hauts-fourneaux à l'usine d'Arbed-Esch' 1939). Investigated parameters included the amount of dust generated by the blast furnaces, particle sizes and the iron (Fe) content of the samples. This analysis was conducted for economic reasons only, with the aim of studying the potential for reusing blast furnace gas after the removal of dust particles. The analysis pointed out that very small dust particles (ranging from 0 to 10 µm) constituted about 25% of the total emitted dust mass; as noted in section 3, such particles can penetrate the respiratory system and pose a threat to human health. In addition, around 48% of all dust released was reported to be made up of Fe: a very high iron content that was not unusual in the context of the Thomas steelmaking process. Even though the 1939 corporate investigation concluded that recycling and filtering techniques would be too expensive to implement, it is clear – in retrospect – that the installation of a dedicated filter could have resulted in significant health gains for both the local population and factory employees.

About two decades later, the scientists Molitor and Barthel conducted a government investigation of the dust concentration in the vicinity of the Differdange iron and steel complex (1956, published in 1958) (Molitor and Barthel 1958). Six passive samplers were placed for a period of about three months in a 2 km radius around the plant, allowing the amount of dust in g per m² per month to be calculated and a compositional analysis to be performed. Molitor and Barthel claimed that insoluble particles such as Fe_2O_3, SiO_2, CaO and SO_4 did not have harmful effects, regardless of particles sizes. However, pulmonary diseases like siderosis (caused by iron dust) were identified decades before this study was published, indicating that this claim was incorrect even at the time of publication (Doig and Mclaughlin 1936). Dust fractions below 0.03 mm (*inhalable* fraction, see section 3) were observed in the samples, including silica, which indeed poses a risk for respiratory diseases and other health issues.[12] The Differdange analysis further indicated that dust concentrations dropped rapidly as the distance from the factory increased, starting with 140 g/m²/month near the centre of pollution and falling to about 40 g/m²/month at a 2 km distance. With the decreasing dust mass, the iron content decreased as well, dropping from 75% to 38.4%

---

[12] Molitor and Mosinger even published a study themselves on industry dust resulting in silicosis and siderosis in 1960 (Molitor and Mosinger 1960).
Knowledge about respiratory diseases caused by dust was even present in ancient times. See: (Rosen 1993, 459–76)

(Molitor and Barthel 1958). These measurements once again show the high contribution of steelmaking to dust pollution levels in the Minett, especially near blast furnaces. However, the values measured by Molitor and Barthel cannot be compared directly to the simulation presented in our analysis, since the passive sampler was unable to capture the smallest fractions of dust (below 5 µm (Heyart 1958); in addition, the measurement units and production site are not the same. The measured values in 1956 nevertheless seem to be far higher than the simulated ones, which indicates that the dust concentrations in our simulation are probably underestimated.

In 1966-1969, Barthel was again involved in dust measurements, this time commissioned by the municipal administration of Esch ('Retombées de poussières: Rapport sur les mesures de retombées en poussière effectuées de 1966 à 1969 à Esch-sur-Alzette' 1972). Barthel's new observations were publicly released in 1972, a year marked by intense worldwide interest in environmental pollution problems – not least with the publication of the Club of Rome's report *Limits to Growth* (Buelens 2022). Compared to the Differdange analysis, the number of measurement points was increased to 52 (spread over the entire surface area of the town), resulting in 16 samplers per km^2. This allowed for a more fine-grained analysis of the accumulated dust generated by multiple iron and steel complexes (Belval, *Terres Rouges* and Schifflange). The measurements were subsequently compared to the West German limit values for atmospheric dust pollution, which were a maximum of 0.42 g/m^2/day in urban zones and 0.85 g/m^2/day in industrial zones ('Retombées de poussières: Rapport sur les mesures de retombées en poussière effectuées de 1966 à 1969 à Esch-sur-Alzette' 1972). Over half of the measurement points were reported to exceed the West German dust limits for *urban* regions (34 in 1966-1967, 35 in 1967-1968 and 28 in 1968-1969), while several points even showed values above the *industrial* limits (9, 6 and 7 respectively). When compared to the simulation model, the measured values are significantly higher (dimensions of measurements in g/day and for the simulation in mg/month). The dust composition showed an iron oxide content of about 65%, varying with distance, indicating a clear link with the steelworks. Even though Barthel (incorrectly) stressed that the dust concentrations had no significant effect on human health, he did call for technical improvements that could bring the pollution levels down to the West German norm for industrial zones.

Lastly, in 1971, Christiane Conter wrote a dissertation on atmospheric pollution, based on measurements performed in 1970-1971 as part of an academic research project at the Belval plant ('Les Composes du Fluor dans la pollution atmospherique' 1971). Conter's work clearly presented an environmental point of view, with a focus on various sources of pollution (steelworks, households and traffic). The dust samples at the Belval steel complex were specifically analysed with regard to fluorine compounds, known for being released during various industrial processes and for their toxic effects on vegetation and animals. Sampling stations were placed near various points in the production chain; the average dust concentration measured was 6 g/m^2/day, which exceeded not only the West German legislative limits but also the measurements undertaken by Barthel in the years before. Just like Molitor and Barthel (Differdange, 1956-1957), Conter analysed dust composition in terms of soluble and insoluble particles. However, her specific focus on fluorinated compounds yielded no concrete results, as possible trace amounts were superimposed by the dust matrix; additionally, the methods used (X-ray spectroscopy and electron microscopy) were not sensitive enough to detect fluorinated compounds ('Les Composes du Fluor dans la pollution atmospherique' 1971).

Despite the lack of comparability between the simulated values and the available measurements, our simulation allows us to isolate the dust pollution potentially generated by a single steelworks in the past and to understand how far it reached overtime. From this point, the analysis could be extended to other industrial plants and other types of pollution, which would probably bring us closer to the existing measurements. Furthermore, the lack of detail in historical measurements, which obviously cannot be measured again (e.g. a precise type of pollution created by a particular plant at a specific point in time), creates an opportunity for the use of modelling and simulation.

8.3. Limitations and lacunae of the dust simulation

Throughout this research, we have presented different decisions taken during the process of simulation, validation and visualization of the data. These decisions simplify and reduce the complexity of the problem, but they have been consciously made to serve a specific purpose. This was to simulate industrial dust pollution from a single factory in a given period, based on historical data; above, we have presented the results of these decisions. This simulation can change perspectives on the past, e.g. with looking at the seasonal changes of wind directions and therefore changes in dust distribution (which is often omitted in historical literature). In future steps, we could modify the model variables (e.g., using data from other sources), or include new variables and use a different model (see section 8.4). In any case, we have proven that data simulation based on historical data offers the option not only to recreate situations in the past, but also to isolate effects that would not be possible to analyse even with real measurements.

Nevertheless, we want to present a final analysis of the main limitations and gaps of our simulation, to which we have already alluded in previous sections. Such limitations become relevant if we seek to understand the total pollution to which the analysed area was exposed. First, not all dust emitting sources in the Belval region were considered; obviously, the steel industry was only one contributor to atmospheric pollution (albeit a very important one) (Metz 1972; Hoffmann 1974; 'Compte-rendu Chambre des Députés, volume 1 - 1975-1976' 1976). Pollution from households, industrial heating and traffic (motor vehicles and trains) accounted for some of the dust particles as well. Second, regarding the mathematical formula (1) used (presented in section 6), a few parameters must be analysed critically. The choice of a Gaussian model in itself involves simplifications. The model assumes that the contaminants come unidirectionally from one point source (one of several chimneys) at a constant emission rate *Q*, with the wind direction always aligned in one axis, considering the time frame of one month (Stockie 2011; Leelőssy et al. 2014). This assumption does not cover intramonth variations and only considers one metallurgical plant (from a total of three in Esch alone). The model is not suitable for low wind speeds, and in terms of meteorological data only averages for Luxembourg City (the 10th and 20th of each month), which is about 20 km from Belval, were used. In Gaussian plume models used in the calculation of atmospheric pollution, the specific features of the surrounding terrain are usually considered (Stockie 2011). As such, our formula could be extended to involve terrain information and a greater radius could be covered to better observe long-range effects. State-of-the-art Gaussian models like AERMOD, CTDM and ADMS try to extend the basic equation to make it more accurate and include more factors like a complex terrain (Leelőssy et al. 2014). Other dispersion models like the Lagrangian, Eulerian or CFD models are more suitable for different scales and applications.

Another factor influencing the results is our use of US data from 1963 (Schueneman 1963) on average dust values for different steelmaking techniques and typical filter installation dates. Steelmaking and filter use vary between individual steel plants. The production data used covered only annual values, so data on monthly production would make the simulation more precise. To better interpret the calculated dust values with regard to regulatory limit values and to determine a possible health effect, it would be useful to identify the dust fractions and pollutants contained therein. This is complicated, however, as historical dust samples have not been preserved and documents on measurements performed in the past (see 7.2) only study a small set of pollutants. Dust pollution values are usually standardised using temperature and pressure, which presents another limitation for interpretation. Lastly, the dust values identified using the model are technique specific (e.g. for Thomas steel) and do not relate to dust generation processes (e.g. sintering) used to set limit values for current EU regulations.

To summarise, the dust simulation of the Belval complex between 1911 and 1997 is just a starting point towards modelling the operation of a real-world process: the total generation and spread of pollution to which the inhabitants of Esch and the surrounding area were exposed. In Section 8.4, we offer a few possible ways of optimising the simulation in future research.

8.4. Optimisation and improvement of the model

Some of the lacunae mentioned in section 8.3 can be improved, thereby opening possibilities for further research. Other emission types like traffic, heating and other industrial plants could be considered, as well as the terrain around the point sources. The overlap of multiple dust generating sources and the resulting plumes, however, involves the risk of making the visualisation too complex and therefore confusing. As we state above, the model and used parameters were chosen on purpose, simulating an isolated phenomena.

The model area could be expanded, and temperature and pressure data could be considered. All these changes would increase the complexity of the mathematical formula. The meteorological data for the years 1911-1948 could be digitised to access additional weather data such as the wind directions and speeds for all months. Moreover, subdividing dust particles into different pollutants and fractions would enhance accuracy, while the generation of dust for each individual ironmaking and steelmaking process could be an interesting parameter to analyse further (especially with regard to the health impact of dust). However, the latter extension of the simulation is at the limit of what is practically feasible, since past dust samples (or contemporary analyses) from the respective processes are required for a precise investigation of the dust, its fractions and its origins. For several short periods, a composition analysis of the dust particles was performed; for other periods, the amount of dust created by different processes in the plant was measured, as explained in section 8.2. Using these data points is a possible way to extend the simulation, even if only for these specific years. Finally, it would be beneficial to look for average dust values produced by Luxembourg's steel industry, monthly production data, filter installation dates and filter efficiency rates. Such data is most likely present in the corporate archives of ARBED, which are unfortunately not completely accessible for researchers. The optimisation possibilities outlined above would take our study beyond its initial scope.

9. Conclusion

Air pollution has a long history: it is not only a problem of the past but also an issue affecting the present and the future. Especially in industrialised regions, atmospheric pollution is not only annoying; it can also be a dangerous companion of daily life. Dust particles in the respirable fraction pose the most worrisome threats to human health, as they can lead to lung diseases. In many parts of the world, the steel industry has contributed significantly to the total mass of airborne dust. Improvements in production techniques and the installation of filters, as well as legislative measures and economic crises, have influenced air pollution over the years.

Our analysis outlines the evaluation of dust production from a single plant in terms of plausibility and health effects. The simulated dust concentrations were compared to measurements and observations from the past, representing the 'scientifically proven reality' during a specific time period. For the simulation, a Gaussian plume model was used; the inserted parameters were analysed critically in terms of their limitations and possible improvements. Given the lack of availability of data regarding the dates of filter installations in Belval, secondary literature on the US case was used.

When the results are compared with dust measurements performed in the past, a clear deviation can be seen. Measurements from the period between the 1930s and the 1970s show that the dust concentration in the Minett was much higher than the simulation shows. Even limit values for industrial zones – taken from West Germany – were exceeded. Moreover, a high iron content and smaller particle fractions were identified. However, only a partial comparison with measured values is possible, since there are differences in terms of units, techniques and measurement area. One of the challenges of modelling and simulation is to find the right balance between complexity (and therefore precision) and interpretability. In this case, given the already added complexity of the interdisciplinary approach, we decided to use a simplified model, which serves as a precedent for creating new ways of generating historical data. At the same time, it allows us to present evidence (albeit on just a fraction of the dust to which inhabitants were exposed) and to understand the impact on health in a specific historical context.

The decision to visualise the pollution as coming from just one chimney also has benefits, as adding multiple dust sources might overload the visualisation (in the latter case, it would no longer be possible to discriminate between the origins of the dust). Furthermore, it has never been possible to measure the isolated dust concentration of a single plant using samplers, since there are always other dust sources. Hence, a simulation provides a suitable heuristic tool for generating this non-existent (historical) data. In sum, our model serves as a basis for studying atmospheric pollution released from one selected source and offers the potential to be extended by looking at other contributing factors.

10. Acknowledgements and funding

We would like to thank Sander Arnout, Yannick Cryns and Cem Tekin from the metallurgical consultancy company Inspyro (Leuven, Belgium), who built the initial simulation. Our thanks also go to all the REMIX and LuxTIME team members who took part in discussions about our research, as well to Sarah Cooper for her language proofreading.

This research has been funded by *Esch-sur-Alzette – European Capital of Culture 2022* and the University of Luxembourg Institute for Advanced Studies Audacity Programme for the Luxembourg Time Machine (LuxTIME).

11. References

150 Jahre Und Kein Bisschen Alt: 150 Joër Wort, 1848-1998. 1998. Luxembourg: Saint-Paul.

‘Annuaire météorologique et hydrographique’. 1949. Annuaire météorologique et hydrographique. Luxembourg: Service météorologique et hydrographique. https://www.agrimeteo.lu/Internet/AM/inetcntrLUX.nsf/cuhome.xsp?src=WY1BE977UP&p1=Q0BOB9JITQ&p3=9VXL1WMS7M.

Annuaire Météorologique et Hydrographique 1949. 1951. Luxembourg: Service météorologique et hydrographique.

Assa, Guy, and Luciano Pagliarini. 1998. *Esch-Uelzecht Gëscht an Haut*. Esch-sur-Alzette: Schortgen.

Back-Hoffmann, Jean. 2021. ‘Allgegenwärtige ARBED!’ *Forum*, no. 419: 42–45.

Bala, Silvana, and Afrim Tabaku. 2010. ‘Chronic Obstructive Pulmonary Disease in Iron-Steel and Ferrochrome Industry Workers’. *Central European Journal of Public Health* 18 (2): 93–98. https://doi.org/10.21101/cejph.a3548.

Banks, Jerry. 1999. ‘Introduction to Simulation’. In *Proceedings of the 1999 Winter Simulation Conference*, 7.

Beach, Brian, and W. Walker Hanlon. 2018. ‘Coal Smoke and Mortality in an Early Industrial Economy’. *The Economic Journal* 128 (615): 2652–75. https://doi.org/10.1111/ecoj.12522.

Brüggemeier, Franz-Josef, and Thomas Rommelspacher. 1992. *Blauer Himmel Über Der Ruhr: Geschichte Der Umwelt Im Ruhrgebiet, 1840-1990*. Essen: Klartext.

Buchler, Georges, Jean Goedert, Antoinette Lorang, Antoinette Reuter, Denis Scuto, Christof Weber, and capybarabooks. 2020. *Esch-sur-Alzette Guide historique et architectural*.

Buelens, Geert. 2022. *Wat We Toen al Wisten: De Vergeten Groene Geschiedenis van 1972*. Amsterdam: Querido Facto.

Butte, Werner, and Birger Heinzow. 2002. ‘Pollutants in House Dust as Indicators of Indoor Contamination’. *Reviews of Environmental Contamination and Toxicology* 175: 1–46.

Carson II, John S. 2005. ‘Introduction to Modeling and Simulation’. In *Proceedings of the 2005 Winter Simulation Conference*, 8.

Champion, Erik Malcolm. 2016. ‘Digital Humanities Is Text Heavy, Visualization Light, and Simulation Poor’. *Digital Scholarship in the Humanities*, November, fqw053. https://doi.org/10.1093/llc/fqw053.

Comission of the European Communities ECSC. 1976. ‘Medical Symposium: Researches on Chronic Respiratory Deseases - Industrial Medicine’. 18. Industrial Health and Medicine Series. Luxembourg.

‘Compte-rendu Chambre des Députés, volume 1 - 1975-1976’. 1976. 53. Luxembourg.

‘Do We Really Need Coal to Make Steel?’ 2020. *Wildsight* (blog). 1 June 2020. https://wildsight.ca/2020/06/01/do-we-really-need-steelmaking-coal/.

Doig, A.T., and A.I.G. Mclaughlin. 1936. ‘X RAY APPEARANCES OF THE LUNGS OF ELECTRIC ARC WELDERS’. *The Lancet* 227 (5875): 771–74. https://doi.org/10.1016/S0140-6736(00)56868-8.

‘Esch-Sur-Alzette Nouveau Plan 1958. Répertoire Des Rues et Places’. 1958.

‘Étude sur le dépoussiérage des gaz des hauts-fourneaux à l’usine d’Arbed-Esch’. 1939. Written by ARBED ingenieurs. Luxembourg. Bibliotheque Nationale Luxembourg.

European Environment Agency EEA. 2013. ‘Rechtsvorschriften zur Luftqualität in Europa’. https://www.eea.europa.eu/de/signale/signale-2013/artikel/rechtsvorschriften-zur-luftqualitaet-in-europa.

European Parliament. Directorate General for Parliamentary Research Services. 2021. *EU Policy on Air Quality : Implementation of Selected Legislation : European Implementation Assessment.* LU: Publications Office. https://data.europa.eu/doi/10.2861/808342.

European Union. 2008. *Richtlinie 2008/50/EG über Luftqualität und saubere Luft für Europa*. https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32008L0050&qid=1419250736800&from=DE.

———. 2017. *Beschluss (EU) 2017/1757 Des Rates*. Vol. 2017/1757. https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:32017D1757&from=FR.

Galison, Peter. 1996. 'Computer Simulations and the Trading Zone'. In *The Disunity of Science: Boundaries, Contexts, and Power*, edited by Peter Galison and David J. Stump, 118–57. Stanford University Press.

Geissler, Stephanie. 2016. *Wem Gehört Die Stadt? Umweltkonflikte Im Städischen Raum Zur Zeit Der Früh- Und Hochindustrialisierung in Aachen Und Duisburg*. Münster & New York: Waxmann.

Glesener, Jeanne E., and Frank Wilhelm. 2009. 'L'image de La Sidérurgie Dans Les Romans Des Luxembourgeois Francophones Joseph Leydenbach Willy Gilson et Nicolas Ries: Des Années Fastes à La Crise Des Années 1970'. In *Terres Rouges: Histoire de La Sidérurgie Luxembourgeoise*, 1:71–95. Terres Rouges: Histoire de La Sidérurgie Luxembourgeoise. Luxembourg: Archives Nationales.

Gochfeld, Michael. 2005. 'Chronologic History of Occupational Medicine': *Journal of Occupational and Environmental Medicine* 47 (2): 96–114. https://doi.org/10.1097/01.jom.0000152917.03649.0e.

Guldi, Jo, and David Armitage. 2014. *The History Manifesto*. Cambridge, United Kingdom: Cambridge University Press.

Han, Qiao, Yang Liu, Xixi Feng, Pan Mao, Ang Sun, Mingya Wang, and Mingshi Wang. 2021. 'Pollution Effect Assessment of Industrial Activities on Potentially Toxic Metal Distribution in Windowsill Dust and Surface Soil in Central China'. *Science of the Total Environment* 759: 11. https://doi.org/10.1016/j.scitotenv.2020.144023.

Hansen, J. 1923. 'Esch et Ses Usines'. Esch-sur-Alzette: Kremer-Müller.

Hemmer, Carlo. 1953. *L'économie Du Grand-Duché de Luxembourg: Deuxième Partie – La Production Secondaire (L'industrie Sidérurgique)*. Luxembourg: Joseph Beffort.

Hoffmann, Sylvie. 1974. 'Kritik in Punkto Umweltschutz - "Elo... Um Richtege Wé...?"' *Democratie 80*, 1974, N 74 edition, sec. Eine persönliche Meinung.

Jarrige, François, and Thomas Le Roux. 2017. *La Contamination Du Monde: Une Histoire Des Pollutions à l'âge Industriel*. Paris: Le Seuil.

Knebeler, Christophe. 2011. 'Wichtige technische Beiträge der ARBED zur Stahlherstellung'. *forum* 304 (ARBED): 36–38.

Knebeler, Christophe, and Denis Scuto. 2010. *Belval: Passé, présent et avenir d'un site luxembourgeois exceptionnel (1911-2011)*. Esch-sur-Alzette: Le Phare.

Koponen, M., T. Gustafsson, K. Kalliomäki, Pirkko-Liisa Kalliomäki, M. Moilanen, and L. Pyy. 1980. 'Dusts in a Steel-Making Plant'. *International Archives of Occupational and Environmental Health* 47 (1): 35–45. https://doi.org/10.1007/BF00378326.

Kumar, Dilip, and Deepak Kumar. 2018. 'Chapter 12 - Dust Control'. In *Sustainable Management of Coal Preparation*, edited by Dilip Kumar and Deepak Kumar, 265–78. Woodhead Publishing. https://doi.org/10.1016/B978-0-12-812632-5.00012-4.

Lake, M. W. 2014. 'Trends in Archaeological Simulation'. *Journal of Archaeological Method and Theory* 21 (2): 258–87. https://doi.org/10.1007/s10816-013-9188-1.

Le Gouvernement du Grand-Duché de Luxembourg. 2016. *Règlement grand-ducal du 14 novembre 2016 concernant la protection de la sécurité et de la santé des salariés contre les risques liés à des agents chimiques sur le lieu de travail.* Vol. A235. https://legilux.public.lu/eli/etat/leg/rgd/2016/11/14/n1/jo.

Lebherz, Daniel S., Christian Zeyen, Jan Hess, Ralph Bergmann, Ingo J. Timm, Thomas Burch, Vera Hildenbrandt, and Claudine Moulin. 2018. 'Towards Scientific Workflows and Computer Simulation as a Method in Digital Humanities'. https://doi.org/10.18420/INFDH2018-03.

Leelőssy, Ádám, Ferenc Molnár, Ferenc Izsák, Ágnes Havasi, István Lagzi, and Róbert Mészáros. 2014. 'Dispersion Modeling of Air Pollutants in the Atmosphere: A Review'. *Open Geosciences* 6 (3): 257–78. https://doi.org/10.2478/s13533-012-0188-6.

'Les Composes du Fluor dans la pollution atmospherique'. 1971. Doctoral Thesis. Luxembourg. 2048. Luxembourg Archives de l'Etat.

Logelin-Simon, Armand. 2006. 'Differdange: Le Décryptage d'une Âme'. *Nos Cahiers* 27 (3–4): 81–104.

Malm, Andreas. 2016. *Fossil Capital: The Rise of Steam Power and the Roots of Global Warming*. Brooklyn: Verso.

Massard-Guilbaud, Geneviève. 2007. 'Quelle Histoire Pour l'environnement?' *Responsabilité & Environnement*, no. 48: 30–36.

'MATLAB - MathWorks'. 2022. 2022. https://www.mathworks.com/products/matlab.html.

McCarty, Willard. 2016. 'Fictions of Possibility: Simulation for the Humanities from Its History in the Technosciences'. *Journal of Siberian Federal University. Humanities & Social Sciences* 9 (7): 1553–61. https://doi.org/10.17516/1997-1370-2016-9-7-1553-1561.

Metz, M. Paul. 1972. 'Recherche et dévelopment en sidérurgie'. In *Revue Technique Luxemburgeoise 1972*, 175–91.

Millim, Anne-Marie. 2014. 'Schooling the Gaze: Industry and Nation-Building in Luxembourgish Landscapewriting, 1900–1940'. *Journal of European Studies* 44 (2): 151–69.

Mlika, Mouna, Rotimi Adigun, and Beenish S. Bhutta. 2022. 'Silicosis'. In *StatPearls*. Treasure Island (FL): StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK537341/.

Moia, Nelly. 1998. *Für Die Katz'*. s.l.: Nelly Moia.

Molitor, Leon, and Josy Barthel. 1958. 'Etude sur la pollution atmosphérique au Grand-Duché de Luxembourg'. Ministère de la Santé Publique Luxembourg. Bibliotheque Nationale Luxembourg.

Molitor, Leon, and Michel Mosinger. 1960. 'Les Poussieres industrielles en pathologie humaine et expérimentale (Silicose et Silico-Anthracose, Sidérose et Sidéro-Silicose)'. 1. Plan et Première Partie 1. Bulletin de la Société des Sciences Médicales. Luxembourg: Haute Autorité de la Communauté Européenne du Charbon et de l'Acier. Bibliotheque Nationale Luxembourg.

Mosley, Stephen. 2001. *The Chimney of the World: A History of Smoke Pollution in Victorian and Edwardian Manchester*. Cambridge: White Horse Press.

Nurul, Ainun Hamzah, Bahri Mohd Tamrin Shamsul, and Ismail Noor Hassim. 2016. 'Assessment of Dust Exposure in a Steel Plant in the Eastern Coast of Peninsular Malaysia'. *Work (Reading, Mass.)* 55 (3): 655–62. https://doi.org/10.3233/WOR-162433.

Parmentier, Isabelle. 2008. *Histoire de l'environnement En Pays de Charleroi 1730-1830: Pollution et Nuisances Dans Un Paysage En Voie d'industrialisation*. Brussels: Académie Royale de Belgique.

Parr, Joy. 2010. 'Our Bodies and Our Histories of Technology and the Environment'. In *The Illusory Boundary: Environment and Technology in History*, 26–42. Charlottesville & London: University of Virginia Press.

Pfister, Christian, and Heinz Wanner. 2021. *Climate and Society in Europe-The Last Thousand Years*. Bern: Haupt.

Pritchard, Sara B., and Carl A. Zimring. 2020. *Technology and the Environment in History*. Baltimore: Johns Hopkins University Press.

Qing, Xiao. 2015. 'Assessment of Heavy Metal Pollution and Human Health Risk in Urban Soils of Steel Industrial City (Anshan), Liaoning, Northeast China'. *Ecotoxicology and Environmental Safety*, 9.

‘Retombées de poussières: Rapport sur les mesures de retombées en poussière effectuées de 1966 à 1969 à Esch-sur-Alzette’. 1972. Esch-sur-Alzette: Institut d’Hygiène et de Santé Publique, Laboratoire des Eaux. 637.3.

Rosen, George. 1993. *A History of Public Health*. Revised Expanded Edition. Baltimore: Johns Hopkins University Press.

Schueneman, Jean J. 1963. *Air Pollution Aspects of the Iron and Steel Industry*. U. S. Public Health Service, Division of Air Pollution.

Scuto, Denis. 1993. *Industriekultur in Esch: Eine Stadtgeschichtliche Wanderung Durch Die Luxemburger Minettemetropole*. Esch-sur-Alzette: Le Phare.

———. 2005. ‘Paul Flesch’. In *100 Joer Esch, 1906-2006*, 60–71. Esch-sur-Alzette: Ville d’Esch-sur-Alzette.

Stockie, John M. 2011. ‘The Mathematics of Atmospheric Dispersion Modeling’. *SIAM Review* 53 (2): 349–72. https://doi.org/10.1137/10080991X.

Su, Xianbo, Rui Ding, and Xinguo Zhuang. 2020. ‘Characteristics of Dust in Coal Mines in Central North China and Its Research Significance’. *ACS Omega* 5 (16): 9233–50. https://doi.org/10.1021/acsomega.0c00078.

Tarr, Joel A. 1996. *The Search for the Ultimate Sink: Urban Pollution in Historical Perspective*. Akron: University of Akron Press.

———. 2021. ‘Land Use and Environmental Change in the Hudson-Raritan Estuary Region, 1700-1980, With an Addendum to 2018’. In *Coastal Metropolis: Environmental Histories of Modern New York City*, 10–38. Pittsburgh: Pittsburgh University Press.

‘The Steel City and a Brief History of Dust Collection’. 2019. *Imperial Systems, Inc.* (blog). 19 June 2019. https://www.isystemsweb.com/the-steel-city-and-a-brief-history-of-dust-collection/.

Tripathi, A, C Souprayen, A Stanley, and N Warrilow. 2018. ‘Computational Fluid Dynamics or Gaussian - Is There a Right Way to Model Gas Dispersion ?’ In *Proceedings of the 10th IMA International Conference on Modelling in Industrial Maintenance and Reliability*. Institute of Mathematics and its Applications. https://doi.org/10.19124/ima.2018.001.26.

Uekötter, Frank. 2020. *Im Strudel: Eine Umweltgeschichte Der Modernen Welt*. Frankfurt/Main: Campus.

Valenti, Carla, Paolo Pozzi, Alessandra Busia, Roberto Mazza, Paolo Bossi, Cinzia De Marco, Ario Alberto Ruprecht, Alessandro Borgini, and Roberto Boffi. 2016. ‘Respiratory Illness and Air Pollution from the Steel Industry: The Case of Piquiá de Baixo, Brazil (Preliminary Report)’. *Multidisciplinary Respiratory Medicine* 11 (1): 41. https://doi.org/10.1186/s40248-016-0077-9.

Van de Maele, Jens. forthcoming. ‘“We Swallow the Dirt.” Popular and Governmental Perceptions of Metallurgical Air Pollution in Luxembourg During the Trente Glorieuses’. In *Rethinking the History of Deindustrialization: Mediating the Transformation of Industrial Cities in Postwar Europe*. Routledge.

WHO. n.d. ‘Hazard Prevention and Control in the Work Environment: Airborne Dust’. WHO Publications. Accessed 21 June 2022. https://www.who.int/publications-detail-redirect/WHO-SDE-OEH-99-14.

Wippich, Cornelia, Jörg Rissler, Dorothea Koppisch, and Dietmar Breuer. 2020. ‘Estimating Respirable Dust Exposure from Inhalable Dust Exposure’. *Annals of Work Exposures and Health* 64 (4): 430–44. https://doi.org/10.1093/annweh/wxaa016.

World Health Organization. 2021. *WHO Global Air Quality Guidelines: Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide*. Geneva: World Health Organization. https://apps.who.int/iris/handle/10665/345329.

Chapter 4: How can data visualization support interdisciplinary research? LuxTIME: studying historical exposomics in Belval

***Contribution statement*:** All authors have contributed to the conceptualization, review and editing of this article. I have given major input on the data visualization theory and Dagny Aurich has contributed to the writing of the section on interdisciplinary work and the discussion of the visualization examples. The visualizations were created by me based on the exchange with Dagny Aurich.

***Submission statement*:** This article has been published on the 29th of September 2023, in Frontiers in Big Data.

frontiers | Frontiers in Big Data

TYPE Methods
PUBLISHED 29 September 2023
DOI 10.3389/fdata.2023.1164885

EDITED BY
Florian Windhager,
University for Continuing Education
Krems, Austria

REVIEWED BY
Georgia Panagiotidou,
King's College London, United Kingdom
Reza M. Salek,
Bruker, Germany

*CORRESPONDENCE
Dagny Aurich
✉ dagny.aurich@uni.lu
Aida Horaniet Ibañez
✉ aida.horanietibanez@uni.lu

[†]These authors have contributed equally to this work and share first authorship

RECEIVED 13 February 2023
ACCEPTED 04 September 2023
PUBLISHED 29 September 2023

CITATION
Aurich D and Horaniet Ibañez A (2023) How can data visualization support interdisciplinary research? LuxTIME: studying historical exposomics in Belval.
Front. Big Data 6:1164885.
doi: 10.3389/fdata.2023.1164885

How can data visualization support interdisciplinary research? LuxTIME: studying historical exposomics in Belval

Dagny Aurich[1]*† and Aida Horaniet Ibañez[2]*†

[1]Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg, [2]Luxembourg Centre for Contemporary and Digital History, University of Luxembourg, Esch-sur-Alzette, Luxembourg

The Luxembourg Time Machine (LuxTIME) is an interdisciplinary project that studies the historical exposome during the industrialization of the Minett region, located in the south of Luxembourg. Exposome research encompasses all external and internal non-genetic factors influencing the health of the population, such as air pollution, green spaces, noise, work conditions, physical activity, and diet. Due to the wide scope of the interdisciplinary project, the historical study of the exposome in Belval involved the collection of quantitative and qualitative data from the National Archive of Luxembourg, various local archives (e.g., the communes of Esch-sur-Alzette and Sanem), the National Library, the Library of National Statistics STATEC, the National Geoportal of Luxembourg, scientific data from other research centers, and information from newspapers and journals digitized in eluxemburgensia.[1] The data collection and the resulting inventory were performed to create a proof of concept to critically test the potential of a multi-layered research design for the study of the historical exposome in Belval. The guiding navigation tool throughout the project was data visualization. It has facilitated the exploration of the data collected (or just the data) and the metadata. It has also been a valuable tool for mapping knowledge and defining the scope of the project. Furthermore, different data visualization techniques have helped us to reflect on the process of knowledge sharing, to understand how the relevance of certain topics changed throughout the project and why, and to learn about the publication process in different journals and the experience of the participants. Data visualization is used not only as a means to an end but also to embrace the idea of *sandcastles* using a speculative and process-oriented approach to advance knowledge within all research fields involved. LuxTIME has proven to be an ideal case study to explore the possibilities offered by different data visualization concepts and techniques resulting in a *data visualization toolbox* that could be evaluated and extended in other interdisciplinary projects.

KEYWORDS

data visualization, historical data, interdisciplinary research, exposome, digital history

---

[1] https://eluxemburgensia.lu/de: eluxemburgensia is a collection of digitized newspapers created by the Bibliothèque nationale du Luxembourg.

1. Introduction

Interdisciplinary collaboration, especially in projects involving researchers across the natural and applied sciences and the humanities, presents many challenges, including the multiplicity of research questions, the use of different methodologies, and the varying underlying assumptions based on different epistemic cultures. Bridging these epistemological and methodological differences to establish a shared understanding requires a substantial intellectual commitment from all participants. In an ideal scenario, such collaboration has the potential to produce new research questions, foster interdisciplinary approaches for the analysis of complex issues, and generate new knowledge based on interactional expertise (Fickers et al., 2022). Data visualization can play a significant role in facilitating such interdisciplinary collaboration.

Data visualization is widely used in research for exploring quantitative data, validating hypotheses, and communicating results, primarily using statistical charts such as bar charts, line charts, or scatter plots (Glivinska, 2021). It is less frequently used to study the research process itself. It is a process through which ideas are explored and a collective discourse is constructed using data visualization to critically observe different mediations (Hinrichs et al., 2019). In this article, we reflect on some of the data visualizations designed during the Luxembourg Time Machine (LuxTIME) project discussed in section 2. Based on the experience obtained during 2.5 years of interdisciplinary collaboration, working with a team of historians, hydrologists, chemists, and data visualization experts, we want to demonstrate that interdisciplinary research can benefit from an "extended" *data visualization toolbox*. We refer to a *toolbox* as a set of data visualization concepts and techniques, and we want to "extend" it, as opposed to only applying techniques frequently used within each specific field.

Specific disciplines can develop certain data visualization conventions. Considering visualization in interdisciplinary projects can help break out of these conventions and create new insights. Statistical graphs are used mostly in the natural, social, and applied sciences with the objective of exploring data to discover trends, patterns, and outliers, and to validate hypotheses or communicate results quickly for decision-making. Visualizations that emphasize esthetics and use metaphors and non-standard visual vocabularies are often found in journalistic, educational, and artistic contexts, while the study of interpretive practices and non-representative approaches predominate in the humanities. Each discipline can benefit from different practices. This becomes even more evident in the case of interdisciplinary research.

Combining these concepts and techniques predominant in different disciplines and applying them in the same context, we developed the *LuxTIME data visualization toolbox* exploring standard statistical graphs and variations thereof, concept maps, visual rhetoric, data humanism, multivariate data glyphs, non-representational approaches and visual elements of interpretation, and data storytelling. The use of different types of visualizations helped to map and exchange knowledge, thereby defining the project scope and the contributions of different disciplines, track timelines, and project deliverables. It also framed the participants' experiences along the way, inviting them to self-reflect on changes throughout the project and to explore the iterative research process. The *toolbox* aims to inspire a wide range of projects, especially those involving different disciplines. It describes and discusses the application of a set of epistemologically distant techniques and concepts from which to begin the exploration, and then add or remove tools, adapting the *toolbox* to each project. Foremost, it extends an invitation to explore and integrate other ways to visualize data, diverging from the traditional techniques commonly employed in each respective discipline.

In the current article, we start with a description of the LuxTIME project. Next, we discuss the methods, including the data collection process, the participants and their roles, the research questions addressed in the interdisciplinary context, the visualization concepts and techniques included in our project, and their application in the results section. We finish with a discussion of the frictions experienced as part of working on an interdisciplinary project with data visualization, how the idea of data visualization as a *sandcastle* (Hinrichs et al., 2019) has been present throughout the project, and finally, how other projects can benefit from our experience.

2. Case study

The Luxembourg Time Machine Project (LuxTIME) is an interdisciplinary research project funded by the Institute for Advanced Studies (IAS) of the University of Luxembourg.[2] Three research institutes engage in this so-called "Audacity Project", namely, the Center for Contemporary and Digital History (C^{2}DH), the Luxembourg Center for Systems Biomedicine (LCSB), and the Luxembourg Institute of Science and Technology (LIST). The main objective is to explore the potential implementation of a national platform ("Luxembourg Time Machine") that would allow scientists and stakeholders to "dive" into the complex past of this country using digital tools and data from different disciplines and fields. It is a subsidiary project of the European Time Machine project[3] adding a new dimension to the past.

By building a digital dataset including information from three very different fields and scientific perspectives, namely, eco-hydrology, environmental cheminformatics, and history, LuxTIME is using a local case (the industrialization of Belval and the Minett region) as a testbed for methodological and epistemological reflections on how to study the impact of environmental changes on the health of the local population, with a regional and long-term perspective. "Contextual information" based on archival evidence is mixed with "scientific evidence" derived from chemical, biological, or medical investigations as the project explores new grounds in interpreting "big data of the past" in a truly interdisciplinary setting. The Belval case study is a pilot project, preceding a project at the national level, the LuxTIME INITIATE.[4]

---

[2] https://luxtimemachine.uni.lu/

[3] https://www.timemachine.eu/

[4] LuxTIME INITIATE aims at building a consortium with the main stakeholders of historical data in Luxembourg (archives, libraries, museums, statistical offices, research institutions, governmental bodies, and private associations) to study the past in an interdisciplinary inter-institutional setting, with both intellectual and technical impact.

The above-mentioned impact of environmental changes on human health is covered by the exposome concept (Miller, 2020). Global and local changes have severe impacts on environmental systems and their inhabitants (Karlsson et al., 2021). The analysis of those changes and their impacts is a challenging task that can, however, be of help to understand and therefore prevent potential future outcomes. This could be, for example, any case of environmental pollution happening in an area resulting in an increase in disease cases as a phenotypic response. Human phenotypes, being a set of observable characteristics or traits, are mainly influenced by two factors and their interactions: the genetic factors described by the *genome* and all non-genetic factors covered by the *exposome* concept. Measuring the genome is difficult, but it is limited to the combination of four nucleotides that are stable over time. The *exposome* of an individual, however, is the measure of environmental influences (e.g., lifestyle, diet, and behavior) and the associated biological responses (Miller and Jones, 2014), changes within the course of a lifetime, and historical developments. Consequently, the *exposome* is influenced by many factors that vary over time and influence each other. Interdisciplinary efforts and data sources are required to come as close as possible to covering the whole picture.

The focus of this article was on how the use of a variety of data visualization concepts and techniques has supported our interdisciplinary (exposome) research.

3. Methods

3.1. Data collection

In this article, we refer to *data* as any collection of values conveying information, whether to represent abstract ideas (e.g., knowledge exchange and relevance of a topic), specific measurements (e.g., concentration of chemicals and number of articles published), or statistics (e.g., census data and steel production). We collected quantitative and qualitative data, mostly from secondary data sources (e.g., statistical archives, existing research), but also included some primary sources (e.g., observations from historical sources and reflections on the process itself). The collection for the LuxTIME project was done through observation, measurement, simulation, and analysis.

The Minett region is known for its past in the iron and steel industry, which was accompanied by many regional changes in terms of environment, socioeconomy, and health (Knebeler and Scuto, 2010). Some of the initial questions of the LuxTIME members included the following: where to look for data (*e.g.,* local and national archives); where to set the geographical and time limits; which topics to focus on (*e.g.,* environmental pollution); and how to obtain scientific data. For example, in the case of performing new chemical analyses, which sample types could be used? The initial objective was to find links between environmental pollution and other influencing factors of the past and disease patterns in the population by looking at archival sources combined with information received by scientific or governmental institutes and current chemical analyses, revealing facts about past exposures. This includes information related to the *historical exposome* in Belval such as datasets, images, text, events, and maps. For each dataset, relevant metadata was also collected, including title, source, reference from the source if available, author, publication date, the team member who collected the information, language(s), description, number of files, digitization status, class, format, type, period covered, geographical area covered, access rights, and the categories and subcategories of the *exposome* covered. Furthermore, we collected information about the process, such as the relevance of the different topics, disciplines involved, period of validity and reasons for change, relation to other disciplines, the amount of information found, and its potential for further analysis. We also collected data about the project deliverables such as the number of publications, the type, the disciplines involved, the different steps in the process (e.g., work starts, submission, and publication), the time, and the experience.

All these quantitative and qualitative data allowed us to study the research questions stated in section 3.3, using two data collection methods. First, a structured and normalized *data inventory* was created, where we included all the pieces of information found through the sources. To date, this table contains 121 records from 17 different information sources, each of them registered with the metadata described above. Second, for the evolution, reflection on the process, and experience, we used the visualization directly to generate the data (that could be extracted later if necessary).

The first steps included contacting governmental and scientific institutes to access past data already collected in the area (*e.g.,* scientific measurements of parameters such as soil or air pollution resulting from industry or other anthropogenic sources). Moreover, measures taken by the industry or the government to enhance life quality and health were investigated. Historical data—not only about environmental pollution but also social and economic data—were retrieved from archives. Newspaper articles, scientific reports and books, pictures, and contemporary witness reports were included. For chemical measurements generating scientific data, sampling campaigns of surface and groundwater were discussed as well as looking at dust samples and soil or biological samples such as mussels, trees, or even teeth. Based on the research on this topic from a chemistry point of view, a review article was published in 2021 (Aurich et al., 2021), helping to plan further project work packages in terms of, e.g., sampling campaigns. One example is the discussion of analyzing human samples to get to know more about past steel-pollution exposure and its health effects. However, the sample bank located in Luxembourg does not provide samples dating back to the steel industry times as it is a fairly new facility. The outcomes of the review showed many possibilities for how to access the historical *exposome* in terms of data and chemical analyses; however, none of them was available or feasible within the project scope and timeline.

3.2. Participants and roles

The core of the research described in this article was conducted during the collaboration between a researcher in data visualization and a researcher in environmental cheminformatics, with occasional feedback and participation from the overall project team that included other researchers in environmental cheminformatics, eco-hydrology, and history. The researcher in environmental cheminformatics already had disciplinary knowledge of data visualization and functioned as a domain expert in the field of the *exposome*, together with the rest of the environmental cheminformatics team. The data visualization researcher collected and studied many techniques and concepts used across different disciplines and then proposed an initial *toolbox* to discuss and experiment with during the visualization sessions are described in section 3.5. The researcher in data visualization also participated in the work of the target domain, *historical exposomics*, and the environmental cheminformatics researcher participated in visualization research. The collaboration was based on a *design-by-immersion* approach with reciprocal immersion, where both the visualization researcher and domain expert engaged with and participated in the work of the other domain, and knowledge emerged from the experiences and interactions (Hall et al., 2020). They not only participated in each other's research approaches and practices but also did archival work, collecting and analyzing historical sources. As discussed in section 5, this approach shapes and enriches the research team and also changes both participants' perspectives on their own fields.

3.3. Research questions

With data visualization as a navigation tool, our goal is to analyze the research questions outlined below.

Question #1: How could we map knowledge and knowledge sharing to define the project scope? The first question aims at exploring the *knowledge gap*, resulting from the different backgrounds and expertise in the project, e.g., history, design, and chemistry (Van Wijk, 2006). To create a "trading zone", i.e., "a space for interactions and negotiations between different knowledge domains" (Fickers and van der Heijden, 2020), we first needed to understand the knowledge within the group and what needed to be shared to define the project scope. In this research question, we explored how to arrive at an initial space, which exists at the boundary of four disciplines in which research can begin; a space in which diverse voices can speak and be heard and differences can be examined to mutually validate diverse perspectives, creating opportunities for mutual learning (Mao et al., 2019).

Question #2: How did the project evolve in terms of scope, relevance of the topics, the information available, and blending of the disciplines in each of these topics? The scope of the project was not permanent, nor was the knowledge or interest of the different disciplines in the different topics, which blended with different intensities throughout the process; as the project progressed, new sources of information appeared, and others were discarded. In this research question, we were interested in the evolution of how, from a series of central themes defined in the previous question, the scope changed throughout the project: which topics gained importance, which ones appeared or disappeared, and why.

Question #3: How could we explore the data and metadata respecting the priorities of the different disciplines? To synthesize disparate datasets in a visualization, especially in a project involving epistemologically distant disciplines, *data frictions* emerge regarding discipline-specific interpretation of the data, methodological approaches, ways of handling uncertainty, and scale and granularity in the datasets (Panagiotidou et al., 2022). Without analyzing in detail the different causes of such frictions, in this research question, we explored how the use of visualization techniques from other disciplines reveals the different priorities, and how through co-construction and exchange throughout the process, a common space of data and metadata visualization can be reached.

Question #4: How could we monitor the project deliverables, including the different steps of the process, the contribution of the different disciplines, and the experience of the participants? In addition to the *knowledge gap* discussed in the first research question, there is often an *interest gap*, caused by the different aims of the researchers, e.g., participation in different types of conferences and publication requirements (Van Wijk, 2006). These differences not only have an impact on the deliverables and, therefore, on the project timeline but also on the experience of the participants. In this research question, we explored the representation of the different steps in the process in time vs. in the experienced temporality.

We have explored each of these questions through visual means. In section 3.4, we introduce the concepts and techniques that we have included in our *toolbox* throughout the project followed by a discussion of the visualizations in which they are applied, in section 4.

3.4. The LuxTIME data visualization toolbox

We refer to *data visualization* as the graphical representation of data, using a variety of visual encoding methods (a representational approach), as well as the use of a graphical representation, to model interpretation and generate or augment data (a non-representational approach). The applications vary (e.g., exploratory analysis, data validation, hypothesis validation, and communication), and the techniques used depend on the intended purpose (e.g., quick decision-making vs. in-depth exploration of multiple narratives). Data visualization (or Dataviz) encompasses other terms including information visualization (InfoVis), information design, scientific visualization (SciVis), information graphics (Infographics), statistical graphics, or exploratory data analysis. The difference between these terms has been largely discussed in previous research (Rhyne et al., 2003; Manovich, 2011; Lankow et al., 2012; Kim et al., 2016).

In this section, we discuss several data visualization concepts and techniques that have been fundamental to our project and, therefore, essential elements of our *data visualization toolbox.*[5] The selection of "tools" for our *toolbox* is based on an extensive literature review and the study of numerous data visualization examples from different disciplines. It aims at integrating the perspectives on the field of the different disciplines (discussed below in relation to each concept) and above all

---

[5] The theory behind the concepts and techniques mentioned below might overlap but each of them has in some way proven to be useful during our project, and therefore, we considered it worthwhile including them all in our *toolbox.*

to experiment with concepts and techniques originating from epistemologically distant disciplines, which are rarely applied in the same context. The *toolbox* is built with an inclusive approach, integrating visualizations frequently used across disciplines (e.g., statistical charts) and potential variations, which also led to rich conversations and outcomes.

Some of these elements may be useful for other interdisciplinary projects; therefore, we present below a brief description and discussion of each of them, including examples from other projects. These methods and techniques are then applied in section 4 to answer our research questions.

3.4.1. Statistical graphs

The most frequently used data visualization techniques across disciplines are all types of statistical charts, such as bar charts to compare magnitudes, line charts to show evolution in time, or scatter plots to analyze relationships (Glivinska, 2021). The main purpose of these graphs is to summarize, validate, and communicate a message effectively. The use of statistical graphs assumes that the audience is data literate, i.e., has some basic knowledge of descriptive statistics; and that the designer[6] is aware of the extensive existing research on the subject.

Statistical graphs have a long history, and there is extensive literature about how to use them correctly to explore and summarize data effectively. However, academic and non-academic data visualization practitioners from different disciplines often fail to apply these theoretical principles. One of the frequent pitfalls is the wrong selection of charts despite the literature about graphical perception and suitability for analytical purposes (Lockwood, 1969; Cleveland and McGill, 1984, 1986; Evergreen, 2017). Other errors include the inappropriate use of colors based on color perception or cultural differences (Rogowitz and Treinish, 1998; Ware, 2004; Silva et al., 2007, 2011), the use of misleading graphs (Cairo, 2019), or poor storytelling (Nussbaumer Knaflic, 2015; Dykes, 2020). Moreover, the frequent lack of polish on elements such as axes, labels, gridlines, annotations, legends, descriptions, and titles (Schwabish, 2021) can create distractions from the core message (Tufte, 1999). These are just a few examples among many other considerations to be examined when creating statistical graphs (e.g., layout, context, transparency, accessibility, and interactivity).

Applied cases of statistical graphics to the LuxTIME are demonstrated in sections 4.3.1 and 4.3.2.

3.4.2. Variations of statistical graphs

We refer to the variation of a statistical graph when, knowing the theory mentioned in the preceding section, the designer decides not to apply one or more of these rules deliberately, e.g., to meet a specific use case requirement. Examples of such variations include duplicating the encoding (i.e., overencoding) to highlight a particular aspect of the graph (e.g., position and color encode the same information), changing the orientation, or overlapping graphs to favor a particular visual effect (e.g., crowdedness). Other options include removing axes, measures, gridlines, or titles to focus the audience's attention on the visualization; replacing predefined geometric shapes with other elements with a rhetorical value (e.g., using the representation of an object instead of rectangles in a bar chart); or any other type of visualization that starts from a statistical graph and modifies it with a purpose, beyond the one initially established for such a graph. It is often a combination of several modifications. Such variations of the conventional statistical graphs open numerous possibilities for the designer, especially in terms of communication.

The *climate stripes* by Ed Hawkins are a variation of a statistical graph, showing a simplified heatmap that has a strong influence on the climate change debate around the world. ShowYourStripes.info[7] registered 89,000 unique visitors to the site worldwide for 30 days during the summer of 2022 (Santoro and Kirkland, 2022). Hawkins emphasized the need for a range of ways of communicating the "climate crisis" because different people learn and experience in different ways and, therefore, justified the need for a range of climate visualizations to talk to different audiences. Another example is the visualization of China's overseas investments by Alberto Lucas López and Cédric Sam[8] where the bars of a bar chart are replaced by semicircles whose area represents the value of the deals and is displayed for many countries, overlapping in the layout. Laura Bronner, Anna Wiederkehr, and Nathaniel Rakich visualized the election night in 2020 *What Blue And Red "Shifts" Looked Like In Every State*[9] using simplified area charts in a tile map of the United States for FiveThirtyEight.[10] Kim Albrecht explores the randomness of success in scientific publications[11] using overlapping timelines in which only the maximums are highlighted. Another visualization example by Alberto Lucas López in collaboration with Ryan Williams and Kaya Berne is *Migration waves,*[12] where variations of area charts are used, through minimalism (no grids, no axes, and no measures), color (to emphasize positive and negative currents), and the superimposition of the graphs for the different countries.

3.4.3. Concept maps

Concept mapping is a visualization technique that uses hierarchical networks of nodes (concepts) and links (relationships) to represent visual knowledge (Romance and Vitale, 1999). The use of a concept map allows the inclusion of cross-links to map the links between concepts in different domains and represent creative leaps in knowledge production (Novak and Cañas, 2008). Joseph D. Novak and Alberto J. Cañas identified two features of concept maps that are important in the facilitation of creative thinking: the hierarchical structure and the ability to search for and characterize

---

[6] Designer refers to the person who designs the visualization (i.e., statistical charts), independent of their main area of expertise.

[7] https://showyourstripes.info/s/globe

[8] https://multimedia.scmp.com/china-overseas-investments/

[9] https://fivethirtyeight.com/features/where-we-saw-red-and-blue-mirages-on-election-night/

[10] https://fivethirtyeight.com

[11] http://sciencepaths.kimalbrecht.com/

[12] https://www.nationalgeographic.com/magazine/graphics/graphic-shows-past-50-years-of-global-human-migration?sf215829698=1&sf217104276=1

new cross-links. In addition to the purpose of enhancing creativity, concept maps are used as a design tool to generate structural organization, for communication purposes to overcome the limitations of the linear nature of the text, to stimulate the learning process by making the interrelationships between concepts explicit, and as an assessment tool to identify misconceptions (Lanzing, 1998). Johannes Wheeldon and Jacqueline Faubert argued that traditional definitions of concept mapping should be expanded to include more flexible approaches to the collection of graphic representations of experience, using concept maps to gather qualitative data from research participants (Wheeldon and Faubert, 2009).

In an interdisciplinary context where mapping knowledge of the different stakeholders is key to understanding the possibilities and developing the project, concept maps are a data visualization technique that helps to navigate the complexity, in terms of the variety of topics and subtopics and how they interrelate (see applied for LuxTIME in section 4.1).

3.4.4. Visual rhetoric

Rhetoric is the study of the communication techniques used to inform, persuade, or motivate a given audience in a particular situation, modifying their conceptions and attitudes toward the object of communication. Visual rhetoric (as an artifact) is the purposeful production or arrangement of colors, forms, and other elements to communicate with an audience (Hill and Helmers, 2009). One case of visual rhetoric in the context of data visualization is mapping rhetoric, which refers to “manipulating the information presentation via the data-to-visual transfer function, the constraints that determine how a piece of information will be translated to a visual feature” (Hullman and Diakopoulos, 2011).

A metaphor is a rhetorical figure, which refers to the cognitive process humans engage in when they reconceptualize a concept from a target domain in terms of another; and, therefore, when visual language is used to perform these functions, it is a visual metaphor (Steen, 2018). Despite the predominance of a minimalist data visualization approach that favors high data-ink ratios,[13] the data visualizations that use visual rhetoric—to engage, communicate, and be memorable using visual metaphors, and elements of embellishment—are still very present, especially in the fields of journalism, information design, or data art. Often these types of visualizations use organic forms, associated with nature, such as plants or rocks. Lima (2014) devoted an entire book to the study of trees in visualization, where he analyzes how throughout history, their trunks, branches, leaves, and fruits have served to represent connections between entities, through different domains of knowledge (e.g., family trees, systems of law, and biological species like trees, see application to LuxTIME in section 4.2).

An example of mapping rhetoric is *What’s Cookin?*[14] by Sarah Emery Clark, where she used the metaphor of preparing a meal to explore and analyze the state of the data visualization industry. Other examples of visual rhetoric include the visualization *Apparel Exports to the US* by Liz Bravo, where she visualized trends in the clothing industry using area charts shaped as sewing patterns; One *Angry Bird* by Periscopic,[15] displaying emotional arcs of the past 10 U.S. presidential inaugural addresses; or *The Great War* by Valentina D’Efilippo,[16] visualizing the fatalities during World War I as a poppy field. In *A View on Despair,*[17] Sonja Kuijpers visualized suicide in the Netherlands in 2017, representing the different categories in a landscape.

3.4.5. Data humanism

In *Data Humanism: The Revolutionary Future of Data Visualization,* Giorgia Lupi advocated for the connection of numbers to knowledge, behaviors, and people as data represent real life; making data unique, contextual, and intimate (Lupi, 2017). To do so, she promotes embracing a certain level of visual complexity, i.e., high-density data visualizations containing multiple attributes; and moving beyond standards, away from conventional graphics, and out-of-the-box solutions, to expand the “data-drawing vocabulary”. She stated that "data is a tool that filters reality in a highly subjective way” and, therefore, it is important to reclaim a personal approach to how data are captured, analyzed, and displayed, as data are imperfect. She urged a paradigm shift to “always sneak context in”, in which data visualization embraces imperfection and approximation, “allowing ways to use data to feel more empathetic, to connect with ourselves and others at a deeper level”.

In their book *Data Feminism*, Catherine D’Ignazio and Lauren F. Klein stated that “refusing to acknowledge context is a way to assert authoritativeness and mastery without being required to address the complexity of what the data actually represent” (D’Ignazio and Klein, 2020). Based on the concept of “situated knowledge”, initially raised by Donna Haraway in the 1980s, they stated that the responsibility of ensuring that the situatedness of data is considered is with the person evaluating the knowledge or building upon it. Yanni A. Loukissas referred to errors in data collection as “signifiers taken out of their original interpretative texts” (Loukissas, 2019). Hannah Schwan, Jonas Arndt, and Marian Dörk (Schwan et al., 2022) identified key aspects of disclosure, i.e., the aspiration to be conscious of the potential effects of the designer’s assumptions. They invited “the viewer into exchanges with the designer, reflections about the visualization, and engagement with an issue” (Dörk et al., 2013) and proposed several representation forms to integrate the disclosure information into the visualizations.

Some examples of visualizations that use information-rich designs and custom visual vocabularies, and highlight the relevance of details and the imperfection of the data include *Data Items:*

---

[13] The data-ink ratio is a concept introduced by Tufte (1999) as the proportion of ink that is used to present actual data compared to the total amount of ink (or pixels) used in the entire display.

[14] https://www.sarahemeryclark.com/work/whats-cookin

[15] https://emotions.periscopic.com/inauguration/

[16] http://poppyfield.org

[17] http://www.studioterp.nl/a-view-on-despair-a-datavisualization-project-by-studio-terp/

A Fashion Landscape,[18] visualizing the role of fashion connecting people and cultures, and *Bruises—The Data We Don't See,*[19] depicting a sensorial picture of a personal journey with a disease, both by Giorgia Lupi. In *Trending seeds,*[20] Valentina d'Efilippo and Lucia Kocincova analyzed and visualized the Twitter social movement #MeToo to understand if social media could become a vehicle to foster social change and reshape traditional views. In *Data Selfi,*[21] Kadambari Komandur explored intersectional feminism. These are just a few examples among many others. Moreover, the concept of data humanism has been discussed in research articles over the last few years. Kim et al. (2018) presented an interface that enables designing and personalizing visual vocabulary to represent data, and they explored how to enable people to determine the representation of their data based on the *Dear Data* project (Kim et al., 2019).[22] Cordell advocated for exploratory, iterative, and dialogic data humanism to foster humanistic engagement with data in an academic context (Cordell, 2019). This concept is applied in several visualizations in section 4.

3.4.6. Multivariate data glyphs

A data glyph is a visual representation of data where the attributes of the graphical entity are defined by the attributes of the data record. It is a visualization technique often used for multivariate data because patterns involving more than two dimensions can often be perceived more easily in this manner (Ward, 2008). Multivariate or multidimensional data consist of a list of records, with multiple columns (variables), which may be either numerical or categorical values. The encoding can map one data attribute to one single graphical attribute; or use redundant mappings (e.g., using tone and shape to display the same variable) to facilitate the interpretation or reinforce a message. Small multiple data visualizations contain several small graphs arranged on a grid, where every representation follows the same structure; and leverages visual constancy, economy of perception, and uninterrupted visual reasoning (Chuah and Eick, 1998). Small multiples can be based on conventional graphs, but also on information-rich glyphs that encode data attributes using customized visual vocabularies. Glyphs can be displayed in a layout based on data variables, data structure (e.g., time and hierarchies), or any other layout (e.g., predefined shape and screen size). The use of glyphs allows the designer to define the level of aggregation, where each glyph is the level of detail selected. They are often used to explore details (e.g., people, objects, and cases) as they allow us to visualize multiple characteristics about each subject (see applied in section 4.3.1).

Take a walk down Fifth Avenue[23] by Molly Morgan uses data glyphs to represent the physical characteristics and the ecological benefits of every tree along Fifth Avenue in New York. Alberto Lucas López in *Our daily faces*[24] depicts every page of the South China Morning Post newspaper over a year, encoding the subjects covered or the length in small glyphs. *Representation of women in politics*[25] by Frederica Fragapane visualizes the top 40 countries in the world by political parity score, using a combination of graphical attributes to represent multiple variables for each country, such as the percentage of seats held by women in local government bodies, in lower and upper houses of national legislatures; the number of female candidates in the most recent elections; the number of elected or appointed heads of state or the geographical area.

3.4.7. Non-representational approaches and interpretation

Inspired by the work of Thrift (2008), Johanna Drucker defined non-representational approaches to modeling interpretation in a graphical environment "as the use of graphical means as a primary method of modeling human-authored interpretation rather than to display preexisting data sets" (Drucker, 2018). In contrast to representational approaches, the existence of data or other representations is not assumed before the interpretative work; the relationship between data and visualization is not unidirectional. Visualization can be the starting point, where we add, for example, connections or annotate reflections and use any visual vocabulary to encode high-level concepts (e.g., contradiction and comparison). This could be later captured as data. As already discussed in previous sections, situated, experiential, and embodied forms of knowledge have been largely researched (in contrast to observer-independent empirical approaches). "A humanistic approach is centered in the experiential, subjective conditions of interpretation" and, therefore, it requires a shift toward non-standard metrics, where the challenge is to design graphical expressions that display interpreted phenomena (Drucker, 2011). Time, for example, is modeled not to imitate its physical dimension but to provide a model that reflects the phenomena under consideration to support a given set of analyses (Aigner et al., 2011). Conventional approaches to timelines are linear, unidirectional, continuous, and structured with a single standard metric unit because they take their structure from the temporal models used in the natural sciences. In humanities, temporality is experienced as asynchronous, variable, broken, and heterogeneous (Drucker, 2021).

There are several examples of data visualizations being used as the starting point to collect and visualize data, for example, in participatory projects using street data walls such as the *Mood Test*[26] by Domestic Data Streamers to analyze people's attitudes toward life; or in data physicalization projects such as the *Data Badges* (Panagiotidou et al., 2020) that invited participants of a conference to make their own customized expressions of their academic profiles. However, examples of interpretive exercises through visualization where graphical attributes are designed to

---

[18] http://giorgialupi.com/data-items-a-fashion-landscape-at-the-museum-of-modern-art

[19] http://giorgialupi.com/bruises-the-data-we-dont-see

[20] http://metoomentum.com/trending.html

[21] https://kadambarik.myportfolio.com/data-selfi

[22] http://www.dear-data.com

[23] https://www.5avestreettrees.com

[24] https://www.lucasinfografia.com/Front-pages-analysis

[25] https://www.behance.net/gallery/138862771/Women-in-politics?tracking_source=search_projects%7Cinformation%20visualization%20data%20art

[26] https://domesticstreamers.com/projects/the-mood-test/

show subjective conditions of interpretation are practically non-existent, beyond the theoretical study presented above and some tool prototypes such as the 3DH project.[27]

When we want to approach analysis from a more humanistic perspective, for example, to analyze an experience, the use of a non-representational approach to data visualization using standard and non-standard metrics, as required, can facilitate the interpretative exercise. In the case of the LuxTIME project, we use this technique to reflect on the evolution of the project's topics in section 4.2 and to understand the participants' experience in section 4.4.

3.4.8. Data storytelling

Data storytelling is a powerful mechanism for sharing insights that involve data, narrative, and visuals to explain, i.e., narrative couples with data, explain; visual couples with data, enlighten; and narrative coupled with visuals, engage (Dykes, 2020). Storytelling makes data interesting, facilitates the understanding of complex subjects, encourages action, and is memorable (Vora, 2019). Brent Dykes identifies nine main tasks as part of the storytelling process: identifying key insights, being aware of biases, having extensive background knowledge, understanding the audience, curating the information, assembling the story, providing narration, choosing the visuals, and adding credibility (Dykes, 2020). Edward Segel and Jeffrey Heer developed a framework of design strategies for narrative visualization in the context of journalistic storytelling, where they placed the visualizations along a spectrum of *author-driven,* i.e., linear structure, heavy messaging, and no interactivity, and *reader-driven* approaches, i.e., highly interactive with no clear path to the story (Segel and Heer, 2010). "Narrative information visualizations rely on rhetorical techniques, to convey a story to users as well as exploratory, dialectic strategies aimed at providing the user with control over the insights gained from interaction." (Hullman and Diakopoulos, 2011).

Framing Luxembourg,[28] a timeline tracing the history of public statistics in Luxembourg, is an example of data storytelling, where the narrative is divided into different chapters (e.g., migration, family, and employment), and *scrollytelling*[29] is used to facilitate the navigation through text, images, and charts.

3.5. Data visualization and feedback sessions

The visualizations presented in section 4 have been co-created by the two main participants, integrating the feedback from the other participants. The work has been developed over more than 2 years of the project, during monthly work sessions. The different types of sessions were not initially defined but were designed during the monthly meetings as the project progressed. We have identified retrospectively 5 types of sessions that took place during the project and have named them to facilitate their discussion and future application.

- *Individual and collective preparation sessions*, where the participants reviewed the data available to date and validated the initial research questions (or reformulated them based on available information). Major advances in the data collection process were discussed in a meeting with the entire team, followed by individual preparation sessions and a final group session to agree on the data and the variables to be used and, in some cases, to reformulate the research questions.
- *Data visualization toolbox discovery sessions*: The data visualization researcher introduced the less-known concepts and techniques during these sessions. Additionally, possible improvements or alternatives were discussed when the *session* was used to review visualizations already created during the *sketching* and *pilot sessions* (e.g., the use of statistical graphs).
- *Sketching sessions*: During these sessions, both participants experimented with different visualization ideas, using R, Python, Excel, or Tableau for more standard charts, and paper and markers to design tailor-made visual vocabularies. Miro was also used for concept maps and brainstorming.
- *Pilot session*: In these sessions, the selected ideas were refined and completed. The result was the data visualizations that were ready to share with the rest of the team. The tools used remain the same as for standard charts, and Adobe Illustrator was used for static visualizations with custom visual vocabularies.
- *Feedback sessions*: Feedback from the rest of the team was collected during these sessions. These exchanges also brought to light the friction between the different disciplines, as we will discuss in section 5.

All these sessions contributed to the creation and application of our *data visualization toolbox*. The types of data available and the research questions that were developed during the *individual and collective preparation sessions* were the basis for the search for suitable visualization techniques. Once these techniques had been collected by the researcher in data visualization, the *data visualization toolbox discovery sessions* allowed the presentation and discussion of these techniques with other researchers. During the *sketching* and *pilot* sessions, the selected tools were applied in a practical way to the project. Finally, the results were discussed during the *feedback sessions* with the entire team. These steps were part of an iterative process that moved back and forth as new data became available, new techniques were added to the toolbox, or new feedback was received.

4. Results

4.1. Working in an interdisciplinary team

Working as an interdisciplinary team and learning how to exchange ideas, data, and experiences were at the core of this research. In this section, we discuss how, with the support of data visualization, we laid the foundations for dialog and defined the main themes of the project and the role of the different disciplines.

---

[27] https://threedh.net

[28] https://www.framingluxembourg.lu, developed by the C2DH in collaboration with STATEC.

[29] *Scrollytelling,* from "scrolling" and "storytelling", is a way to display content that unfolds as the user scrolls.

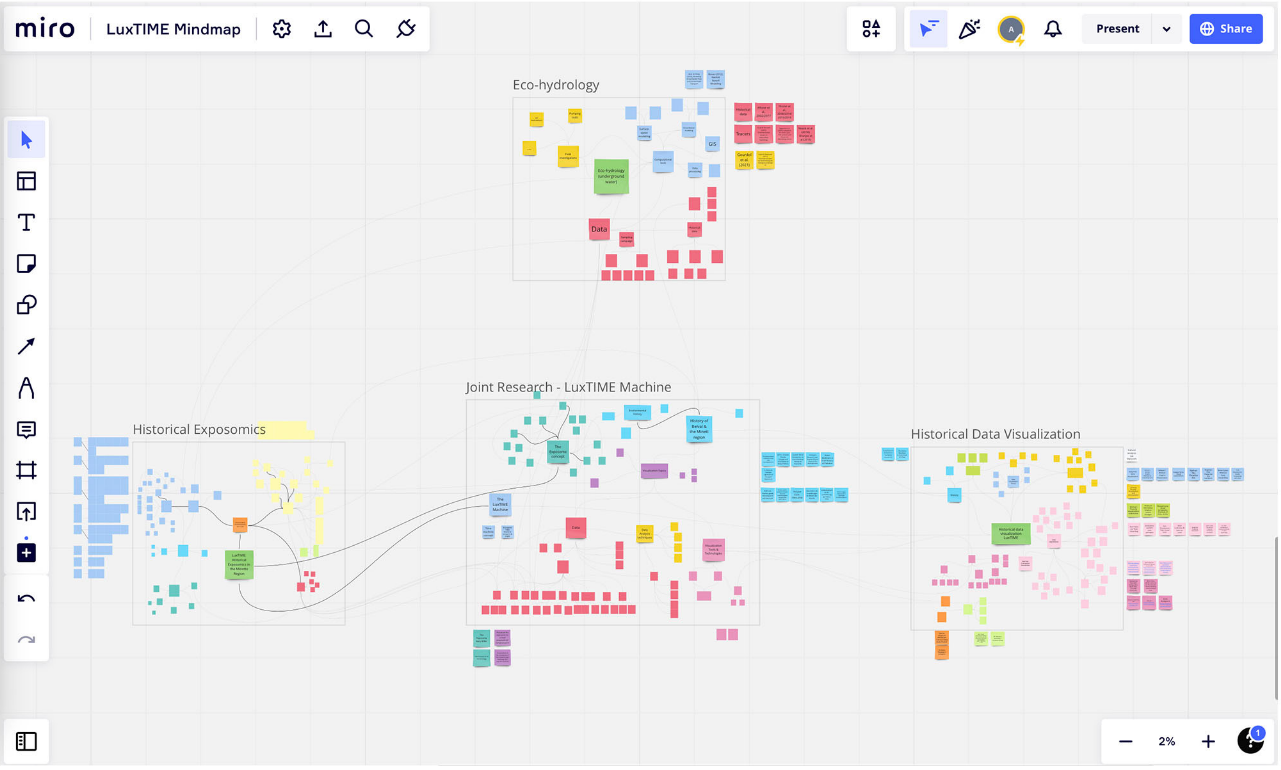


FIGURE 1
LuxTIME project concept map developed using the Miro online visual collaboration platform.

We, therefore, focus in this section on research *question #1: How could we map knowledge and knowledge sharing to define the project scope?*

Our first interdisciplinary challenge was to map the knowledge in the team to understand the skills of each team member and how they could contribute to the project. The initial team consisted of three PhD students, three supervisors, and five supporting researchers from the different research centers involved. The foundation of any interdisciplinary project is the interest in understanding other disciplines and finding out what level of learning and collaboration is required to produce new "joint knowledge" [*trading zone* concept, see (Kemman, 2021)].

To support this process, we used a *concept map*, where we mapped the knowledge of the different disciplines and the "joint knowledge" and how they related to each other. The concept map was used to add literature about the different topics, as a starting point for the other participants to familiarize themselves with the topics, and to collect the references to the first datasets. This visualization also served as both an internal and external communication tool. The interactive online platform Miro[30] was used for this purpose. Figure 1 shows an example of this collaborative platform, which was also used in interactive workshops such as the *UniTalks LuxTime Machine: Back to the future*, where the participants (e.g., researchers from other areas, librarians, and other university stakeholders) were encouraged to complete the concept map by adding terms and links, after reading a list of guiding questions such as "what exposures would you consider understanding the *exposome* of the population in the Minett region over the last 200 years?" and "where would you look for information?". The result of the workshop was an enriched version of the concept map that helped us to identify new ideas and points of view.

The final version of the concept map is shown in Figure 2. In the initial versions, we only mapped three disciplines: history, environmental cheminformatics, and eco-hydrology, displayed in different colors; and the interdisciplinary project "overlap", in gray. In later versions, we decided to map *data visualization* as a fourth discipline, since it is not only a tool that helps to achieve a technical task but a branch of knowledge that also contributes at a theoretical level. Certain topics moved from outside "specific knowledge" to inside "joint knowledge", such as "the exposome", which originated in the domain of environmental cheminformatics but became the central focus of the project; or the industrial history of the Minett region, which initiated exclusively under the expertise of the historians and evolved to become a major area of joint knowledge.

4.2. Analyzing the process: topics, information available, and participation

As discussed above, the project involves several disciplines, with a broad area of joint overlap. Once the contributions of the

---

[30] https://miro.com

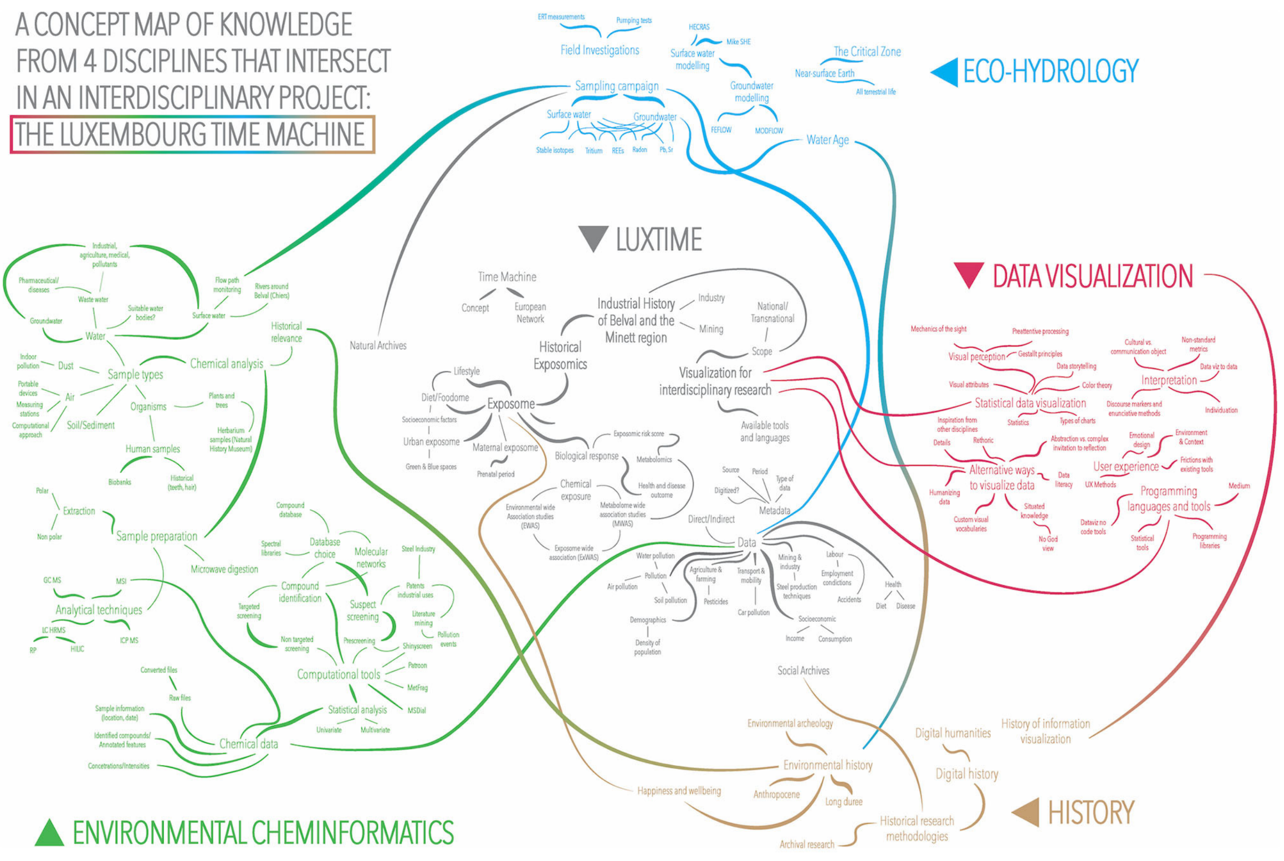


FIGURE 2
Final LuxTIME project concept map.

participants were understood, the project scope was defined and revised throughout the project. In this section, we reflected on the process: which topics were the most relevant at each moment and why, which disciplines participated in different topics, which topics had sufficient information for further development, and which ones fell out of the project scope. Thus, here, we focus on research *question #2: How did the project evolve in terms of scope, relevance of the topics, available information, and blending of the different disciplines in each of these topics?*

The concept map allowed us to see independent snapshots of the topics at different points in time in the project, but it did not allow us to compare their evolution throughout the project in the same view, in terms of the appearance and disappearance of certain subjects, to what extent they had been integrated into the different disciplines, how relevant they were considered at each time, and what were the reasons behind these changes. The data to do this analysis were not collected during the project in a structured way, since, in this case, we wanted to use the visualization directly as a reflection tool (which could generate data later, if necessary).

The result of this exercise was the visualization in Figure 3, where we created a timeline with three checkpoints (January 2021, January 2022, and January 2023), around which we randomized the themes. We included a separation line "inside/outside" of the project, which allowed us to visualize which topics and around what time had been excluded. Each circle represents a topic, and the size (3 levels) indicates how important it was considered at a certain point of time in the project. The circles have a black border at the last check point before being excluded from the project scope. The color represents discipline (this color is kept consistent throughout the different visualizations in which the discipline variable is displayed). In addition, we added annotations on the lines linking topics over time to explain how a topic is integrated into different disciplines, why a topic is excluded, or how it is merged into a different topic.

In the visualization, we observed that at the beginning of the project, there were many different discipline-specific topics with lower relevance, except for the data and the tools that were common to all disciplines from the beginning. In the second checkpoint, topics such as the field investigation integrated both eco-hydrology and environmental cheminformatics with increased relevance, but then, it was excluded from the scope due to changes in the team. We also saw how the concept of historical exposomics became more important, integrating the historical and hydrological perspectives, thanks to the research seminars, literature review, and writing workshops.[31] At the last checkpoint,

---

[31] LuxTIME Seminar Series organized throughout the project to facilitate the exchange of knowledge and stimulate the discussions within and outside the working group: https://luxtimemachine.uni.lu/#1620717521789-01362bff-416e.

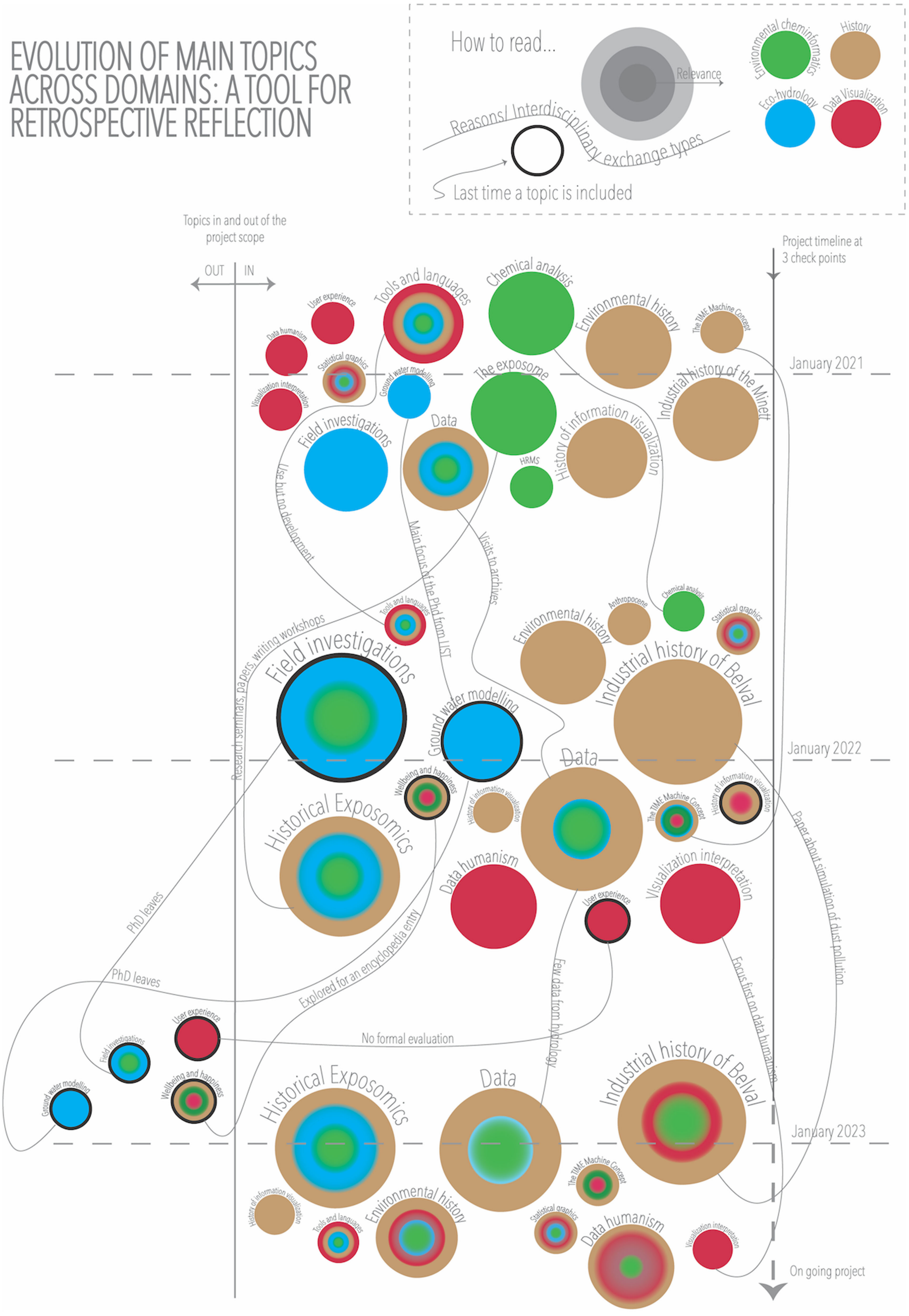


FIGURE 3
Evolution of main topics of LuxTIME.

only a few topics remained, with higher relevance and a blend of history, environmental cheminformatics, eco-hydrology, and data visualization. Overall, we saw a prioritization of topics, knowledge transfer, and shifting contributions among the project members, representing the functioning interdisciplinarity aspect of the project.

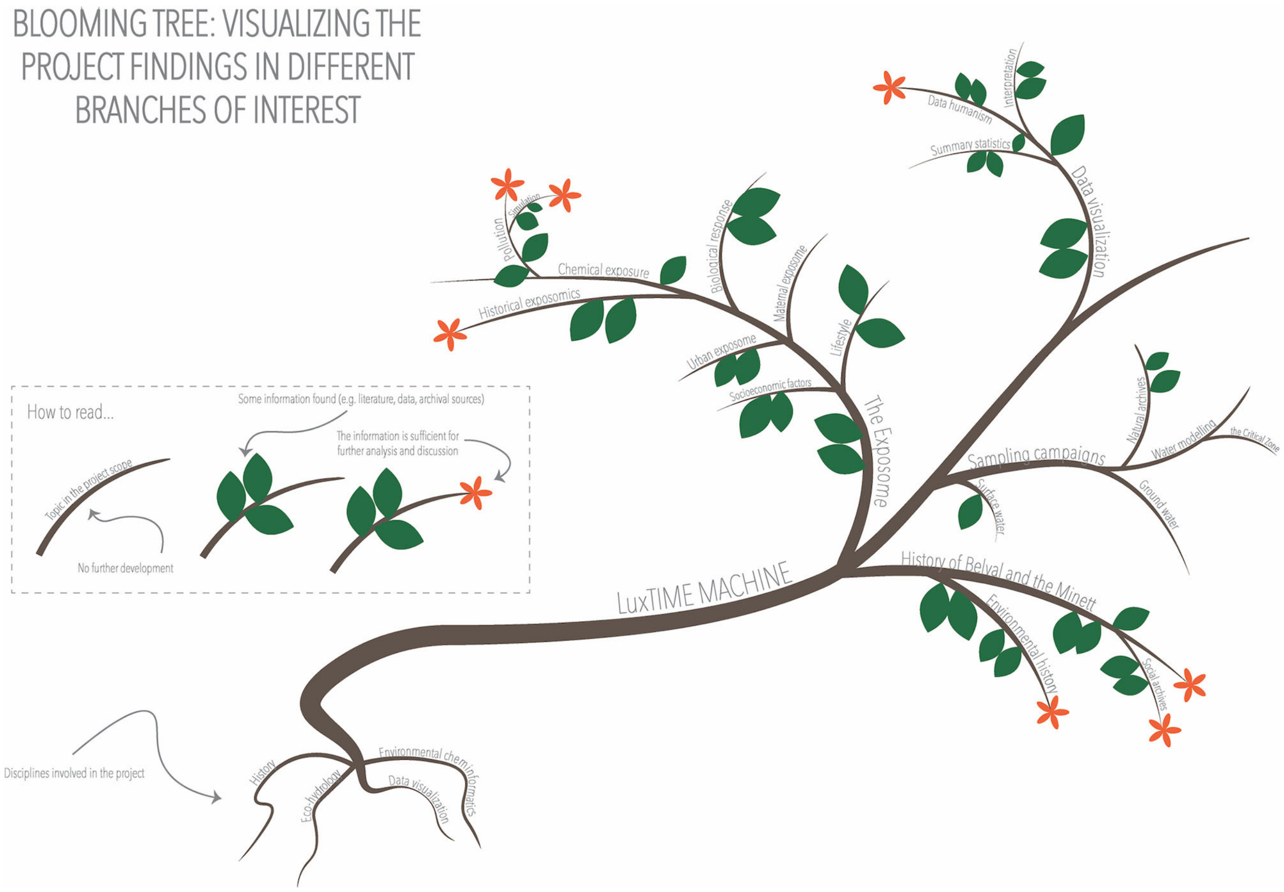


FIGURE 4
LuxTIME tree visualizing the project developments.

This visualization applied the principles of a non-representative approaches are explained in section 3.4.7, where data are not collected before the visualization is created. We used graphical features such as shapes (e.g., circles and lines), size, color, and annotations; to reflect on the evolution of the project. We drew a line connecting circles when we noticed a connection and increased or decreased the size of a circle after discussing the relative relevance of a topic at a given point in time. The existence of these elements can be registered as data, but such data did not exist beforehand, and the exercise started with the interpretative process.

After defining the topics of interest, we collected all the information found through public entities, research centers, libraries, historical archives, and other sources presented in section 3.1. In the visualization in Figure 4, the objective was to display the areas with the most potential based on the information found, i.e., the "flourishing branches" (see visual rhetoric is described in section 3.4.4 at a given moment). We defined three levels: The first was just the branch, which stated that the topic had been considered and researched; the second level was illustrated with a leafy branch, indicating that some information had been found (e.g., historical sources mentioning the topic, existing previous research, and literature available); and the third one, a flowering branch, depicted the possibility of further analysis to extend the research (e.g., data can be extrapolated to the Minett region, it concerns the period of interest, or a projection can be done). The tree, at the same time, allowed us to separate the branches into themes and sub-themes and to represent the roots of the project, the four disciplines. As we can see in Figure 4, the most developed branches, at the time of visualization, were pollution, data humanism, historical exposomics, environmental history, and the discovery of social archives about the history of the Minett region. Other topics for which some information had been found but was not yet sufficient for further analysis included other areas of the *exposome* (e.g., urban exposome, lifestyle, and biological responses) or the sampling campaigns. This visualization could be annotated to explain why these areas had been further developed, which types of data had been found, where, and what analyses they allowed. It could also be combined with a second visualization showing the details, an animation (e.g., showing a tree whose leaves appear, the flowers bloom and then wilt, and the leaves fall off), or adding interactivity that has been neglected in this first visualization phase.

The use of the tree allowed us to use multiple visual metaphors, the roots representing the disciplines, support, and

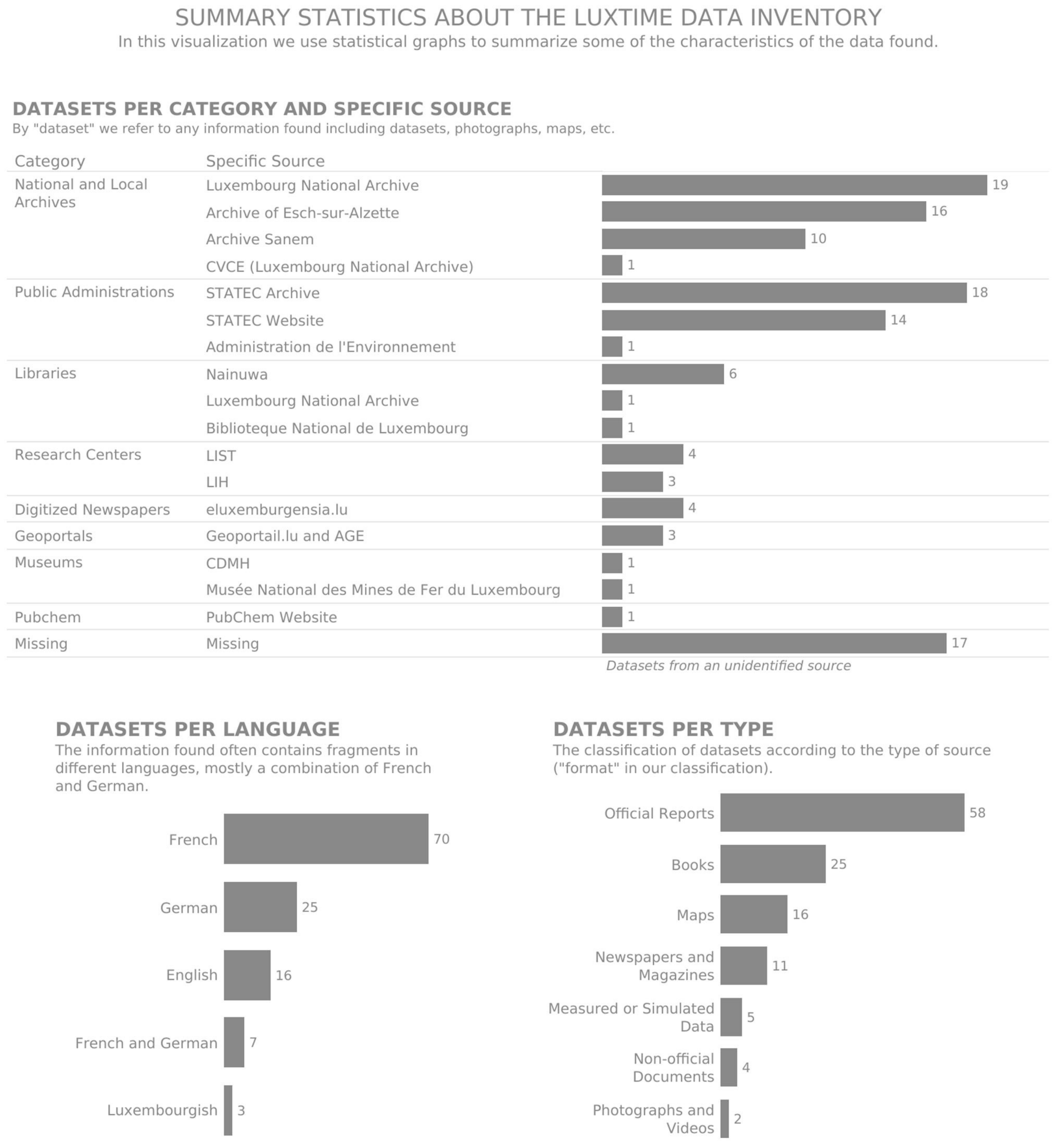


FIGURE 5
Summary statistics about the LuxTIME data inventory.

nourish the project; the branches grow and branch out as the project progresses, and the leaves and flowers bloom due to multiple factors. This visualization showed the progress of the project without using any conventional numerical or graphical charts and could be used to communicate with all kinds of audiences. It also showed which branches of the project were "blooming" at first glance, as a starting point to discuss how to move forward.

4.3. Exploring data and metadata

One of the most challenging aspects of the project was the data collection. In addition to the complexity due to the wide range of topics, there were also different types of sources (e.g., texts, images, and maps) and archives. The inventory of the information continues, as well as the analysis of the different datasets. This section focuses on research question #3 using selected examples:

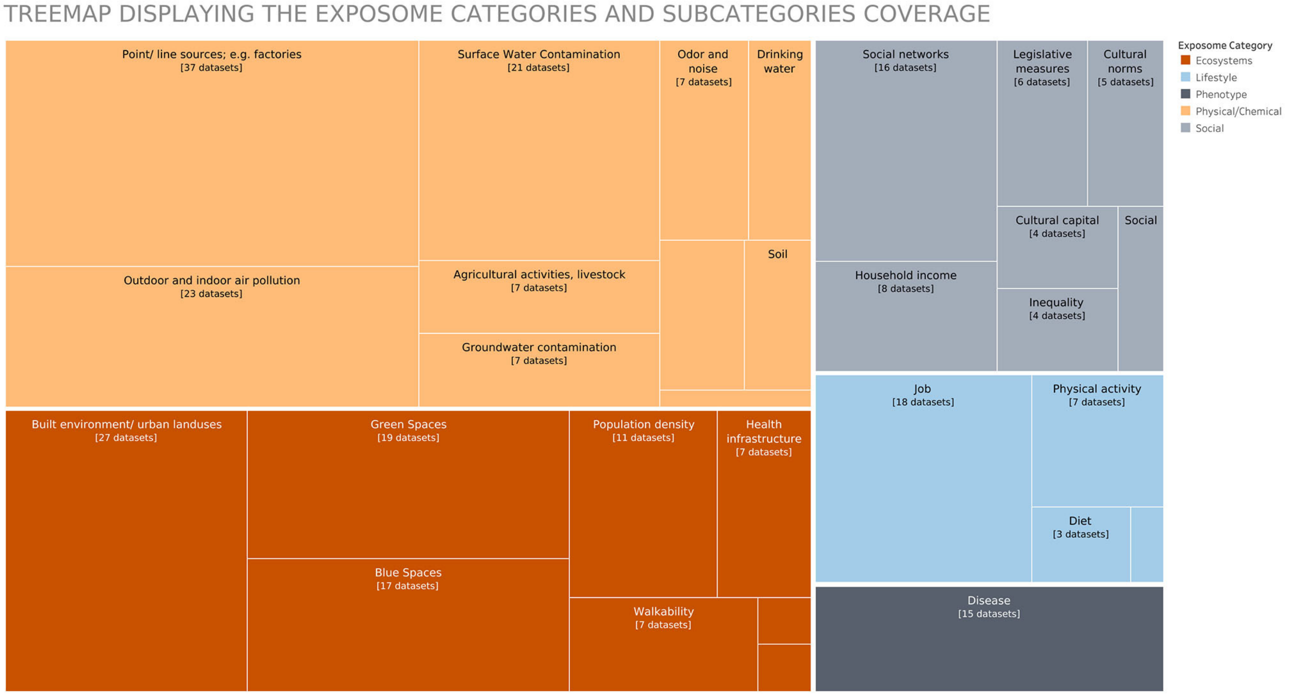


FIGURE 6
Treemap displaying the data inventory by exposome category and subcategory.

How could we explore the data and metadata respecting the priorities of the different disciplines?

4.3.1. Visualizing metadata of the information collected

The use of statistical graphs, such as bar charts allowed us to explore individual variables of the inventory, answering questions such as *How many datasets per source did we have?* This question could be easily answered using summary statistics (Figure 5), as the inventory contains 21 datasets from the Luxembourg National Archive, 18 from STATEC, 16 from the Archive of Esch-sur-Alzette, etc.

How many datasets covered every category and subcategory of the exposome? We did a subdivision of the *exposome* into five categories (Ecosystems, Lifestyle, Social, Physical/Chemical, and Phenotype) and 43 subcategories. In Figure 6, we visualized this using a treemap, with the number of datasets related to the different categories and subcategories. Most datasets belonged to several classification categories, which are not displayed in the treemap. The analysis of this overlap would require a different visualization. In the graph, we could see that the categories with the most information were physical/chemical and ecosystems, followed by social and lifestyle.

The visualizations in Figures 5, 6 allowed us to summarize the metadata, an analysis that could be extended to all the collected metadata in the inventory (e.g., time period, geography, and authors). However, we could not include multiple variables in the same view and still see the details of each of the collected datasets. For this purpose, we used multivariate data glyphs (see Figure 7). One of the principles of *data humanism* (Lupi, 2017) already introduced in section 3.4.5 is the use of dense and unconventional data visualizations to promote exploration, as it requires the reader to become familiar with the visual encoding, and it layers multiple visual narratives for the readers to follow their own interest "since clarity does not need to come all at once".

As we see in Figure 7, we selected a series of graphical attributes customized to the visualization. The color of the outer pentagon represents the source. In the inventory, we collected 121 datasets from 17 sources, which we grouped for visualization into national and local archives, museums, libraries, research centers, the geoportal, public administrations, and eluxemburgensia. The inner pentagon, in gray, represents the areas of the *exposome* covered in the dataset. The outer square with a thick gray border indicates that the source has been digitized, and each corner of the inner square indicates in which language it is available. At the bottom of the pentagon, the period covered by the data was indicated, with the base coinciding with the period of interest of the project, between 1,850 and 1,950. The concentric triangles at the top right indicate from the inside out the geographical area covered by the data, from the Minett (smallest triangle) to a global scope (largest). This visualization, despite requiring more time to explore, would allow us to see different aspects at the same time, without losing sight of the granularity of the dataset. Further iterations of glyphs and layouts will be evaluated before final implementation.[32]

---

[32] All visualizations are prototypes developed throughout the project, at the end of which, a final version will be integrated into the project website https://luxtimemachine.uni.lu/.

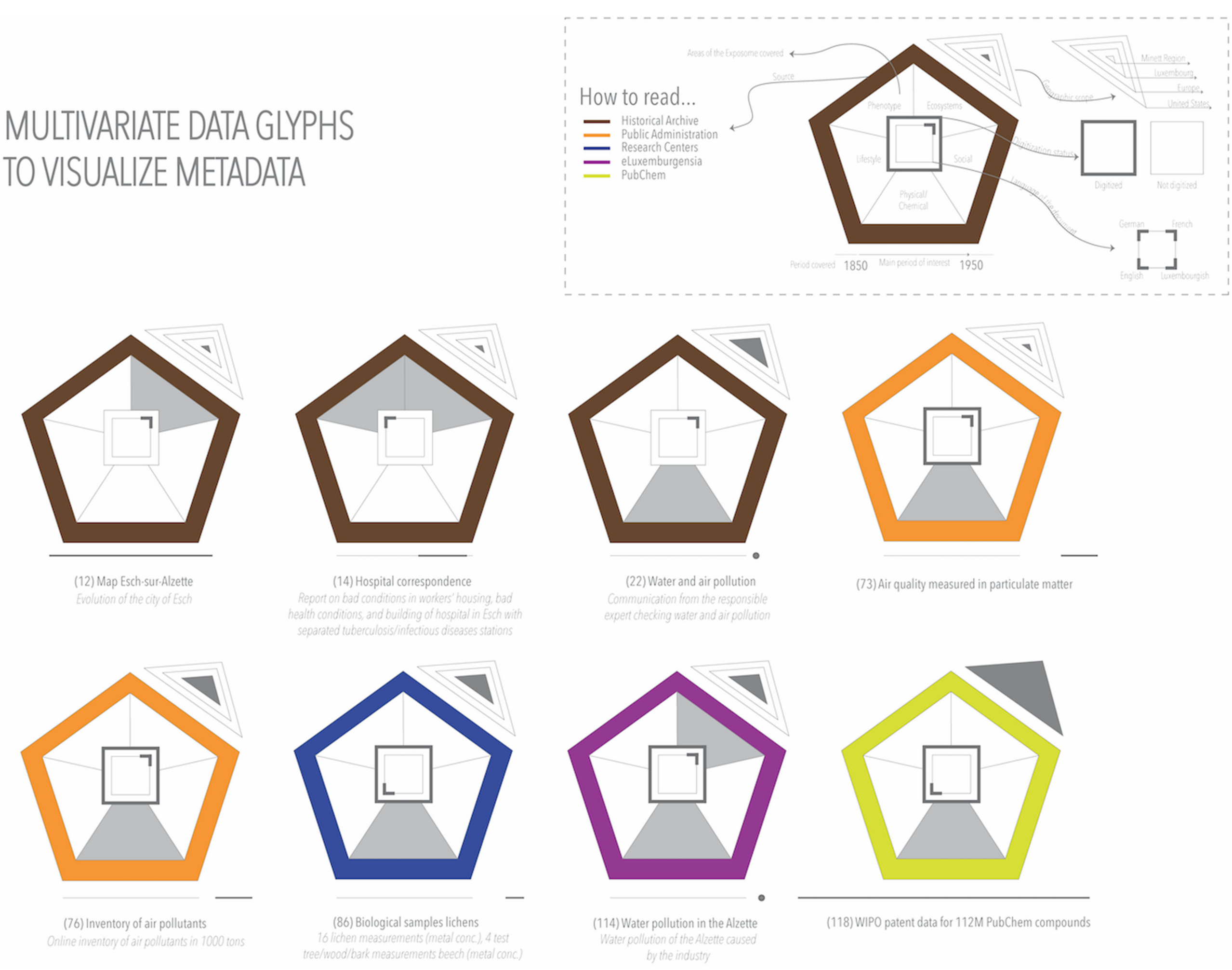


FIGURE 7
Multivariate data glyphs.

4.3.2. Exploring a dataset: number of chemicals registered over time

After having explored the metadata of the data inventory, we started to analyze the datasets. Given the variety of datasets in the inventory, each required a particular analysis to define the required data visualizations. We have chosen, as an example, a set of data about the number of chemicals registered in the Chemical Abstracts Service (CAS) registry[33] over time.

Figure 8 shows a classical representation of quantitative (chemical) data using a line chart. As an alternative, in Figure 9, *chemical stripes* (Arp et al., 2023) (inspired by the *warming* or *climate stripes* discussed in section 3.4.2) are shown, presenting trends of chemical registrations in the CAS registry since 1965, with low numbers in light red and high numbers in dark red. In Figure 9, the color hue was chosen to alert about the situation of increasing chemical numbers, while the color values (from light to dark) allowed the use of red and made it possible for readers with color deficiencies to see the differences.

The numbers of chemicals registered in the CAS registry do not correspond directly to the number or amount of chemicals in use. The stripes could be created for different compound classes or a group of several chemicals, having the registration numbers for these individual requests. The general trend remains the same, irrespective of the view or data: The number of chemicals in use and present in the environment increases along with the number of chemicals. The total number of registered unique substances in the CAS registry lies over 200 million substances currently, with 10–20 million new entries added per year. As in the case of the *climate stripes*, this simplified heatmap conveyed a clear message about the increasing number of chemicals. As discussed in section 3.4.2, this is an example of a variation of a statistical graph, a heatmap, where the graph elements (e.g., axes, legend, and tick marks) are removed to draw attention to a single message, the increasing or decreasing values through the color hue. The objective is not to provide exact values to the reader but to show a noticeably clear trend through minimalism in the visualization.

---

[33] https://www.cas.org/cas-data/cas-registr

FIGURE 8
Line chart: Number of unique substances registered in the CAS registry per year. Source: CAS registry.

FIGURE 9
Stripes: Number of unique substances registered in the CAS Registry per year. Source: CAS registry.

4.4. Monitoring project deliverables and experience

Publications are a fundamental part of the results of a research project, including the various parts of the publication process and the experience of the participants. However, the analysis of this process was not only to optimize the logistics of the project but to embrace a hermeneutic analysis based on interpretation, to understand how the publication processes differ among disciplines, what kind of publications predominate and why, and how it changes in more or less interdisciplinary, individual, or shared publications. We used data visualization not only for the purpose of quantitative analysis (e.g., how many papers had been published) but as a tool for close reading, to develop a deeper understanding of the process itself. In this last section of the results, we focus on *research question #4: How could we monitor the project deliverables, including the different steps of the process, the contribution of the different disciplines, and the experience of the participants?*

In previous examples, we have discussed two time-based visualizations: the evolution of the project scope around three checkpoints throughout our project (Figure 3), and a simplified heatmap to visualize the change of a variable (e.g., CAS registration numbers; Figure 9) over time. Next, we were interested in monitoring the progress of planned publications during the project. We wanted to know the start date of the work, the date of publication, and also the intermediate steps (e.g., when it is submitted and the progress toward acceptance). We also wanted to know which disciplines participated in each publication, whether authors from outside the project team were involved, and what type of publication it was (e.g., article, poster). The visualization technique most frequently used to represent time intervals is Gantt charts. However, a simple Gantt chart did not allow us to represent all the variables of interest in this analysis.

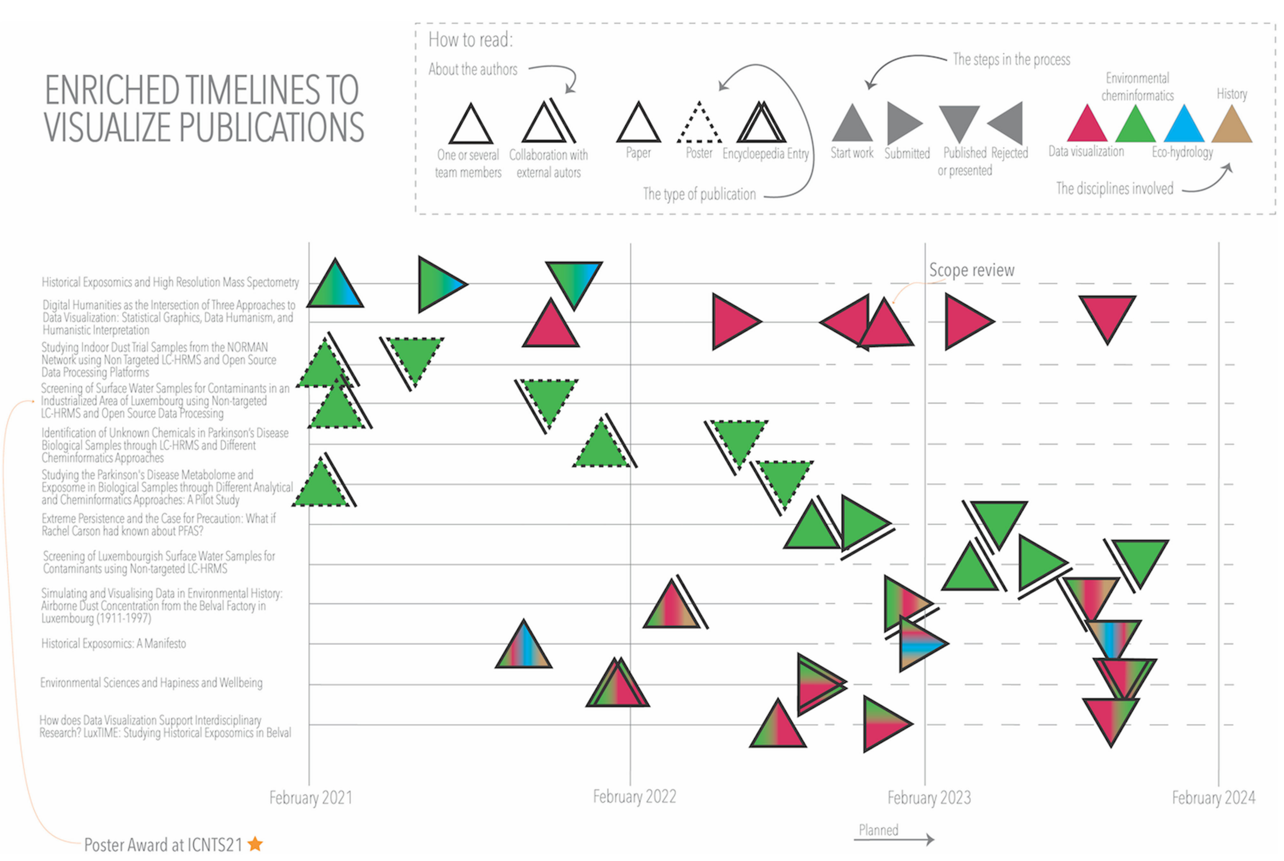


FIGURE 10
Team publications.

In Figure 10, we can see how we could enrich the information shown in a temporal diagram. By combining multivariate data glyphs and placing them at specific points on the timeline, we could divide the process into different steps between the time of starting the work and the publication. We displayed a timeline for each publication, with a triangle at each stage of the process, rotating to the right when the process moves forward (e.g., submission), and to the left when it "goes backwards" (e.g., rejection/revision leading to a restart of the submission process). The color of the triangles indicates that there are individual publications for each discipline, as well as combinations of two or more disciplines. Publication types for environmental cheminformatics (green) include three posters (dashed triangles), seven articles, and an encyclopedia entry. We can also see that in five publications, authors from outside the project are involved. This visualization allows us to see at which moments work accumulates and why. The rotation effect of the triangle helped us to understand parts of the process, for example, the effect of "going backwards" even as time moved forward, due to a rejection that required a restart. At the same time, the rotation of the triangle prevented us from seeing when the event occurred with precision (i.e., vertex or center), but as in this visualization, we did not need exact dates; we chose to keep the triangular shape and take advantage of the rotation effect.

The time scale could be visualized in more detail (e.g., with monthly details), adding more steps to the process (e.g., differentiate acceptance and publication), or factoring in the time perception of different participants, using non-standard intervals. For example, one of the authors might have perceived the initial time spent working on a publication differently from the time spent working after the rejection of the initial version (see Figure 11). The visualization in Figure 11 highlights the difference between time and temporality, the latter being relational. It opens the visualization techniques to graphical methods that represent experiential temporality, a subjective experience that depends on many psychological and physiological factors. If different participants of the project were to repeat the visualization for different publications, the challenges related to the different points of view of the multiple sources would need to be accounted for as discursive temporality. The visualization of time—based on experience—allows us to explore how the different participants experience the project (e.g., which moments are perceived as most stressful). This visualization technique, where the timeline is not standard, is probably the one that takes epistemological differences the furthest.

5. Discussion

In this article, we have presented the LuxTIME *data visualization toolbox*, including several data visualization concepts and techniques from epistemologically distant disciplines.

10.3389/fdata.2023.1164885

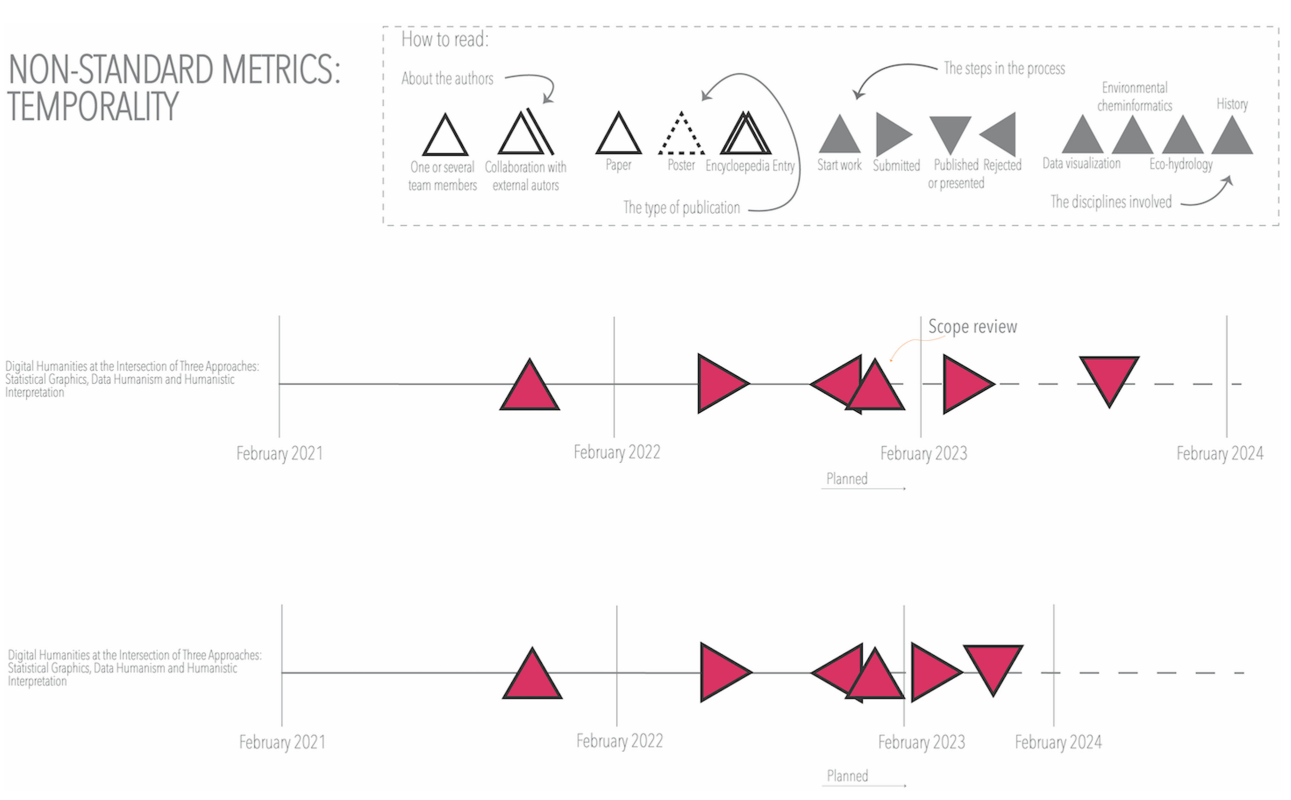


FIGURE 11
Non-standard time intervals.

This *toolbox* has facilitated an interdisciplinary collaboration among researchers in history, environmental cheminformatics, eco-hydrology, and data visualization, learning and applying concepts and techniques that reflect the paradigms of the other disciplines (e.g., integrating the use of rhetoric or temporality in natural sciences). Through the data collection process, the numerous exchanges on data visualization, the sketching and prototype development sessions, and the feedback collection and implementation, we have experienced and learned about the frictions inherent in interdisciplinary work.

First, the different perspectives on the level of granularity are worth noting, a friction already described by Panagiotidou et al. (2022). Especially, in the visualization process, we had several discussions about how important the details were in the visualizations (i.e., how we reached the results through several iterations with positive and challenging experiences: failure, team changes, and learning opportunities) vs. just visualizing the result. Second, by integrating the ideas of *data humanism* and *interpretation* into our *toolbox*, we increased the time needed to explore some of the visualizations, which often triggered the "discomfort" of not being able to immediately arrive to a clear conclusion. For example, in multivariate visualizations, there are several levels of information that cannot be extracted at first glance, but they allow us to show many facets of the metadata in a single view if we are willing to spend more time in the exploration. Third, the most accepted concepts were the integration of rhetorical mapping, and the variations of statistical graphics, notably the minimalist approach (i.e., no axes, tick marks, or grids), probably because it already had a strong precedent with the climate stripes.

Finally, one of the elements that we added to the *toolbox* at a later stage, because the need emerged, was the correct use of statistical graphics. Although statistical graphics were often used in all the disciplines involved, the visualizations generated did not always respect established theories on data visualization research. These discussions (e.g., color theory) highlighted the need to collaborate with researchers in data visualization so as not to perpetuate errors within the disciplines. Moreover, the fact of including concepts and techniques that serve different objectives (e.g., drawing quick conclusions using perception theory vs. exploring in-depth multiple narratives through multivariate visualizations with custom visual vocabularies) in our *toolbox*, created a series of discussions about how blurred the lines are between such "tools" and the paradigms they come from.

Throughout the research, the idea of data visualization as a *sandcastle* has been very present, especially the use of data visualization as part of the *speculative process*, and not just to present the results once the work is done (Hinrichs et al., 2019). This research is an attempt to use data visualization not only to explore and communicate about the domain data and metadata but also to collect and explore forms of thinking and creating knowledge through visual means. To make this visual means more specific, we have collected a series of existing concepts and techniques, our *toolbox*, to encourage participants to rethink what they already know but also to defamiliarize themselves with the usual methods and use data visualization as an *aesthetic provocation* (Hinrichs et al., 2019) to open up new perspectives for their own disciplines and the interdisciplinary work at LuxTIME. Throughout the process, visualization in its role as a *mediator* promotes an open and critical discourse (Hinrichs et al., 2019).

All the visualization concepts and techniques discussed in this article are just a proposal for a *data visualization toolbox* suitable for many research fields. Such concepts and techniques are not new but are rarely combined across disciplines or within the framework of a single project. In LuxTIME, we wanted to experiment with different data visualization techniques in a practical way, through our own process of interdisciplinary learning and exchange, during the definition of the project, and to explore the data and metadata found in relation to our main theme: historical exposomics in the Minett region. The aim was to extend the "go-to" *data visualization toolbox* and to explore, validate, and communicate, benefiting from techniques researched and applied across different disciplines. The use of a variety of techniques allowed us to look at data from different perspectives, analyzing the process quantitatively, qualitatively, and in an interpretative manner (e.g., combining Gantt charts with more flexible and detailed views, using non-standard timelines to express the experience of participants), and alternating and combining the use of statistical graphs with the use of metaphors or other graphic elements, whose aim is not necessarily to communicate quickly and accurately but to foster emotions.

This combination of statistical graphics and their variations, the use of information design elements generally present in the so-called *data art*, as well as the integration of interpretative elements gave us an *extended data visualization toolbox* to navigate our project. Probably because of the interdisciplinary nature of the project, this combination was more obvious, but we believe it could be useful for projects of all kinds. Such a *playground* is necessary to be able to formally evaluate the combined use of these techniques in future. This exploratory practice refers to the notion of "thinkering", composed of the verbs tinkering and thinking that describes the action of playful experimentation with digital tools for the interpretation and presentation of history (Fickers and van der Heijden, 2020). There is no one-size-fits-all toolbox for every research project, as the toolbox concept is built on the idea of flexible rearrangements of tools depending on the research questions, needs, and aims of a project. The toolbox is therefore the result of a "co-design" process based on situated knowledge practices.

The visualizations presented in this article are prototypes that will evolve further toward the end of the project, incorporating other techniques (e.g., interactivity, direct visualization, and enhanced ways of storytelling). Moreover, after having experimented with several types of visualizations separately to address different research questions, future steps will include a review of the connections between different visualizations and how they are linked in the overall narrative.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: Data will be published at the end of the project in 2024. In this article we just discuss the data visualizations. Requests to access these datasets should be directed to aida.horanietibanez@uni.lu and dagny.aurich@uni.lu.

Ethics statement

Ethical approval was not required for the study involving human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was not required from the participants in accordance with the national legislation and the institutional requirements.

Author contributions

All authors participated in the conception and design of the study, performed the data collection, data analysis, data interpretation, drafting of the article, contributed to the manuscript revision, and read and approved the submitted version.

Funding

This research has been funded by the University of Luxembourg Institute for Advanced Studies Audacity Program for the Luxembourg Time Machine (LuxTIME).

Acknowledgments

The authors would like to thank Emma Schymanski and Andreas Fickers for the detailed feedback on this study and the LuxTIME team members (Christophe Hissler, Laurent Pfister, Reinhard Schneider, and Lars Wieneke) and Johanna Drucker who took part in the discussion on our research, as well as Jennifer Yang for providing the history of the CAS registry numbers.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Aigner, W., Miksch, S., Schumann, H., and Tominski, C. (2011). *Visualization of Time-Oriented Data*. London: Springer London.

Arp, H. P. H., Aurich, D., Schymanski, E. L., Sims, K., and Hale, S. E. (2023). Avoiding the next silent spring: our chemical past, present, and future. *Environ. Sci. Technol.* 57, 6355–6359. doi: 10.1021/acs.est.3c01735

Aurich, D., Miles, O., and Schymanski, E. L. (2021). historical exposomics and high resolution mass spectrometry. *Exposome* 1, osab007. doi: 10.1093/exposome/osab007

Cairo, A. (2019). *How Charts Lie: Getting Smarter About Visual Information*. 1st ed. New York, N.Y.: W.W. Norton and Company.

Chuah, M. C., and Eick, S. G. (1998). Information rich glyphs for software management data. *IEEE Comput. Graph. Appl.* 18, 24–29. doi: 10.1109/38.689658

Cleveland, W. S., and McGill, R. (1984). Graphical perception: theory, experimentation, and application to the development of graphical methods. *J. Am. Stat. Assoc.* 79, 531–554. doi: 10.1080/01621459.1984.10478080

Cleveland, W. S., and McGill, R. (1986). An experiment in graphical perception. *Int. J. Man-Mach. Stud.* 25, 491–500. doi: 10.1016/S0020-7373(86)80019-0

Cordell, R. (2019). “Teaching humanistic data analysis,” in *Digital Scholarship, Digital Classrooms: New International Perspectives on Research and Teaching: Proceeding of the Gale Digital Humanities Day at the British Library*.

D’Ignazio, C., and Klein, L. F. (2020). *Data Feminism*. Cambridge, MA: The MIT Press.

Dörk, M., Feng, P., Collins, C., and Carpendale, S. (2013). “Critical InfoVis: exploring the politics of visualization” in *CHI’13 Extended Abstracts on Human Factors in Computing Systems* (Paris France: ACM), 2189–2198.

Drucker, J. (2011). *Humanities Approaches to Graphical Display*. Boston, MA: DHQ. p. 5.

Drucker, J. (2018). Non-representational approaches to modeling interpretation in a graphical environment. *Digit. Scholarsh. Humanit.* 33, 248–263. doi: 10.1093/llc/fqx034

Drucker, J. (2021). *TimeCapsule Report on Research 2001-2021*. UCLA, Department of Information Studies and Design.

Dykes, B. (2020). *Effective Data Storytelling: How to Drive Change With Data, Narrative and Visuals*. Hoboken, NJ: John Wiley and Sons, Inc.

Evergreen, S. D. H. (2017). *Effective Data Visualization: The Right Chart for the Right Data*. Los Angeles, CA: SAGE.

Fickers, A., Tatarinov, J., Fickers, A., Tatarinov, J., and van der Heijden, T. (2022). “Digital history and hermeneutics – between theory and practice: an introduction,” in *Digital History and Hermeneutics: Between Theory and Practice* (Berlin: De Gruyter).

Fickers, A., and van der Heijden, T. (2020). *Inside the Trading Zone: Thinkering in a Digital History Lab*. Boston, MA: DHQ. p. 14. Available online at: http://digitalhumanities.org/dhq/vol/14/3/000472/000472.html (accessed September 05, 2023).

Glivinska, A. (2021). *The 25 Best Data Visualizations of 2023 [Examples]*. Visme. Available at: https://visme.co/blog/best-data-visualizations/ (accessed January 13, 2023).

Hall, K., Wm., Bradley, A. J., Hinrichs, U., Huron, S., Wood, J., et al. (2020). Design by immersion: a transdisciplinary approach to problem-driven visualizations. *IEEE Trans. Vis. Comput. Graph.* 26, 109–118. doi: 10.1109/TVCG.2019.2934790

Hill, C. A., and Helmers, M. (2009). *Defining Visual Rhetorics*. Reprinted from Lawrence Erlbaum ed. New York, NY: Routledge.

Hinrichs, U., Forlini, S., and Moynihan, B. (2019). In defense of sandcastles: Research thinking through visualization in digital humanities. *Digit. Scholarsh. Humanit.* 34, i80–i99. doi: 10.1093/llc/fqy051

Hullman, J., and Diakopoulos, N. (2011). Visualization rhetoric: framing effects in narrative visualization. *IEEE Trans. Vis. Comput. Graph.* 17, 2231–2240. doi: 10.1109/TVCG.2011.255

Karlsson, O., Rocklöv, J., Lehoux, A. P., Bergquist, J., Rutgersson, A., Blunt, M. J., et al. (2021). The human exposome and health in the Anthropocene. *Int. J. Epidemiol.* 50, 378–389. doi: 10.1093/ije/dyaa231

Kemman, M. (2021). *Trading Zones of Digital History*. Berlin: De Gruyter.

Kim, M. C., Zhu, Y., and Chen, C. (2016). How are they different? A quantitative domain comparison of information visualization and data visualization (2000–2014). *Scientometrics* 107, 123–165. doi: 10.1007/s11192-015-1830-0

Kim, N. W., Im, H., Riche, N. H., Gajos, K., and Pfister, H. (2018). *Fostering Data Humanism With DataPortraits: Empowering People to Create a Personalized Visual Vocabulary*. IEEE VIS Poster (VIS).

Kim, N. W., Im, H., Henry Riche, N., Wang, A., Gajos, K., and Pfister, H. (2019). “DataSelfie: empowering people to design personalized visuals to represent their data” in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow Scotland Uk: ACM), 1–12.

Knebeler, C., and Scuto, D. (2010). *Belval: Passé, présent et avenir d’un site luxembourgeois exceptionnel (1911-2011)*. Esch-sur-Alzette: Le Phare.

Lankow, J., Ritchie, J., and Crooks, R. (2012). *Infographics: The Power of Visual Storytelling*. Hoboken, N.J: John Wiley and Sons, Inc.

Lanzing, J. (1998). Concept mapping: tools for echoing the minds eye. *J. Vis. Lit.* 18, 1–14. doi: 10.1080/23796529.1998.11674524

Lima, M. (2014). *The Book of Trees: Visualizing Branches of Knowledge*. New York, NY: Princeton Architectural press.

Lockwood, A. (1969). *Diagrams: A Visual Survey of Graphs, Maps, Charts and Diagrams for the Graphic Designer*. London; New York, NY: Studio Vista; Watson-Guptill.

Loukissas, Y. A. (2019). *All Data are Local: Thinking Critically in a Data-Driven Society*. Cambridge, MA: The MIT Press. doi: 10.7551/mitpress/11543.001.0001

Lupi, G. (2017). *Data Humanism: The Revolutionary Future of Data Visualization*. PRINT Mag. Available online at: https://www.printmag.com/article/data-humanism-future-of-data-visualization/ (accessed January 10, 2023).

Manovich, L. (2011). What is visualisation? *Vis. Stud.* 26, 36–49. doi: 10.1080/1472586X.2011.548488

Mao, Y., Wang, D., Muller, M., Varshney, K. R., Baldini, I., Dugan, C., et al. (2019). How data scientists work together with domain experts in scientific collaborations: to find the right answer or to ask the right question? *Proc. ACM Hum.-Comput. Interact.* 3, 1–23. doi: 10.1145/3361118

Miller, G. W. (2020). *The Exposome - A New Paradigm for the Environment and Health*. Cambridge, MA: Academic Press.

Miller, G. W., and Jones, D. P. (2014). The nature of nurture: refining the definition of the exposome. *Toxicol. Sci.* 137, 1–2. doi: 10.1093/toxsci/kft251

Novak, J. D., and Cañas, A. J. (2008). *The Theory Underlying Concept Maps and How to Construct and Use Them. Tech. Rep. IHMC CmapTools*. Available online at: http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf (accessed September 05, 2023).

Nussbaumer Knaflic, C. (2015). *Storytelling With Data: A Data Visualization Guide for Business Professionals*. Hoboken, NJ: Wiley.

Panagiotidou, G., Gorucu, S., and Vande Moere, A. (2020). Data badges: making an academic profile through a DIY wearable physicalization. *IEEE Comput. Graph. Appl.* 40, 51–60. doi: 10.1109/MCG.2020.3025504

Panagiotidou, G., Poblome, J., Aerts, J., and Vande Moere, A. (2022). Designing a data visualisation for interdisciplinary scientists. How to transparently convey data frictions? *Comput. Support. Coop. Work CSCW.* 31, 633–667. doi: 10.1007/s10606-022-09432-9

Rhyne, T., Tory, M., Munzner, T., Ward, M., Johnson, C., and Laidlaw, D. H. (2003). “Information and scientific visualization: separate but equal or happy together at last” in *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* (Seattle, WA, USA: IEEE), 611–614.

Rogowitz, B. E., and Treinish, L. A. (1998). Data visualization: the end of the rainbow. *IEEE Spectr.* 35, 52–59. doi: 10.1109/6.736450

Romance, N. R., and Vitale, M. R. (1999). Concept mapping as a tool for learning: broadening the framework for student-centered instruction. *Coll. Teach.* 47, 74–79. doi: 10.1080/87567559909595789

Santoro, C., and Kirkland, C. (2022). The warming stripes - Inspiring a movement. *Nightingale Mag.* 2022, 52–60.

Schwabish, J. A. (2021). *Better Data Visualizations: A Guide for Scholars, Researchers, and Wonks*. New York, NY: Columbia University Press.

Schwan, H., Arndt, J., and Dörk, M. (2022). Disclosure as a critical-feminist design practice for Web-based data stories. *First Monday* 27, 11. doi: 10.5210/fm.v27i11.12712

Segel, E., and Heer, J. (2010). Narrative visualization: telling stories with data. *IEEE Trans. Vis. Comput. Graph.* 16, 1139–1148. doi: 10.1109/TVCG.2010.179

Silva, S., Madeira, J., and Santos, B. S. (2007). “There is more to color scales than meets the eye: a review on the use of color in visualization” in *2007 11th International Conference Information Visualization (IV’07)* (Zurich, Switzerland: IEEE), 943–950.

Silva, S., Sousa Santos, B., and Madeira, J. (2011). Using color in visualization: a survey. *Comput. Graph.* 35, 320–333. doi: 10.1016/j.cag.2010.11.015

Steen, G. (2018). *Visual Metaphor: Structure and Process*. Amsterdam; Philadelphia: John Benjamins Publishing Company.

Thrift, N. J. (2008). *Non-Representational Theory: Space, Politics, Affect*. 1st ed. London: Routledge.

Tufte, E. R. (1999). *The Visual Display of Quantitative Information*. 17th ed. Cheshire; Conn: Graphics Press.

Van Wijk, J. J. (2006). Bridging the gaps. *IEEE Comput. Graph. Appl.* 26, 6–9. doi: 10.1109/MCG.2006.120

Vora, S. (2019). *The Power of Data Storytelling*. Thousand Oaks, CA: SAGE Publications India Pvt Ltd.

Ward, M. O. (2008). "Multivariate Data Glyphs: Principles and Practice," in *Handbook of Data Visualization Springer Handbooks Comp.Statistics (Berlin, Heidelberg: Springer Berlin Heidelberg),* 179–198.

Ware, C. (2004). *Information Visualization: Perception for Design*. San Francisco, CA: Morgan Kaufman.

Wheeldon, J., and Faubert, J. (2009). Framing experience: concept maps, mind maps, and data collection in qualitative research. *Int. J. Qual. Methods* 8, 68–83. doi: 10.1177/160940690900800307

Conclusion and further perspectives

Breaking out of duality

For many years, there has been only two ways of doing data visualization: "rule-abiding" vs. "rule-breaking", "right" vs. "wrong", "scientific" vs. "artistic", and so many other mutually exclusive terms. A way to escape this restrictive duality has been to create new terms for everything that does not refer to the "true data visualization" practice, "the correct one", and "the scientific one", to escape the criticism for not adhering to the predefined. To be able to experiment with new methods and techniques, we speak of "information design" or "data art", although both use data represented through a series of visual attributes, and are therefore data visualizations, as much as a bar chart is. We often teach techniques such as creating custom visual vocabularies, as "other ways to visualize data", or "how to visualize data differently". In fact, I have entitled the gallery presented in Chapter 1 "alternative ways to visualize data". While research methods are accepted and taught on an epistemological continuum that encompasses quantitative, qualitative, critical, and creative methods; data visualization is not. These "alternative" forms, the "other way", is an enormous reduction of all that this category encompasses. It is a box of lost objects, where we put everything, often ideas with little relation between them, and that does not allow us to study them in detail. Even when other types of data visualization, beyond *statistical graphics*, are accepted and researched (e.g., visualisation rhetoric), their use is often justified exclusively in the context of data dissemination with aesthetic purposes relating to outcomes such as engagement or memorability. These are certainly reasons for the use of these approaches, but this implies that they have no place in other contexts, such as the exploration of scientific data.

It is difficult to know if "data art" has a place in science, if it is not taught, experimented, and researched in different contexts. Teaching different approaches to data visualization involves extending the definition of visualization literacy (beyond the understanding of statistical graphics). What happens when we use vocabularies to visualize data that the reader does not need to know beforehand? How long does it take for the reader to become familiar with this type of data visualization, to become "data literate"? Is it possible that we are all data literate by default, if the data display includes instructions on how to read it? In a recent project carried with a Thinkering grant[10]: "From The Historical Archive To The Citizens: Visualizing Census Data From Brill Street in 1922" (Aida Horaniet Ibañez et al., 2023) (see Figure 1), we hypothesized that experiencing the census data in a physical form and without statistical charts, would allow the current residents to engage with historical data in a new way. This brief experiment made me wonder what would happen if we moved away from the usual evaluation methods, where a data visualization is better the quicker a message is conveyed and with fewer errors, where there is always a trade-off between complexity and comprehension. To get out of this duality implies defining new evaluation metrics adapted to the specific needs of the project, which are not always efficiency and precision. What new metrics would be adequate and in which cases? Maybe some user experience metrics used in web analysis such as number and profile of the users, accessibility, time to interest, different paths followed by the user (multiple narrative exploration), on-visualization time (i.e., the longer the better), number of

---

[10] Thinkering is a word composed of the verbs thinking and tinkering which together convey a sense of playful experimentation. Each year the Luxembourg Centre for Contemporary and Digital History (C2DH) of the University of Luxembourg funds a handful of small-scale research projects that have a relatively high chance of 'failure' but foster 'creative uncertainty' and encourage researchers in different disciplines to collaborate.

conversations/questions triggered? Perhaps a metric that estimates the time until the visualization can be navigated i.e., time to literacy?

Another barrier yet to be overcome is the disciplinary silos in the use of data visualization. Despite the obvious historical and current interdisciplinarity in the field of data visualization, there is still a reluctance to cross disciplinary boundaries. Moreover, all disciplines can leverage different methods. The quantitative methods based on objective measurements for numerical analysis are not exclusive to the sciences, nor are critical methods of inquiry derived from an appreciation of human values, exclusive to the humanities. In Chapter 4 of this research, we have discussed an attempt to blur these boundaries, exploring different visualization techniques to visualize objective and subjective phenomena across the disciplines involved, not without friction. Through this experimentation and the proposed toolbox, we are not trying to justify the use of any technique in any context, but to consider more options. Concrete options that will allow us in the future to define a formal evaluation based on relevant metrics adapted to each project.

The underlying objective is to take out little by little all that we have put in that common box of "the other way", to be able to research it in detail, and in context. To help us understand what we have grouped in that category, in Chapter 1, a collection of data visualizations has been manually curated, because the use of automatic techniques would only reinforce the analysis of already studied collections (i.e., only research articles, published in specific journals, with a certain length, including the same comparable sections). To collect data visualizations using different approaches, inside and outside academia, dealing with topics in different disciplines, we are faced with a plurality of sources, currently difficult to automate. However, we need such collections for our research, which in the future can be extended, completed, and analysed from new perspectives. The proposed gallery, it is just the starting point.

It is the starting point to understand what tools support the creation of such visualizations and what remains to be developed. As I discussed in Chapter 1, there are many data visualization tools, but most of them serve a single approach, *statistical graphics*. To perform other types of visualizations, such as those discussed throughout this manuscript, the user needs advanced programming skills (and time for customization) or must sacrifice interactivity and connection to the data by using design tools. This entire manuscript, not only through the gallery of data visualizations, but also through the proposed classification of approaches across disciplines and practices, their practical implementation, and the proposal of an extended "data visualization toolbox", is an attempt to concretize how to integrate these "other" types of visualizations in research (and beyond). We can then formally evaluate them, incorporate them into the definition of visualization literacy, develop tools to support them, teach them, and finally stop referring to them as "the other way", “the wrong way”, “the alternative way”.

Advancing data visualization and interdisciplinary research

In the introduction, I have discussed at length the advantages and challenges of interdisciplinary research. Throughout the chapters, we have seen practical cases of its implementation, the use of data visualization and the frictions that arise. In recent years, an increasing number of studies call for making these frictions explicit and propose ways to address them through actionable design considerations (Panagiotidou, Poblome, et al., 2022). The integration of data visualization sessions is fundamental in interdisciplinary projects. These sessions might integrate training in techniques belonging to different approaches, design by immersion activities (Hall et al., 2020), participatory design and prototyping (Jänicke et al., 2020), or collaborative decision-making through data physicalisation (Cazacu et al., 2023), among others. The objective of these sessions is to use data visualization as a tool to understand the practices of the other participants, to learn about the other domains, to reflect and self-critique the participants' own practices, to develop a common language, to produce common knowledge, and to make explicit and overcome frictions that arise during the process.

Once participants in an interdisciplinary project have produced results by navigating such frictions, they often face the challenge of publishing. The different disciplines, governed by the rules within their epistemological practices, pose a barrier to the publication of interdisciplinary research. I have discussed in the introduction, the case of Data Visualisation for the Digital Humanities vs. the Digital Humanities for Data Visualization, each of them with their corresponding venues, on the visualization (e.g., IEEE VIS, ACM CHI Conference on Human Factors in Computing Systems) and the digital humanities' side (e.g., Digital Humanities Quarterly, Digital Scholarship in the Humanities). In LuxTIME, the publishing challenge extended to three disciplines in addition to data visualization: history (digital history and digital humanities), environmental cheminformatics (the exposome) and eco-hydrology. As a result, this manuscript includes 4 chapters corresponding to four publications published in four different journals and disciplines.

During this project, I have identified two main challenges when publishing articles where data visualizations have a central role, which I believe require dedicated future research. First, the lack of peer review on the content of the visualizations themselves, where most of the feedback concerns the text. There is an urgent need to discuss and develop best practices for peer review of data visualisations in scientific articles, not as images that enrich the content, but as a central piece of scholarship. Second, the restrictive rules on size and image quality accepted in the different journals that often do not allow detailed visualizations. This encourages the use of only *statistical graphics* where the data-ink ratio is low and therefore easy to integrate, and disfavours other types of information-dense visualizations, where image quality and size, are paramount for exploration. The alternative is usually to create a dedicated website to integrate the data visualizations and add the link to the article. This represents an important amount of additional work to develop and maintain it, and for example, in the case of static visualizations it does not add any value. There are new journals that embrace new submission formats more suitable for data visualization, such as the Journal of Digital History, where Chapter 3 has been published as a Jupyter notebook in Python, or the recently launched Journal of Visualization and Interaction (Journal of Visualization and Interaction, 2022) with an alternative experimental GitHub track.

Throughout this and many other research projects, data visualization has proven to be a powerful facilitation, exploration, and communication tool. It has a fundamental role in interdisciplinary projects, so it is necessary to formalize its integration throughout the project and continue to investigate new forms of practice across the entire epistemological continuum. In future research, an evaluation methodology could be developed to determine how interdisciplinary such practices really are. The evaluation could perhaps consider how many of the three proposed approaches the project engages in, how many methods and techniques are included in the project's *data visualisation toolbox*, how they support quantitative, qualitative, rhetorical, and creative inquiry, and how well this serves the needs of the project. It is also very important that the review of the data visualizations forms part of the peer review (in the context of the article) in the different journals. Finally, it is necessary to continue with the implementation of "visualization friendly" publication formats that allow exploring all types of visualizations, without requiring the creation of additional support material (e.g., website).

I would like to end this section by addressing more generally interdisciplinary research, and more specifically how it is learned. During the last three years, I have been exposed to numerous interdisciplinary experiences, not only in the context of the LuxTIME project, but also through collaborations with other faculties and institutions. Everything I have learned has been through hands-on experience, and this is how I think interdisciplinarity is mostly taught, as I have rarely encountered theoretical courses on the subject. However, looking back, I believe that an introductory theoretical training in interdisciplinary work would probably have facilitated the process. This training would ideally include at least an introduction to humanistic and scientific epistemologies, theories, and methods, to understand what constitutes knowledge and how it is created in the different disciplines. It would address the different types of collaboration and the roles involved, how to define common vocabularies, different forms of knowledge sharing, potential frictions and how to negotiate joint outcomes. It would also cover how to write and publish together, and where. Such theoretical training could be supported by the many existing practical examples, but I believe it would be of great help to any researcher embarking on an interdisciplinary project, especially between distant epistemologies.

LuxTIME: From Belval to a national case study

The Belval Time Machine was defined as a proof of concept for a potential future national project that would potentially cover other research topics. In this section, we present the lessons learned, and the recommendations for a future application at national level.

One of the most important challenges at the beginning of this project was knowing where to find relevant information. Therefore, our first contribution is an inventory of physical and digital archives and delegations, where we have found information related to our case study[11]. Even if this is only an initial discovery of the available information and its metadata (that could be further discussed and completed in future research), the list of sources will save a significant amount of time for future researchers in the same field. In view of a national project, such an inventory could be initiated from the institutions, so that researchers come to the project knowing where to start looking for information. In national research on topics other than the exposome, researchers being able to share where and what they have found in a database accessible to new researchers would be of great value in advancing research at national level. This may seem obvious, but such an infrastructure or collaboration does not currently exist in Luxembourg at either the local or national level, and it consumes a lot of resources in any research project, as we have experienced for ourselves. In our project, being a proof of concept, we have focused on the discovery of information to create an inventory as complete as possible in the study of the exposome. It was therefore not part of this research to digitize the sources found. Such digitization and its dissemination entail additional challenges (e.g., resolution, completeness, access rights). Later projects would advance these efforts from different perspectives. LuxTIME INITIATE, a one-year project, was aimed at creating a formal consortium of the main stakeholders for the study of historical data in Luxembourg, including the mapping of digitization efforts. The Doctoral Training Unit, Deep Data Science of Digital History (D4H)[12] builds on the learnings of previous projects, to continue to bridge research

---

[11] The source inventory with the mapping of the Exposome categories and subcategories can be found in the Appendix I. .

[12] Deep Data Science involves the University of Luxembourg's Centre for Contemporary and Digital History, the Faculty of Science, Technology and Medicine, the Faculty of Humanities, Education and Social Sciences; and the

between humanities and sciences, based on the concepts of "interdisciplinary trading zones" and "digital hermeneutics".

Second, based on the data inventory, we present a logical data model (see Figure 9 and Appendix II) that could serve as reference for future collections, to avoid the back and forth of adding metadata once the collection is already at an advanced stage. This structure could be transferred to any database for source management at the national level. However, it can also be used on a small scale by researchers in other projects to start organizing their sources. Specific to our research is the mapping of categories and subcategories of the exposome, and to the Luxembourgish context, the mapping and translation of different languages.

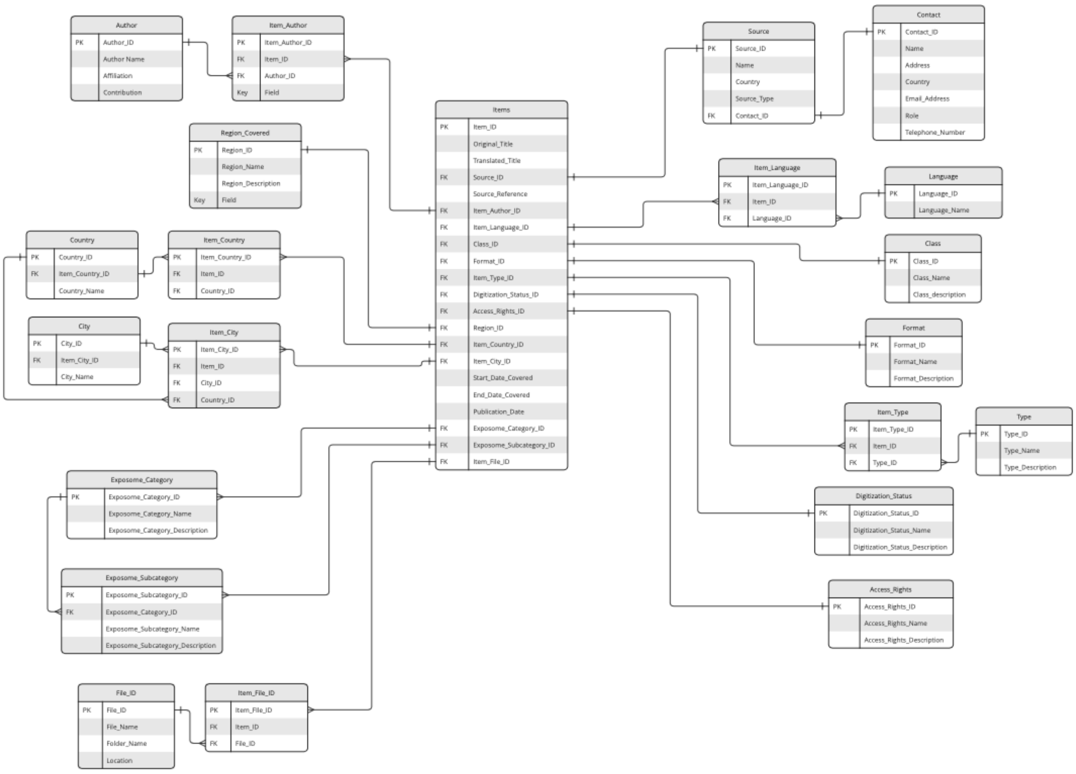


Figure 9 LuxTIME Entity Relationship Diagram

Finally, we propose the use of data visualization as a tool to facilitate, explore, communicate, and reflect on the process of any interdisciplinary project, be it at the local or national level. There are numerous tools ready to explore data and metadata with a statistical approach, as we have shown in Chapter 3. Such off-the-shelf data visualization tools and programming languages do not require new developments, and would be relatively simple to implement, for example the use of Python or R with Juypter notebooks, or tools such as Tableau. However, as I have argued throughout this manuscript, to use only statistical graphics just for exploration and communication is to consider only a very narrow part of the epistemological continuum, that of quantitative and perhaps some qualitative methods. Once a national database is in place, where researchers can find available sources on their topic of interest across different disciplines, the easiest way to explore that database is by integrating an existing visualization tool. There are still no standard tools that allow us to integrate data humanism or humanistic

---

Luxembourg Institute of Science and Technology and the Institute of Socio-Economic Research: https://www.c2dh.uni.lu/news/d4h-data-science-meets-digital-history.

interpretation, but that does not mean that we cannot benefit from their use. An important step is to spread the implementation of data visualization workshops at different stages of the interdisciplinary projects, where different techniques are explored with different objectives, even in analogue form if required. For example, creating a resource library to facilitate different types of data visualization workshops.

If we use the study of pollution as an example, a future researcher could open the national research data platform, where he/she could find different sources of information, including measurements made by research centres such as the LCSB or LIST, historical data on different types of industry and their production, legal data on the implementation of different environmental laws, data from public administration about official reports, historical maps, models on the dispersion of different types of pollution used in previous data simulations, etc. From that point, the researcher could go to the resource library on data visualisation to explore the data from different perspectives. Perhaps to start summarizing the data with a data visualization tool for *statistical graphics*, already integrated in the platform, to explore how many data sources are available, what periods they cover, what types of pollution they include, etc. Afterwards, the researcher could refer to a repository of visualisation workshops that would support other perspectives related to data humanism and interpretation, such as the simultaneous visualization of different pollution sources in different environments, the representation of people's experiences, or the role of the borders. This second step is difficult to explore using standard graphs, but it is equally relevant for an interdisciplinary approach to such analysis, and even its initial exploration in analogue form could be used as sketches for later digitisation.

Humanities theories and data visualization

In Chapter 1, I have argued the leading position of the digital humanities to benefit from and to advance the field of data visualization. In addition to developing new data visualization tools that integrate techniques from data humanism and humanistic interpretation, the digital humanities can (and must) continue to permeate humanities' theories and practices into the data visualization field. As discussed in Chapter 3, we propose a toolbox with different techniques and concepts in data visualization, but future versions could also include different analytical frameworks. Hermeneutics, and more specifically digital hermeneutics, provide a framework for the critical and self-reflective use of technologies, exploring how data visualization must adjust its use for interpretative activities (Kleymann & Stange, 2021). It enables historians to critically reflect on the various interventions of digital research infrastructures (Fickers & Tatarinov, 2022). Semiotics theory provides a framework for studying visual vocabularies used to encode data as well as any other elements of a data visualization as polysemic signs, to understand what they mean and why they have been selected beyond their literal meaning. Enunciation theory, as discussed by Drucker, can be applied to data visualization to explore the specific ways graphical features articulate subject positions (Drucker, 2017a). Critical theory attempts to reveal and challenge power structures. For example, feminist data visualization disavows binary distinctions; challenges claims of objectivity, neutrality and universalism; examines power; considers context; recognizes embodied and affective experiences as ways of knowing, and promotes working backwards to discover the actors involved in generating a particular dataset (D'Ignazio & Klein, 2016). Historiography explores the different perspectives from which the history of information visualization, data and statistics has been studied, and why certain angles and interpretations prevailed.

These and many other analytical frameworks allow us not only to evaluate data visualization from new perspectives, but also to concretize its implementation and development, advancing the field of data visualization in a way that is relevant to all disciplines, and therefore can better support interdisciplinary collaborations. All of these occupy an important place in the *data visualisation toolbox*, along with the concepts and techniques explored in this research, and others yet to come, adding to it.

Bibliography

Aida Horaniet Ibañez, Daniel Richter, & Joëlla Van Donkersgoed. (2023). *From The Historical Archive To The Citizens: Visualizing Census Data From Brill Street in 1922*. https://nightingaledvs.com/from-the-historical-archive-to-the-citizens-visualizing-census-data-from-brill-street-in-1922/

Alexander, J., Jansen, Y., Hornbæk, K., Kildal, J., & Karnik, A. (2015). Exploring the Challenges of Making Data Physical. *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, 2417–2420. https://doi.org/10.1145/2702613.2702659

Aurich, D., Miles, O., & Schymanski, E. L. (2021). Historical exposomics and high resolution mass spectrometry. *Exposome*, *1*(1), osab007. https://doi.org/10.1093/exposome/osab007

Barthel, C., Kirps, J., Centre d'études et de recherches européennes Robert Schuman, & Archives nationales (Luxembourg) (Eds.). (2009). *Terres rouges: Histoire de la sidérurgie luxembourgeoise*. Centre d'études et de recherches européennes Robert Schuman, Ministère d'état : Archives nationales de Luxembourg, Ministère de la culture, de l'enseignement supérieur et de la recherche.

Bederson, B. B., & Shneiderman, B. (2003). *The craft of information visualization: Readings and reflections*. Morgan Kaufmann.

Bellhouse, D. R. (2023). *The flawed genius of William Playfair: The story of the father of statistical graphics*. University of Toronto Press.

Beniger, J. R., & Robyn, D. L. (1978). *Quantitative Graphics in Statistics: A Brief History*. *32*(1).

Bertin, J., & Berg, W. J. (2010). *Semiology of graphics: Diagrams, networks, maps* (1st ed). ESRI Press : Distributed by Ingram Publisher Services.

Biltgen, F. (Ed.). (2005). *100 Joer Esch: 1906 - 2006*. Binsfeld.

Bonneuil, C., & Fressoz, J.-B. (2016). *The shock of the Anthropocene: The earth, history, and us*. Verso.

Boyd Davis, S., Vane, O., & Kräutli, F. (2021). Can I believe what I see? Data visualization and trust in the humanities. *Interdisciplinary Science Reviews*, *46*(4), 522–546. https://doi.org/10.1080/03080188.2021.1872874

C2DH, University of Luxembourg. (2021). *Journal of Digital History—Article Fingerprint*. https://journalofdigitalhistory.org/en/issue/jdh001

C2DH, University of Luxembourg. (2023). *Luxembourg Memorial de la Shoah*. https://memorialshoah.lu/en

Cairo, A. (2016). *The truthful art: Data, charts, and maps for communication*. New Riders.

Campbell, L. M. (2005). Overcoming Obstacles to Interdisciplinary Research. *Conservation Biology*, *19*(2), 574–577. https://doi.org/10.1111/j.1523-1739.2005.00058.x

Canali, S. (2020). What Is New about the Exposome? Exploring Scientific Change in Contemporary Epidemiology. *International Journal of Environmental Research and Public Health*, *17*(8), 2879. https://doi.org/10.3390/ijerph17082879

Cazacu, S., Mulder, I., Vande Moere, A., & Steenberghen, T. (2023). Revealing the role of values in developing a garden data ecosystem through a reflective participatory design approach. *The 11th International Conference on Communities and Technologies (C&T)*, 156–162. https://doi.org/10.1145/3593743.3593777

Clark, S. E. (2022). What's cookin. *Winner of the 2022 Data Visualization Society 'State of the Industry' Award for Most Beautiful Design*. https://www.sarahemeryclark.com/work/whats-cookin

Dataphys. (2021). *Dataphys Wiki*. http://dataphys.org

Denis Scuto. (2019). *Une histoire contemporaine du Luxembourg en 70 chroniques*. Fondation Robert Kriep, Center for Contemporary and Digital History.

D'Ignazio, C., & Klein, L. F. (2016). *Feminist Data Visualization*.

D'Ignazio, C., & Klein, L. F. (2020). *Data feminism*. The MIT Press.

Dima Yarovinsky. (2018). *Visualising Terms and Conditions*. https://www.designboom.com/readers/dima-yarovinsky-visualizes-facebook-instagram-snapchat-terms-of-service-05-07-2018/

Djavaherpour, H., Samavati, F., Mahdavi-Amiri, A., Yazdanbakhsh, F., Huron, S., Levy, R., Jansen, Y., & Oehlberg, L. (2021). Data to Physicalization: A Survey of the Physical Rendering Process. *Computer Graphics Forum*, *40*(3), 569–598. https://doi.org/10.1111/cgf.14330

Domestic Data Streamers. (2013). *The mood test*. https://domesticstreamers.com/projects/the-mood-test/

Domestic Data Streamers. (2014). *A visualization of life and death*. https://domesticstreamers.com/projects/lifeline/

Drucker, J. (2011). Humanities Approaches to Interface Theory. *Culture Machine*, *12*.

Drucker, J. (2017a). Information visualization and/as enunciation. *Journal of Documentation*, *73*(5), 903–916. https://doi.org/10.1108/JD-01-2017-0004

Drucker, J. (2017b). Non-representational approaches to modelling interpretation in a graphical environment. *Digital Scholarship in the Humanities*, *33(2)*. https://doi.org/10.1093/llc/fqx034

Drucker, J. (2020). *Visualization and interpretation: Humanistic approaches to display*. The MIT Press.

Drucker, J., Polack, P., & Santachiara, P. (2022). Heterochronologies: A platform for correlation and research in temporal graphics. *Digital Humanities Quarterly*, *16*(3).

Dulli Frühauf & Rene Kollwelter. (1987). *Umwelt Atlas für Luxembourg*. RTL Edition.

Few, S. (2012). *Show me the numbers: Designing tables and graphs to enlighten* (2. ed). Analytics Press.

Fickers, A., & Tatarinov, J. (Eds.). (2022). *Digital History and Hermeneutics: Between Theory and Practice*. De Gruyter. https://doi.org/10.1515/9783110723991

Frederica Fragapane. (2022). *Barriers to reporting*. https://www.behance.net/gallery/185533329/Barriers-to-reporting

Friendly, M. (2008). *Milestones in the history of thematic cartography, statistical graphics, and data visualization*.

Friendly, M., & Denis, D. J. (n.d.). *Milestones project*. Retrieved 12 November 2023, from https://www.datavis.ca/milestones/

Galison, P. (1996). Computer Simulations and the Trading Zone. In P. Galison & D. J. Stump (Eds.), *The Disunity of Science: Boundaries, Contexts, and Power* (pp. 118–157). Stanford University Press.

Gibbons, M., & Nowotny, H. (2001). The Potential of Transdisciplinarity. In J. T. Klein, R. Häberli, R. W. Scholz, W. Grossenbacher-Mansuy, A. Bill, & M. Welti (Eds.), *Transdisciplinarity: Joint Problem Solving among Science, Technology, and Society* (pp. 67–80). Birkhäuser Basel. https://doi.org/10.1007/978-3-0348-8419-8_7

Giorgia Lupi. (2018). *Connecting the Dots on Income Equality*. https://www.rand.org/about/nextgen/art-plus-data/giorgia-lupi/connecting-the-dots-on-income-inequality.html

Gitelman, L. (Ed.). (2013). *'Raw data' is an oxymoron*. MIT Press.

Goedert, J., Lorang, A., & Maas, J. (2023). *Paysages du fer: Mutations du bassin luxembourgeois-lorrain de la minette du XVIIIe au XXIe siècle*. Editions Gérard Klopp.

Hall, K. Wm., Bradley, A. J., Hinrichs, U., Huron, S., Wood, J., Collins, C., & Carpendale, S. (2020). Design by Immersion: A Transdisciplinary Approach to Problem-Driven Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *26*(1), 109–118. https://doi.org/10.1109/TVCG.2019.2934790

Hinrichs, U., Forlini, S., & Moynihan, B. (2019). In defense of sandcastles: Research thinking through visualization in digital humanities. *Digital Scholarship in the Humanities*, *34*(Supplement_1), i80–i99. https://doi.org/10.1093/llc/fqy051

Jänicke, S. (2016). Valuable Research for Visualization and Digital Humanities: A Balancing Act. *Workshop on Visualization for the Digital Humanities*, *7*, 5.

Jänicke, S., Franzini, G., Cheema, M. F., & Scheuermann, G. (2015). On Close and Distant Reading in Digital Humanities: A Survey and Future Challenges. In R. Borgo, F. Ganovelli, & I. Viola (Eds.), *Eurographics Conference on Visualization (EuroVis)—STARs*. The Eurographics Association. https://doi.org/10.2312/eurovisstar.20151113

Jänicke, S., Kaur, P., Kuzmicki, P., & Schmidt, J. (2020). Participatory Visualization Design as an Approach to Minimize the Gap between Research and Application. *VisGap - The Gap between Visualization Research and Visualization Software*, 8 pages. https://doi.org/10.2312/VISGAP.20201108

Jansen, Y., Dragicevic, P., Isenberg, P., Alexander, J., Karnik, A., Kildal, J., Subramanian, S., & Hornbæk, K. (2015). Opportunities and Challenges for Data Physicalization. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3227–3236. https://doi.org/10.1145/2702123.2702180

Jean Goedert, Lucien Steil, & Paul Weidig. (1991). *Al Esch. Déclin et Renaissance du Vieil-Esch*. Editions 'Al-Esch'.

Journal of Visualization and Interaction. (2022). *Journal of Visualization and Interaction*. https://www.journalovi.org

Karyda, M., Wilde, D., & Kjarsgaard, M. G. (2021). Narrative Physicalization: Supporting Interactive Engagement With Personal Data. *IEEE Computer Graphics and Applications*, *41*(1), 74–86. https://doi.org/10.1109/MCG.2020.3025078

Katri Huutoniemi, K. (2017). Evaluating interdisciplinary research. In *The Oxford Handbook of Interdisciplinarity* (pp. 309–320). Oxford University Press.

Kemman, M. (2021). *Trading Zones of Digital History*. De Gruyter. https://doi.org/10.1515/9783110682106

Kirk, A. (2016). *Data visualisation: A handbook for data driven design*. SAGE.

Kleymann, R., & Stange, J.-E. (2021). Towards Hermeneutic Visualization in Digital Literary Studies. *Digital Humanities Quarterly*.

Knebeler, C., & Scuto, D. (2010). *Belval: Passé, présent et avenir d'un site luxembourgeois exceptionnel (1911 - 2011)*. Éd. Le Phare.

Koeser, R. S., Doroudian, G., Budak, N., & Li, X. (2020). *Data Beyond Vision*. https://doi.org/10.5281/ZENODO.3713670

Laloux, F., & Wilber, K. (2014). *Reinventing organizations: A guide to creating organizations inspired by the next stage of human consciousness* (First edition (revised)). Nelson Parker.

Laurie Frick. (n.d.). *Childhood Hunger*. Retrieved 12 November 2023, from https://www.lauriefrick.com/childhood-hunger

Leboutte, R., Puissant, J., & Scuto, D. (1998). *Un siècle d'histoire industrielle, 1873-1973, Belgique, Luxembourg, Pays-Bas: Industrialisation et sociétés*. SEDES.

Liina Klauss, Skye Morét, & Moritz Stefaner. (2021). *Perpetual Plastic*. https://perpetual-plastic.net

Lima, M. (2014). *The book of trees: Visualizing branches of knowledge* (First edition). Princeton Architectural Press.

Lima, M. (2017). *The Book of Circles: Visualizing spheres of knowledge*. Princeton Architectural Press.

Lockton, D., Brawley, L., College, V., Ulloa, M. A., Prindible, M., Forlano, L., Rygh, K., Fass, J., Herzog, K., & Nissen, B. (2019). *Tangible thinking: Materializing how we imagine and understand systems, experiences, and relationships*.

Lorang, A. (1994). *Luxemburgs Arbeiterkolonien und billige Wohnungen: 1860 - 1940*. Ministère du Logement.

Loukissas, Y. A. (2019). *All data are local: Thinking critically in a data-driven society*. The MIT Press.

Lupi, G. (n.d.). *Giorgial Lupi*. http://giorgialupi.com

Lupi, G. (2017). *Data Items: A Fashion Landscape at The Museum of Modern Art*. http://giorgialupi.com/data-items-a-fashion-landscape-at-the-museum-of-modern-art

Lupi, G., & Posavec, S. (2016). *Dear data* (First edition). Princeton Architectural Press.

Lupi, G., & Posavec, S. (2015). *Dear Data Exhibition at MoMA*. https://www.moma.org/collection/works/215813

Marcel SCHROEDER. (1961). *Terres Rouges*. Bourg-Bourger.

Michela Lazzaroni. (2023). *The Poet's Journey*. https://www.behance.net/gallery/175499297/The-Poets-Journey-La-Lettura-dataviz

Miller, G. W., & Jones, D. P. (2014). The Nature of Nurture: Refining the Definition of the Exposome. *Toxicological Sciences*, *137*(1), 1–2. https://doi.org/10.1093/toxsci/kft251

Minett Stories. (2022). https://minett-stories.lu/en/story/air-pollution-visualized

Munzner, T. (2015). *Visualization analysis and design*. CRC Press.

Nadieh Bremer. (2020). *Space Wars*. https://www.visualcinnamon.com/portfolio/space-wars/

Nwosu, C. (2022). *Build Up Nepal for Viz for Social Good*. https://public.tableau.com/app/profile/chimdi.nwosu/viz/VizForSocialGood-BuildUpNepal/Dashboard1

Panagiotidou, G., Gorucu, S., & Vande Moere, A. (2020). Data Badges: Making an Academic Profile Through a DIY Wearable Physicalization. *IEEE Computer Graphics and Applications*, *40*(6), 51–60. https://doi.org/10.1109/MCG.2020.3025504

Panagiotidou, G., Lamqaddam, H., Poblome, J., Brosens, K., Verbert, K., & Moere, A. V. (2022). Communicating Uncertainty in Digital Humanities Visualization Research. *IEEE Transactions on Visualization and Computer Graphics*, 1–11. https://doi.org/10.1109/TVCG.2022.3209436

Panagiotidou, G., Poblome, J., Aerts, J., & Vande Moere, A. (2022). Designing a Data Visualisation for Interdisciplinary Scientists. How to Transparently Convey Data Frictions? *Computer Supported Cooperative Work (CSCW)*. https://doi.org/10.1007/s10606-022-09432-9

Pentagram. (2021). A Warming Planet. *Bulleting of the Atomic Scientists*. http://giorgialupi.com/bullettin-of-the-atomic-scientists

Perovich, L. J., Wylie, S. A., & Bongiovanni, R. (2021). Chemicals in the Creek: Designing a situated data physicalization of open government data with the community. *IEEE Transactions on Visualization and Computer Graphics*, *27*(2), 913–923. https://doi.org/10.1109/TVCG.2020.3030472

Rendgen, S. (2019). *History of information graphics*. Taschen.

Rhoten, D. (2004). Interdisciplinary Research: Trend or Transition. *Items and Issues*, *5*(1–2).

Sánchez, R. T., Santos, A. B., Vicente, R. S., & Gómez, A. L. (2019). Towards an Uncertainty-Aware Visualization in the Digital Humanities. *Informatics*, *6*(3), 31. https://doi.org/10.3390/informatics6030031

Schreibman, S., Siemens, R. G., & Unsworth, J. (2016). *A new companion to digital humanities*. Wiley/Blackwell.

Schwabish, J. A. (2021). *Better data visualizations: A guide for scholars, researchers, and wonks*. Columbia University Press.

Scuto, D. (1993). *Industriekultur in Esch: Eine stadtgeschichtliche Wanderung durch die Luxemburger Minettemetropole*. Le Phare.

Shirly Wu. (2019). *One Amongst Many*. https://oneamongstmany.com

Sonja Kujpers. (n.d.). *The Hotel New Hampshire*. Retrieved 12 November 2023, from http://www.studioterp.nl/the-hotel-new-hampshire-a-data-art-project-by-studio-terp/

Sonja Kujpers. (2019). *A View on Dispair*. https://www.studioterp.nl/a-view-on-despair-a-datavisualization-project-by-studio-terp/

South China Morning Post. (2020). *Arrested Hong Kong protesters*. https://multimedia.scmp.com/infographics/news/hong-kong/article/3088009/one-year-protest/index.html

Stefan Krebs. (2022). The Emergence of the Minett. *Minett Stories*. https://minett-stories.lu/en/story/the-emergence-of-the-minett

Stefanie Posavec. (2021). *Updating Happiness: Wellcome Collection*. https://www.stefanieposavec.com/updating-happiness

Stefanie Posavec. (2023). *Data Murmurations: Points in flight (People Like You)*. https://peoplelikeyou.ac.uk/system/

Stefanie Posavec & Miriam Quick. (2015). *Air Transformed*. https://www.stefanieposavec.com/airtransformed

Steffen, W., Grinevald, J., Crutzen, P., & McNeill, J. (2011). The Anthropocene: Conceptual and historical perspectives. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *369*(1938), 842–867. https://doi.org/10.1098/rsta.2010.0327

Swaminathan, S., Shi, C., Jansen, Y., Dragicevic, P., Oehlberg, L. A., & Fekete, J.-D. (2014). Supporting the design and fabrication of physical visualizations. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3845–3854. https://doi.org/10.1145/2556288.2557310

Tanzania dLab. (2018). *Data Zetu Project*. https://medium.com/data-zetu/young-artists-use-fashion-and-data-to-promote-dialog-on-sexual-health-517429662ec2

Testot, L., & Throssell, K. (2020). *Cataclysms: An environmental history of humanity*. University of Chicago Press.

Thierry Kruchten, Jessika Maria Rauch, Emile Hengen, Sabrina Notka, Marie Mathieu, Jennifer Müller, & Salomé Jeko. (2021). *Ein Streifzug durch den Süden Luxemburgs & das benachbarte Frankreich* (1. Auflage). Éditions Guy Binsfeld.

Trausch, G. (2000). *ARBED in the Luxembourg society*. ARBED.

Tufte, E. R. (1997). *Visual Explanations*. Graphics PR.

Tufte, E. R. (1999). *The visual display of quantitative information* (17. print). Graphics Press.

Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley Pub. Co.

Undocumented Migration Project. (2019). *Hostile Terrain 94: Deaths at the US/Mexico Border*. https://hostileterrain94.wordpress.com

Vermeulen, R., Schymanski, E. L., Barabási, A.-L., & Miller, G. W. (2020). The exposome and health: Where chemistry meets biology. *Science*, *367*(6476), 392–396. https://doi.org/10.1126/science.aay3164

Victor Eischen & Roger Fournelle. (1986). *Esch-sur-Alzette à la belle époque*. Imprimerie Saint-Paul.

Voit, P., Francke, T., & Bronstert, A. (2023). Accounting for operational irrigation options in mesoscale hydrological modelling of dryland environments. *Hydrological Sciences Journal*, *68*(5), 670–684. https://doi.org/10.1080/02626667.2023.2187296

Warde, P., Robin, L., & Sörlin, S. (2018). *The environment: A history of the idea*. Johns Hopkins University Press.

Ware, C. (2004). *Information visualization: Perception for design*. Morgan Kaufman.

Wenger-Trayner, É. (2008). *Communities of practice: Learning, meaning, and identity* (18th printing). Cambridge University Press.

Wild, C. P. (2005). Complementing the Genome with an 'Exposome': The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology. *Cancer Epidemiology Biomarkers & Prevention*, *14*(8), 1847–1850. https://doi.org/10.1158/1055-9965.EPI-05-0456

Wimmer, A. (2023). Layered Legacies. How Multiple Histories Shaped the Attitudes of Contemporary Europeans. *Sociological Science*, *10*, 1–46. https://doi.org/10.15195/v10.a1

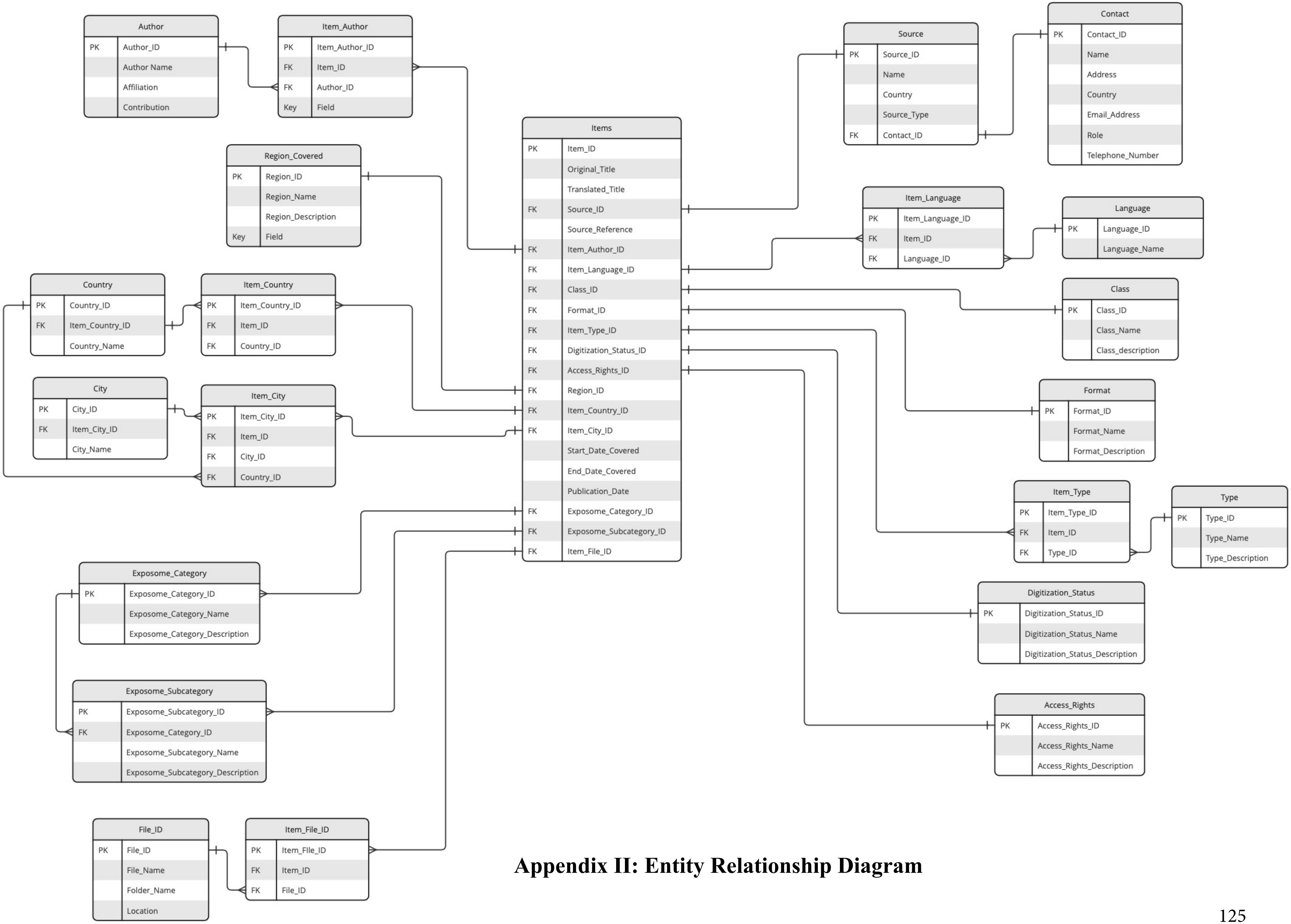
Appendix I: Information inventory and exposome mapping

ID	Title	English Title	Source	Language	Description	Digitised	Class	Format	Cover_Type	From	To	Geo_City	Geo_Country	Exposome_Category	Exposome_Subcategory
1	Complaint letter	Complaint letter	Luxembourg National Archive	DE	Report on a complaint about water pollution.	NO	Newspaper	Newspapers and Magazines	Date	24/06/1931	24/06/1931	Wiltz	Luxembourg	Physical/Chemical	Surface Water Contamination
2	Rapports concernant la pollution des cours d'eau (Wiltz; Alzette...) - rapport sur la défense des cours d'eau contre la pollution des usines - projet de loi concernant le curage, l'entretien et l'amélioration des cours d'eau - arrêté du 9 septembre 1929	Reports concerning the pollution of watercourses (Wiltz; Alzette...) - report on the defense of watercourses against pollution from factories - bill concerning the cleaning, maintenance and improvement of watercourses - order of September 9, 1929	Luxembourg National Archive	DE	Report of the 'oberförster'/head forester about the river pollution of the Alzette.	NO	Letter	Official Reports	Date	05/10/1929	05/10/1929	Mersch	Luxembourg	Physical/Chemical	Surface Water Contamination
3	Burage, entretien et amelioration des cours d'eau	Burying, maintenance and improvement of waterways	Luxembourg National Archive	FR	Law of 16 May 1929. Decree of 9 September 1929 on the purification of waste water.	NO	Text	Official Reports	Date	25/04/1929	25/04/1929	None	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories, Surface water contamination
4	En Dag an der Minière	A day in the mine	Luxembourg National Archive	FR	Exhibition on more than one hundred years of mining tradition. Inauguration in Belvaux.	NO	Newspaper	Newspapers and Magazines	Date	19/09/1991	19/09/1991	Belvaux	Luxembourg	Ecosystems, Physical/Chemical	Built environment/ urban landuses, Point/ line sources; e.g. factories
5	Reichs Gesetzblatt Nr. 21	Imperial Law Gazette No. 21	Luxembourg National Archive	DE	Legislation change in 'Reichsbeamtengesetz' from 1873.	NO	Text	Newspapers and Magazines	Date	24/05/1907	24/05/1907	Düsseldorf	Germany,Luxembourg	Social	Legislative measures
6	Königliche Großherzogliche Gendarmerie Compagnie - Verzeichnis	Royal Grand Ducal Gendarmerie Company - Directory	Luxembourg National Archive	FR,DE	Industry inventory Esch-sur-Alzette	NO	Document	Official Reports	Date	30/11/1887	30/11/1887	Esch-sur-Alzette	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories
7	Historische Skizze Über der Stadt Esch anlässlich der Erhebung zur Stadt	Historical sketch About the town of Esch on the occasion of its elevation to the status of a city	City Archive of Esch-sur-Alzette	DE	Historical development of city Esch (sketch).	NO	Text	Official Reports	Period	700	1871	Esch-sur-Alzette	Luxembourg	Ecosystems	Built environment/ urban landuses
8	Das Buch der Geschicke und Geschichte der Minettgegend	The book of the fate and history of the Minett area	City Archive of Esch-sur-Alzette	DE	Description of life and accomplishments of August & Norbert Metz (iron industry Metz & Cie).	NO	Book	Books and Research Articles	Period	1812	1917	Minett Region	Luxembourg	Social, Physical/Chemical	Social networks, Point/ line sources; e.g. factories
9	Plan d'Esch-s-Alzette. Métropole du basin minier luxembourgeois	Map of Esch-s-Alzette. Metropolis of the Luxembourg mining basin	City Archive of Esch-sur-Alzette	FR	Map of Esch.	NO	Map	Maps	Year	1982	1982	Esch-sur-Alzette	Luxembourg	Ecosystems	Built environment/ urban landuses
10	Unabhängigkeitsfeier des Katons Esch	Independence celebration of the caton Esch	City Archive of Esch-sur-Alzette	DE	Esch city between 1839 and 1939.	YES	Magazine	Newspapers and Magazines	Period	1839	1939	Esch-sur-Alzette	Luxembourg	Ecosystems, Social, Physical/Chemical	Built environment/ urban landuses, Health infrastructure, Population density, Green Spaces, Blue Spaces, Social capital, Social networks, Cultural norms, Cultural capital, Point/ line sources; e.g. factories
11	Plan Esch-sur-Alzette 1825-1956	Map Esch-sur-Alzette 1825-1956	City Archive of Esch-sur-Alzette	FR	Evolution of the city of Esch.	NO	Map	Maps	Period	1825	1956	Esch-sur-Alzette	Luxembourg	Ecosystems	Built environment/ urban landuses
12	Umwelt Atlas für Luxembourg	Environment Atlas for Luxembourg	City Archive of Esch-sur-Alzette	DE	Water systems in Luxembourg, groundwater, drinking water, water networks.	YES	Book	Books and Research Articles	Period	1960	1986	None	Luxembourg	Ecosystems, Physical/Chemical	Blue Spaces, Drinking water contamination, Groundwater contamination, Surface water contamination
13	Krankenhaus Korrespondenz 1928-1937	Hospital correspondence 1928-1937	City Archive of Esch-sur-Alzette	DE	Report on bad conditions in worker housings, bad health conditions, building of hospital in Esch with seperated tuberculosis/infectious diseases stations.	NO	Document	Official Reports	Period	1907	1937	Esch-sur-Alzette	Luxembourg	Ecosystems, Phenotype	Health infrastructure, Disease
14	Die Escher Waldschule als Kindertagesheim	The Esch Forest School as a day care center for children	City Archive of Esch-sur-Alzette	DE	Description of the benefits and neccessity of a forest school in Esch.	NO	Document	Official Reports	Year	1930	1930	Esch-sur-Alzette	Luxembourg	Ecosystems, Lifestyle, Physical/Chemical	Green Spaces, Physical Activity, Odor and noise, Outdoor and indoor air pollution
15	Bulletin d'information et d'action sanitaires N°3	Health Information and Action Bulletin N°3	STATEC Archive	FR	Statistics about sicknesses and death rates by year, region and age range.	NO	Text,Dataset,Visualization	Official Reports	Period	1963	1966	None	Luxembourg	Phenotype	Disease
16	Bulletin d'information et d'action sanitaires N°18	Health Information and Action Bulletin N°18	STATEC Archive	FR	Monthly concentrations in different cities including Ech/Alzette.	NO	Text,Dataset,Visualization	Official Reports	Period	1972	1973	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution
17	Bulletin d'information et d'action sanitaires N°19	Health Information and Action Bulletin N°19	STATEC Archive	FR	Monthly concentrations in different cities including Ech/Alzette, measures to protect the environment, water pollution.	NO	Text,Dataset,Visualization	Official Reports	Period	1973	1974	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution, Surface water contamination
18	Les composes du fluor dans la pollution atmospherique	Fluorine compounds in air pollution	Luxembourg National Archive	FR	Study of the emissions of the luxembougish steel industry and X-ray diffraction and electron microscopy.	NO	Text,Dataset,Visualization,Map	Official Reports	Year	1971	1971	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution
19	Parliamentary debate 1976 session 53	Parliamentary debate 1976 session 53	Missing	DE	Minutes of the parliamentary debate 1976.	NO	Text	Official Reports	Year	1976	1976	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution, Drinking water contamination, Groundwater contamination, Surface water contamination
20	Revue Technique Luxembourgeois 1972	Luxembourg Technical Review 1972	Missing	FR	Report about the corrosivity of the atmosphere in the mining basin, raw materials, evolution of the techniques and production capacity and quality.	NO	Book	Books and Research Articles	Year	1972	1972	None	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories, Outdoor and indoor air pollution
21	Communication about the responsible expert checking water and air pollution	Communication about the responsible expert checking water and air pollution	Luxembourg National Archive	FR	Communication about the responsible expert checking water and air pollution.	NO	Letter	Official Reports	Year	1973	1973	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution, Drinking water contamination, Groundwater contamination, Surface water contamination
22	Klärverfahren für Gerbereiabwasser	Clarification process for tannery wastewater	Luxembourg National Archive	DE	Wastewater treatment for tannery sewage.	NO	Letter	Official Reports	Year	1930	1930	None	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories, Surface water contamination
23	Carte des chemins de feu et des bassins miniers	Map of fire roads and coalfields	Luxembourg National Archive	FR	Map of the mines and railways in Differdange, Belvaux, Esch/Alzette, Dudelange, Rumelange.	NO	Map	Maps	Missing	Missing	Missing	None	Luxembourg	Ecosystems, Physical/Chemical	Built environment/ urban landuses, Point/ line sources; e.g. factories
24	Neuzeitliche Abwässer-Reinigung	Modern wastewater purification	Luxembourg National Archive	DE	Description of wastewater treatment procedure.	NO	Text	Official Reports	Year	1930	1930	None	Luxembourg,Germany	Physical/Chemical	Point/ line sources; e.g. factories, Surface water contamination
25	Treatment of Municipal Sewage	Treatment of Municipal Sewage	Luxembourg National Archive	EN	The applications of Dorr Equipment to modern sewage treatment and water purification plants.	NO	Book	Books and Research Articles	Year	1930	1930	None	United States	Physical/Chemical	Point/ line sources; e.g. factories, Surface water contamination
26	Images Source Belval	Images Source Belval	Archive of the Municipality of Sanem	FR	Postcards, advertisements, photographs from Source Belval in 1907, 1912, 1925, 1929 and when they reopened it in 2000 and 2010.	YES	Image	Photographs and videos	Period	1907	2010	Belval	Luxembourg	Ecosystems, Lifestyle, Phenotype	Blue Spaces, Diet, Disease
27	Belval Metzerlach Reflets 1960-1980 20e anniversaire Cercle Vocal	Belval Metzerlach Reflets 1960-1980 20th anniversary Cercle Vocal	Archive of the Municipality of Sanem	FR,DE	Anniversary of Cercle Vocal Belval - Metzerlach 1960 - 1980, Source Belval.	YES	Magazine	Newspapers and Magazines	Period	1960	1980	Belval	Luxembourg	Ecosystems, Lifestyle, Social, Phenotype	Blue Spaces, Diet, Social networks, Disease
28	Belval Neighbors. Bieleser Hausnimm	Belval Neighbors. Bieleser House Numbers.	Archive of the Municipality of Sanem	DE	List of families living in different streets.	YES	Document	Non-official document	Period	1920	1930	Belval	Luxembourg	Ecosystems, Social	Built environment/ urban landuses, Population density, Social networks
29	La source Bel-Val. Un petit historique	The Bel-Val spring. A little history.	Archive of the Municipality of Sanem	FR	Key people and dates, water schemas, international sales, photographs, closing, potential reopening of source Belval	YES	Document	Official Reports	Period	1855	2010	Belval	Luxembourg	Ecosystems, Lifestyle, Social, Phenotype	Blue Spaces, Diet, Social networks, Disease
30	D'Seelebunn	D'Seelebunn (Air cargo carrier).	Archive of the Municipality of Sanem	FR,DE	History of the Seelebunn, aerial transport between Ottance and Differdange.	YES	Book	Books and Research Articles	Period	1906	1980	Belvaux	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories
31	Bieles. Gemeng Suessem 1900-1929 Sectioun C	Municipality of Sanem 1900-1929 Section C.	Archive of the Municipality of Sanem	LU	Map of Sanem.	YES	Map	Maps	Period	1900	1929	Sanem	Luxembourg	Ecosystems	Built environment/ urban landuses
32	Photographs Sanem 1940-1994	Photographs Sanem 1940-1994.	Archive of the Municipality of Sanem	LU	Photos of Sanem from 1940 to 1990.	YES	Image	Photographs and videos	Period	1940	1990	Sanem	Luxembourg	Ecosystems, Physical/Chemical	Built environment/ urban landuses, Green Spaces, Point/ line sources; e.g. factories
33	ARBED in the Luxembourg Society	ARBED in the Luxembourg Society.	Nainuwa C2DH Digital Archive System	EN	Photo documentary of the ARBED, key actors.	YES	Book	Books and Research Articles	Period	1911	1990	None	Luxembourg	Social, Physical/Chemical	Social networks, Point/ line sources; e.g. factories
34	Der Luxemburger Bergmann in seiner Arbeitswelt	The Luxembourg miner in his working environment.	Nainuwa C2DH Digital Archive System	DE	Evolution of the Minette Region (geology aspects, mine workers life).	YES	Book	Books and Research Articles	Period	1840	1997	None	Luxembourg	Ecosystems, Lifestyle, Social, Physical/Chemical	Built environment/ urban landuses, Job, Social networks, Social capital, Occupational exposures

35	Industriekultur in Esch	Industrial culture in Esch.	Nainuwa C2DH Digital Archive System	DE	Culture, housing, hospitals and schools (Waldschule) and buildings including maps of urbanism and evolution and historical timeline.	YES	Book	Books and Research Articles	Period	770	1993	Esch-sur-Alzette	Luxembourg	Ecosystems, Social, Physical/Chemical	Built environment/ urban landuses, Health infrastructure, Green Spaces, Cultural capital, Point/ line sources; e.g. factories
36	Legacy of the Past	Legacy of the Past.	Nainuwa C2DH Digital Archive System	EN	Industrial tourism south of Luxembourg: museums, worker neighborhoods, blast furnaces, innovation .	YES	Book	Books and Research Articles	Missing	Missing	Missing	None	Luxembourg	Ecosystems, Social, Physical/Chemical	Built environment/ urban landuses, Green Spaces, Social networks, Cultural capital, Point/ line sources; e.g. factories
37	Luxemburger Hüttenwerke. Im Wandel der Zeit	Luxembourg Steelworks. In the course of time.	Nainuwa C2DH Digital Archive System	DE	History of the Luxembourgish iron and steel works.	YES	Book	Books and Research Articles	Missing	Missing	Missing	None	Luxembourg	Ecosystems, Lifestyle, Social, Physical/Chemical	Built environment/ urban landuses, Job, Point/ line sources; e.g. factories
38	The Rise of Luxembourg from Independance to Success	The Rise of Luxembourg from Independance to Success.	Nainuwa C2DH Digital Archive System	EN	History of Luxembourg: medieval origins, independent country, steel industry, world wars I and II, banks, european institutions .	YES	Book	Books and Research Articles	Period	1815	2015	None	Luxembourg	Ecosystems, Lifestyle, Social, Physical/Chemical	Built environment/ urban landuses, Population density, Green Spaces, Job, Social capital, Cultural capital, Point/ line sources; e.g. factories
39	Census Data Bassin Minier BAMI 1900	Census data mining basin.	Human Migration Documentation Centre	DE	Census data including family members full names, profesions, gender, marital status, nationality, address, birthdates and religion.	YES	Dataset	Non-official document	Period	1895	1919	Esch-sur-Alzette, Differdange, Rumelange, Dudelange	Luxembourg	Lifestyle, Social	Job, Social networks, Cultural norms
40	Carte Peletier 1928	Peletier Map 1928.	Luxembourg National Museum of Iron Mines	FR	Map of mines of Differdange and Esch-Rumelange in 1928 (border with France).	YES	Map	Maps	Year	1928	1928	Differdange, Esch-sur-Alzette, Rumelange	Luxembourg	Ecosystems	Built environment/ urban landuses
41	Map of closed large landfills	Map of closed large landfills.	Geoportail.lu	EN	Map of the closed large landfills in Luxembourg.	YES	Map	Maps	Missing	Missing	Missing	None	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories, Soil contamination, Groundwater contamination
42	Sampling Location with coordinates based on the Luxembourg Reference Frame	Sampling Location with coordinates based on the Luxembourg Reference Frame.	Geoportail.lu, Water Management Administration	EN	Surface water sampling positions of AGE.	YES	Map	Maps	Year	2019	2019	None	Luxembourg	Ecosystems, Physical/Chemical	Blue Spaces, Surface water contamination
43	Waste water treatment plants	Waste water treatment plants.	Geoportail.lu	EN	Map of the waste water treatment plants in Luxembourg.	YES	Map	Maps	Missing	Missing	Missing	None	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories, Surface water contamination
44	Esch-sur-Alzette Karte 1918	Esch-sur-Alzette Map 1918.	Missing	DE	Map of Esch sur Alzette in 1918.	YES	Map	Maps	Year	1918	1918	Esch-sur-Alzette	Luxembourg	Ecosystems	Built environment/ urban landuses, Walkability
45	Inventaire achitectural de la ville d'Esch	Architectural inventory of the city of Esch.	Missing	FR	Map of Esch sur Alzette in 1906 - 1956.	YES	Map	Maps	Period	1906	1956	Esch-sur-Alzette	Luxembourg	Ecosystems	Built environment/ urban landuses, Walkability
46	Dorf Esch sur Alzette 1842	City Esch sur Alzette 1842.	Missing	DE	Map of Esch sur Alzette in 1842.	NO	Map	Maps	Year	1842	1842	Esch-sur-Alzette	Luxembourg	Ecosystems	Built environment/ urban landuses, Walkability
47	Zukunftsplan Stadtbauplan Skizze für die Stadt Esch-sur Alzette	Future plan urban development plan sketch for the city of Esch-sur Alzette.	Missing	DE	Plan for the construction of the city of Esch sur Alzette.	YES	Map	Maps	Missing	Missing	Missing	Esch-sur-Alzette	Luxembourg	Ecosystems	Built environment/ urban landuses, Walkability, Green Spaces
48	25e Anniversaire de la Ligue National Luxembourgeoise du coin de terre et du foyer	25th Anniversary of the Luxembourg National League of Land and Home.	City Archive of Esch-sur-Alzette	FR,DE	Initiative to give a piece of land to the workers. Social and economic work. International movement and achivements in Luxembourg. National Congress in Esch in 1951.	YES	Document	Newspapers and Magazines	Period	1929	1951	None	Luxembourg	Ecosystems, Social	Built environment/urban landuses, Green spaces, Inequality, Legislative measures, Social networks
49	Année jubilaire 1927-1967 Office International du Coin de Terre et des Jardins Ouvriers	Jubilee Year 1927-1967 Office International du Coin de Terre et des Jardins Ouvriers.	City Archive of Esch-sur-Alzette	FR,DE	40th anniversary of the gardens for workers. It also discusses social housing, leasure, evolution of the organisation, municipal parks and urbanism.	YES	Magazine	Newspapers and Magazines	Period	1927	1967	None	Luxembourg	Ecosystems, Social	Built environment/urban landuses, Green spaces, Walkability, Inequality, Social networks
50	Eist Waasser	Our Water.	City Archive of Esch-sur-Alzette	DE	Water systems, fround water, drinking water, chemicals in the water, network, industry and water, organisms in water, laws.	YES	Book	Books and Research Articles	Period	1972	1988	None	Luxembourg	Ecosystems, Physical/Chemical	Blue Spaces, Point/ line sources; e.g. factories, Drinking water contamination, Groundwater contamination, Surface water contamination
51	Ecole en forêt 1904-1930	School forest 1904-1930.	City Archive of Esch-sur-Alzette	DE	Documents about the school of the forest.	YES	Document	Official Reports	Period	1904	1930	None	Luxembourg	Ecosystems, Lifestyle, Physical/Chemical	Green Spaces, Physical Activity, Odor and noise, Outdoor and indoor air pollution
52	Le Recensement de l'agriculture au 15 mai 1988	Census of Agriculture as of May 15, 1988.	City Archive of Esch-sur-Alzette	FR	Overview statistics of the agriculture in Luxembourg: exploitations, machines, surface.	YES	Text,Dataset	Official Reports	Period	1950	1988	None	Luxembourg	Ecosystems, Physical/Chemical	Green Spaces, Agricultural activities, livestock
53	Recensement de la population au 1er mars 1991 - Canton d'Esch	Population census as of March 1, 1991 - Canton of Esch.	City Archive of Esch-sur-Alzette	FR	Statistics about population by gender, commune of residence, age, family situation, work place, nationality, professional status, living conditions (e.g. type of building, surface...).	NO	Text,Dataset,Visualization	Official Reports	Year	1991	1991	Esch-sur-Alzette	Luxembourg	Ecosystems, Lifestyle, Social	Built environment/ urban landuses, Population density, Job, Social networks, Cultural norms
54	Syndicat des Eaux du Sud Koerich	South Koerich Water Syndicate.	City Archive of Esch-sur-Alzette	FR	Main reservoirs, network of water pipes, history, analysis of produced water, technical installations, evolution of the nature of water (spring, underground, surface).	YES	Text,Dataset,Visualization,Image	Official Reports	Period	1960	1976	Koerich	Luxembourg	Ecosystems, Physical/Chemical	Blue Spaces, Drinking water contamination, Groundwater contamination, Surface water contamination
55	Bulletin Trimestriel Novembre 1933. Tableau économique	Quarterly Bulletin November 1933. Economic table.	Missing	FR,DE	Population movements per Canton, currencies average rate Luxembourg Stock Exchange, savings bank.	NO	Text,Dataset	Official Reports	Period	1928	1932	None	Luxembourg	Ecosystems, Social	Population density, Household income, Social networks
56	Les Archives de la Division d'ARBED Esch-Schifflange 1870 - 1940 Inventaire	The Archives of the ARBED Division Esch-Schifflange 1870 - 1940 Inventory.	The Virtual Knowledge Centre for Europe (CVCE)	FR	Inventory of the archived documents for the Esch-Schifflange factory: accounting, real state, management, health insurance for workers, labor council, sales, workers register.	YES	Document	Official Reports	Missing	Missing	Missing	Schifflange, Esch-sur-Alzette	Luxembourg	Ecosystems, Lifestyle, Social	Health infrastructure, Job, Household income, Social networks
57	Population Statistics	Population Statistics .	STATEC Website	FR	Population per canton and commune, density of population, gender, nationality, wages.	YES	Dataset	Official Reports	Period	1821	2022	None	Luxembourg	Ecosystems, Lifestyle, Social	Population density, Job, Household income, Social networks, Cultural norms
58	Air Quality Statistics	Air Quality Statistics.	STATEC Website	FR	Concentrations in ug/m3 in the air from 1990.	YES	Dataset	Official Reports	Period	1990	2020	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution
59	Forest Areas	Forest Areas.	STATEC Website	FR	Forest areas in ha by ecological region.	YES	Dataset	Official Reports	Period	2000	2010	None	Luxembourg	Ecosystems	Green Spaces
60	Les poussieres industrielles en pathologie humaine et expérimentale (silicose et silico-anthracose sidérose et sidéro-silicose)	Industrial dusts in human and experimental pathology (silicosis and silico-anthracosis siderosis and siderosilicosis).	National Library of Luxembourg	FR	Includes tables containing birth date, 'silicogenic' station (ARBED mine, iron/steel foundry) + working years, other silicogenic stations (outside ARBED)+working years, silicose stadium (I-III), continued to work/ changed station/ retired/age of death.	NO	Text,Dataset	Official Reports	Period	1920	1960	None	Luxembourg	Lifestyle, Physical/Chemical, Phenotype	Job, Outdoor and indoor air pollution, Occupational exposures, Disease
61	Etude sur la Pollution de l'Alzette	Study on the Pollution of the Alzette.	Missing	FR	Hydrographic data of the Alzette between Esch and Ettelbruck, pollution sources, sample analysis from 1953.	NO	Text,Dataset,Visualization,Map	Books and Research Articles	Year	1953	1953	River between Ettelbruck and Esch-sur-Alzette	Luxembourg	Ecosystems, Physical/Chemical	Blue Spaces, Surface water contamination
62	Entreprises et environnement en France de 1960 à 1990. Les chemins d'une prise de conscience	Companies and the environment in France from 1960 to 1990. The paths of awareness.	Missing	FR	Environmental protection evolution between 1960 and 1990 in France.	NO	Book	Books and Research Articles	Period	1960	1990	None	France	Physical/Chemical	Point/ line sources; e.g. factories; Outdoor and indoor air pollution
63	Zeitschritt für Unternehmensgeschichte	Time step for company history.	Missing	DE	History of metal and chemical industries between 1850 and 1933.	NO	Book	Books and Research Articles	Period	1850	1933	None	Germany	Physical/Chemical	Point/ line sources; e.g. factories
64	La pollution atmospherique dans la siderurgie	Air pollution in the steel industry.	Missing	FR	Air pollution in steelworks.	NO	Book	Books and Research Articles	Missing	Missing	Missing	None	Missing	Physical/Chemical	Outdoor and indoor air pollution
65	ARBED Un Demi-siècle d'histoire industrielle 1911-1964	ARBED Half a century of industrial history 1911-1964.	Missing	FR	Document for managing newcomers of the ARBED presenting the company and the different mines and factories (including Esch Belval), it includes photos, and newspaper extracts, extraction rights. production details, mineral processing techniques…	YES	Book	Books and Research Articles	Period	1911	1964	None	Luxembourg	Lifestyle, Physical/Chemical	Job, Point/ line sources; e.g. factories
66	Sitzung von 9 Mai 1923 Deliberationsregistern Esch an der Alzette	Meeting of 9 May 1923 Deliberation registers Esch sur Alzette	Luxembourg National Archive	FR,DE	Request approval: Terms of contract from ARBED, Forest school project.	NO	Document	Official Reports	Date	09/05/1923	09/05/1923	Esch-sur-Alzette	Luxembourg	Ecosystems, Lifestyle, Social, Physical/Chemical	Green Spaces, Physical Activity, Legislative measures, Odor and noise, Outdoor and indoor air pollution
67	Contrat d'association sand but lucratif Administration communale d'Esch	Contract of association sand profit aim Administration communale d'Esch.	Luxembourg National Archive	FR	Contract of creation of the school forest between the city of Esch, ARBED, and the association Terres Rouges.	NO	Document	Official Reports	Missing	Missing	Missing	None	Luxembourg	Ecosystems, Lifestyle, Social, Physical/Chemical	Green Spaces, Physical Activity, Legislative measures, Odor and noise, Outdoor and indoor air pollution
68	Bericht über Tätigkeit die Waldschularzt	Report on activity the forest school doctor.	Luxembourg National Archive	DE	Report of the doctor in the forest school about height/ weight/ chest size gain, diseases, hygenic conditions in the schools.	NO	Document	Official Reports	Missing	Missing	Missing	Esch-sur-Alzette	Luxembourg	Ecosystems, Lifestyle, Physical/Chemical, Phenotype	Health infrastructure, Green Spaces, Physical Activity, Odor and noise, Outdoor and indoor air pollution, Disease

69	Architectural designs Forest School	Architectural designs Forest School.	Luxembourg National Archive	FR	Architectural designs Waldschule.	NO	Map	Maps	Missing	Missing	Missing	Esch-sur-Alzette	Luxembourg	Ecosystems, Lifestyle, Physical/Chemical	Built environment/ urban landuses, Health infrastructure, Green Spaces, Physical Activity, Odor and noise, Outdoor and indoor air pollution
70	Ville d'Esch Ecole en plein air	City of Esch open air school.	Luxembourg National Archive	FR	Map of Esch including the buildings of the forest school.	NO	Map	Maps	Missing	Missing	Missing	Esch-sur-Alzette	Luxembourg	Ecosystems, Lifestyle, Physical/Chemical	Built environment/ urban landuses, Health infrastructure, Green Spaces, Physical Activity, Odor and noise, Outdoor and indoor air pollution
71	Registered contagious diseases per year	Registered contagious diseases per year.	STATEC Website	FR	Registered contagious diseases per year.	YES	Dataset	Official Reports	Period	1998	2019	None	Luxembourg	Phenotype	Disease
72	Air quality measured in particulate matter (PM) 2007-2020, Esch, Lux, Differdange, Schifflange	Air quality measured in particulate matter (PM) 2007-2020, Esch, Lux, Differdange, Schifflange.	Environmental Administration	FR	Air quality measured in particulate matter (PM) 2007-2020, Esch, Lux, Differdange, Schifflange.	YES	Dataset	Measured or simulated Data	Period	2007	2020	Esch-sur-Alzette, Differdange, Luxembourg, Schifflange	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution
73	Surveillance et évaluation de la qualité de l'air- Publications périodiques	Air quality monitoring and assessment - periodic publications.	STATEC Website	FR	Dust fallout/ substances contained in plants/PM/NO2/substances in PM.	YES	Dataset	Official Reports	Period	2007	2021	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution, Soil contamination
74	Air quality 1990-2020	Air quality 1990-2020.	STATEC Website	FR	Specification (chemicals, etc.) and year.	YES	Dataset	Official Reports	Period	1990	2020	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution
75	Inventory of air pollutants (in 1000 tons) 1990 - 2018	Inventory of air pollutants (in 1000 tons) 1990 - 2018.	STATEC Website	FR	Inventory of air pollutants (in 1000 tons).	YES	Dataset	Official Reports	Period	1990	2018	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution
76	Inventory of greenhouse gaz emissions (GHG) (in 1000 tons) by main CRF sectors 1990 - 2018	Inventory of greenhouse gaz emissions (GHG) (in 1000 tons) by main CRF sectors 1990 - 2018.	STATEC Website	FR	Specification (CRF sector, etc.), measured chemicals (CH4, NO2, CO2) and year.	YES	Dataset	Official Reports	Period	1990	2018	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution
77	Status of water bodies (in %) 2009 and 2015	Status of water bodies (in %) 2009 and 2015.	STATEC Website	FR	Specification (good, bad…) and year.	YES	Dataset	Official Reports	Period	2009	2015	None	Luxembourg	Ecosystems, Physical/Chemical	Blue Spaces, Surface water contamination
78	Water data: dissolved+colloidal fractions (<1.0μm, <0.45μm and <0.22 μm),	Water data: dissolved and colloidal fractions (<1.0μm, <0.45μm and <0.22 μm).	Luxembourg Institue of Science and Technology (LIST)	EN	171 measurements of water (+colloids), looking at several parameters (pH, metals, etc.).	YES	Dataset	Measured or simulated Data	Period	2005	2013	None	Luxembourg	Ecosystems, Physical/Chemical	Blue Spaces, Surface water contamination
79	Water data: suspended sediments	Water data: suspended sediments.	Luxembourg Institue of Science and Technology (LIST)	EN	184 measurements of sediments/effluents, looking at several parameters (metal concentrations).	YES	Dataset	Measured or simulated Data	Period	2005	2013	None	Luxembourg	Ecosystems, Physical/Chemical	Blue Spaces, Surface water contamination
80	Problematic industrial waste (in tons) 1994 - 2019	Problematic industrial waste (in tons) 1994 - 2019.	STATEC Website	FR	Specification (waste type) and year.	YES	Dataset	Official Reports	Period	1994	2019	None	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories, Soil contamination, Groundwater contamination
81	Phytosanitary status (in %) 1986 - 2020	Phytosanitary status (in %) 1986 - 2020.	STATEC Website	FR	Specification (damage status of trees) and year.	YES	Dataset	Official Reports	Period	1986	2020	None	Luxembourg	Ecosystems	Green Spaces
82	Iron and steel-production (in 1 000 tonnes) 1990 - 2020	Iron and steel-production (in 1 000 tonnes) 1990 - 2020.	STATEC Website	FR	Production by product and year.	YES	Dataset	Official Reports	Period	1990	2020	None	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories
83	Production of construction materials 1938 - 2014	Production of construction materials 1938 - 2014.	STATEC Website	FR	Production of construction materials 1938 - 2014.	YES	Dataset	Official Reports	Period	1938	2014	None	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories
84	Production, exports, imports and consumption of beer (in hl) 1966 - 2019	Production, exports, imports and consumption of beer (in hl) 1966 - 2019.	STATEC Website	FR	Production, exports, imports and consumption of beer (in hl) 1966 - 2019.	YES	Dataset	Official Reports	Period	1996	2019	None	Luxembourg	Ecosystems, Lifestyle, Physical/Chemical	Food/alcohol outlets, Alcohol use, Point/ line sources; e.g. factories
85	Biological samples lichens	Biological samples lichens.	Luxembourg Institue of Science and Technology (LIST)	EN	16 lichen measurements (metal conc.), 4 test tree/wood/bark measurements beech (metal conc.).	YES	Dataset	Measured or simulated Data	Period	2005	2010	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution
86	Dust waste incinerator	Dust waste incinerator.	Luxembourg Institue of Science and Technology (LIST)	EN	1 dust measurement (metal conc.).	YES	Dataset	Measured or simulated Data	Year	2013	2013	None	Luxembourg	Physical/Chemical	Outdoor and indoor air pollution
87	ARBED Timeline	ARBED historical timeline.	STATEC Archive	FR	Timeline of the history of ARBED until 1971.	NO	Book	Books and Research Articles	Period	1847	1971	None	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories
88	Annuaire statistique rétrospectif 1973 Luxembourg	Retrospective statistical yearbook 1973 Luxembourg.	STATEC Archive	FR	Employees in surface and underground mining industry by type of occupation, nationality, wages, industrial production from 1939.	NO	Dataset	Official Reports	Period	1939	1973	None	Luxembourg	Lifestyle, Social, Physical/Chemical	Job, Household income, Social networks, Cultural norms, Occupational exposure
89	L'assurance-maladie 1933 par le Comite Central des Causes de maladie	Medicare 1933 by the central committee on causes of illness.	STATEC Archive	FR	General sickness cases, industrial deaths per year apd. 1913, sickness benefits of the factories and specific cases (coverage in which cases, apd of which day and with which salary).	NO	Dataset	Official Reports	Period	1913	1933	None	Luxembourg	Lifestyle, Social, Physical/Chemical, Phenotype	Job, Household income, Legislative measures, Occupational exposure, Disease
90	Rapport Annuel de l'Inspection du Travail et des Mines Année 1949	Annual report of the labour and mining inspection year 1949.	STATEC Archive	FR	Number of accidents per mine, by causes, work incapacity, type of injuries, wages, concessions, labor disputes, paid vacations, legal paid holidays.	NO	Dataset	Official Reports	Missing	Missing	Missing	None	Luxembourg	Lifestyle, Social, Physical/Chemical, Phenotype	Job, Household income, Legislative measures, Occupational exposure, Disease
91	Entwicklungsgeschichte der Luxemburgischen Eisenindustrie im XIX Jahrhundert	History of the development of the Luxembourg iron industry in the XIX century.	STATEC Archive	DE	History of the steel industry in the XIX century. Maps lorraine-luxembourg mining basin, analysis table of the Luxembourg alluvial ores according to Reuter and van Kerckhoff. Konzessions on a map and date of the concession.	NO	Book	Books and Research Articles	Period	1800	1900	Lorraine	Luxembourg,France	Physical/Chemical	Point/ line sources; e.g. factories
92	Rapport annuel de l'Office National du Travail 1960	Annual report of the National Labor Office 1960.	STATEC Archive	FR	Unemployment and training figures 1960.	NO	Dataset	Official Reports	Year	1960	1960	None	Luxembourg	Lifestyle, Social	Job, Household income, Inequality
93	La démographie du Luxembourg passé, présent et avenir. Rapport au president du gouvernement (1ère partie). Avril 1978	The demography of Luxembourg past, present and future. Report to the President of the Government (Part 1). April 1978	STATEC Archive	FR	Population movements, projections of different scenarios.	NO	Dataset	Official Reports	Missing	Missing	Missing	None	Luxembourg	Ecosystems	Built environment/ urban landuses, Population density
94	La croissance de l'économie luxembourgeoise	The growth of the Luxembourg economy.	STATEC Archive	FR	Steel production, per capita income, female labor, evolution of public investments, private savings.	NO	Book	Books and Research Articles	Period	1851	Missing	None	Luxembourg	Lifestyle, Social	Job, Household income, Inequality
95	Annuaire météorologique et hydrologique	Meteorological and hydrological yearbook.	STATEC Archive	FR	Annual meteorological and hydrological report.	NO	Dataset	Official Reports	Period	1983	Missing	None	Luxembourg	Ecosystems	Blue Spaces, Climate
96	Rapport & Bilan ARBED pour l'exercise 1953	ARBED Report & Balance Sheet for the year 1953.	STATEC Archive	FR	ARBED report.	NO	Document	Non-official document	Period	1953	Missing	None	Luxembourg	Social, Physical/Chemical	Household income, Point/ line sources; e.g. factories
97	Memorandum sur l'agriculture et les problèmes de l'environnement	Memorandum on agriculture and environmental problems.	Missing	FR	Use of fertilizers in the EU and USA (N, P205, K20), environmental damage index, annual production of pollutants.	NO	Dataset	Official Reports	Period	1956	1969	Europe, US	Europe,United States	Physical/Chemical	Agricultural activities, livestock, Soil contamination
98	Indicateurs Rapides Series	Quick Indicators Series.	City Archive of Esch-sur-Alzette	FR	Car accidents (Serie F), motor vehicles (D), Industrial Production (B).	NO	Dataset	Official Reports	Period	1978	1979	None	Luxembourg	Physical/Chemical	Point/ line sources; e.g. factories
99	Maps Esch and Belval	Maps of Esch and Belval.	Archive of the Municipality of Sanem	LU	Several maps of Esch and Belval site.	NO	Map	Maps	Period	1982	1929	Sanem	Luxembourg	Ecosystems	Built environment/ urban landuses, Walkability
100	Les Eaux minérales luxembourgeois. Theories physico-chimiques modernes	The Luxembourg mineral waters. Modern physico-chemical theories.	Archive of the Municipality of Sanem	FR	Modern physical-chemical theories presented to the Science Section of the Grand-Ducal Institute.	NO	Document	Official Reports	Missing	Missing	Missing	None	Luxembourg	Physical/Chemical	Drinking water contamination, Groundwater contamination, Surface water contamination
101	Application thérapeutiques	Therapeutic applications.	Archive of the Municipality of Sanem	FR	Therapeutical applications of mineral water in Luxembourg.	NO	Document	Non-official document	Missing	Missing	Missing	None	Luxembourg	Phenotype	Disease
102	Aspects économiques et financiers du développement de l'industrie sidérurgiques luxembourgeoise	Economic and financial aspects of the development of the Luxembourg steel industry.	Missing	FR	Forest surface, agriculture, steel production techniques.	NO	Text,Dataset	Official Reports	Period	1830	1889	None	Luxembourg	Ecosystems, Physical/Chemical	Green Spaces, Point/ line sources; e.g. factories, Agricultural activities, livestock
103	La Métallurgie du Luxembourg	The metallurgy of Luxembourg.	Missing	FR	Production amounts per factory, technical description of the furnaces and techniques, number of employees, prices, maps of the factories, production of electrical energy, production of fertilizers.	NO	Text	Books and Research Articles	Missing	Missing	Missing	None	Luxembourg	Lifestyle, Ecosystems, Physical/Chemical	Job, Point/ line sources; e.g. factories, Agricultural activities, livestock
104	Structure economique des industries fondamentales du grand-duche de Luxembourg	Economic structure of the basic industries of the Grand Duchy of Luxembourg.	Missing	FR	Production amounts per factory, technical description of the furnaces and techniques, number of employees, prices, maps of the factories, production of electrical energy, production of fertilizers.	NO	Text	Books and Research Articles	Missing	Missing	Missing	None	Luxembourg	Lifestyle, Ecosystems, Physical/Chemical	Job, Point/ line sources; e.g. factories, Agricultural activities, livestock

105	La siderurgie luxembourgeoise	The Luxembourg steel industry.	STATEC Archive	FR	Production amounts per factory, technical description of the furnaces and techniques, number of employees, prices, maps of the factories, production of electrical energy, production of fertilizers.	NO	Text	Books and Research Articles	Missing	Missing	Missing	None	Luxembourg	Lifestyle, Ecosystems, Physical/Chemical	Job, Point/ line sources; e.g. factories, Agricultural activities, livestock
106	La région industrielle de la haute-alzette	The industrial region of the Upper Alzette.	STATEC Archive	FR	Production amounts per factory, technical description of the furnaces and techniques, number of employees, prices, maps of the factories, production of electrical energy, production of fertilizers.	NO	Text	Books and Research Articles	Missing	Missing	Missing	Upper Alzette	Luxembourg	Lifestyle, Ecosystems, Physical/Chemical	Job, Point/ line sources; e.g. factories, Agricultural activities, livestock
107	Entretien et administration des cours de l'eau. Loi du mai 1929.	Maintenance and administration of water courses. Law of May 1929.	Luxembourg National Archive	FR	Water purification by-law.	NO	Letter,Text	Official Reports	Missing	Missing	Missing	None	Luxembourg	Ecosystems, Physical/Chemical	Blue Spaces, Surface water contamination
108	Map Esch-sur-Alzette 1856	Map Esch-sur-Alzette 1856.	Luxembourg National Archive	FR	Map of the city of Esch. In the folder there are also other maps of Kayl and Dudelange.	NO	Map	Maps	Year	1856	1856	Esch-sur-Alzette	Luxembourg	Ecosystems	Built environment/ urban landuses, Walkability
109	Apercu general agriculture luxembourgeois	General overview of Luxembourg agriculture.	Luxembourg National Archive	FR	Statistics about the situation of the agriculture per Canton: cereals produces, animals, wood, roots, dairy products, wine.	NO	Text,Dataset	Official Reports	Period	1890	1933	None	Luxembourg	Physical/Chemical	Agricultural activities, livestock
110	L'Immigration Etrangere a Differdange au debut du XXe siecle	Foreign Immigration in Differdange at the beginning of the 20th century.	STATEC Archive	FR	Population analysis.	NO	Text	Thesis	Period	1898	1914	None	Luxembourg	Ecosystems	Population density
111	Etat de la population d'apres les résultats du recensement du 1er décembre 1900	State of the population according to the results of the census of December 1, 1900.	STATEC Archive	FR	Demographic statistics.	NO	Text	Official Reports	Period	1900	1903	None	Luxembourg	Ecosystems	Population density
112	Evolution et croissance de la ville d'Esch-sur-Alzette Vol. XX	Evolution and growth of the city of Esch-sur-Alzette Vol. XX.	STATEC Archive	FR	Demographic statistics about Esch-sur-Alzette.	NO	Text	Official Reports	Missing	Missing	Missing	Esch-sur-Alzette	Luxembourg	Ecosystems	Population density
113	La pollution des eaux de l'Alzette	Water pollution in the Alzette.	eluxemburgensia.lu	FR	Water pollution of the Alzette caused by the industry.	YES	Newspaper	Newspapers and Magazines	Year	1954	1954	None	Luxembourg	Ecosystems, Physical/Chemical	Blue Spaces, Point/ line sources; e.g. factories , Surface water contamination
114	Le billet Lorraine	The Lorraine ticket.	eluxemburgensia.lu	FR	Complaints of the fishermen about the pollution in the Luxembourgish border caused by the industry.	YES	Newspaper	Newspapers and Magazines	Year	1956	1956	Thionville	France	Ecosystems, Physical/Chemical	Blue Spaces, Point/ line sources; e.g. factories , Surface water contamination
115	L'aspect moderne de l'épuration des eaux usées urbaines et industrielles	The modern aspect of urban and industrial wastewater treatment.	eluxemburgensia.lu	FR	Conference and the engineers and industrial practitioners forum about pollution of water.	YES	Newspaper	Newspapers and Magazines	Year	1954	1954	None	Missing	Ecosystems, Physical/Chemical	Blue Spaces, Point/ line sources; e.g. factories , Surface water contamination
116	Belvaux, le movement de la population	Belvaux, the movement of the population.	eluxemburgensia.lu	FR	Number of inhabitants in 1938 in Belvaux.	YES	Newspaper	Newspapers and Magazines	Period	1937	1938	Belvaux	Luxembourg	Ecosystems	Population density
117	WIPO patent data for 112M PubChem compounds	WIPO patent data for 112M PubChem compounds.	PubChem Website	EN	WIPO patents for every chemical registered in PubChem.	YES	Dataset	Official Reports	Period	1782	2022	None	Missing	Physical/Chemical	Point/ line sources; e.g. factories
118	Medical Symposium: Researches on chronic respiratory diseases	Medical Symposium: Researches on chronic respiratory diseases.	Missing	EN	Industrial Health and Medicine series No 18.	YES	Document	Official Reports	Period	Missing	1976	None	Luxembourg	Phenotype	Disease
119	Geographical variation of overweight, obesity and related risk factors: Findings from the European Health Examination Survey in Luxembourg, 2013-2015	Geographical variation of overweight, obesity and related risk factors: Findings from the European Health Examination Survey in Luxembourg, 2013-2015.	Luxembourg Institute of Health	EN	Paper on geographical variation of overweight, obesity and related risk factors in Luxembourg.	YES	Text	Books and Research Articles	Period	2013	2015	None	Luxembourg	Phenotype	Disease
120	Depression burden in luxembourg: Individual risk factors, geographic variations and the role of migration, 2013–2015 European Health Examination Survey	Depression burden in luxembourg: Individual risk factors, geographic variations and the role of migration, 2013–2015 European Health Examination Survey.	Luxembourg Institute of Health/Journal of Affective Disorders	EN	Paper on depression burden in Luxembourg.	YES	Text	Books and Research Articles	Period	2013	2015	None	Luxembourg	Phenotype	Disease
121	Hypertension burden in Luxembourg: Individual risk factors and geographic variations, 2013 to 2015 European Health Examination Survey	Hypertension burden in Luxembourg: Individual risk factors and geographic variations, 2013 to 2015 European Health Examination Survey.	Luxembourg Institute of Health/Medicine journal	EN	Paper on hypertension burden in Luxembourg.	YES	Text	Books and Research Articles	Period	2013	2015	None	Luxembourg	Phenotype	Disease



Appendix II: Entity Relationship Diagram