

A Synoptic Visualization Framework for the Multi-Perspective Study of Biography and Prosopography Data

Florian Windhager^{1*}, Paolo Federico², Saminu Salisu¹, Matthias Schlögl³, and Eva Mayr¹

¹ Danube University Krems ² Vienna University of Technology ³ Austrian Academy of Sciences

ABSTRACT

The investigation of biography data as consistently time-oriented information connecting multiple data dimensions, can be supported by multiple visualization perspectives. Biographical and prosopographical database projects contain temporally structured datasets connecting events, places, people, institutions with a variety of relations between them. We discuss challenges emerging from scholarly reasoning with these complex aggregated data, and present a visualization concept as a basis for future investigations. Specific attention is dedicated to the discussion of possible synergies and combinations (coherence techniques) to combine information and insights from multiple views and perspectives into more coherent visual analytics environments, supporting the creation of integrated and shared mental models of (time- or history oriented) data from our collective past.

Keywords: Biography data, prosopography data, information visualization, visual analytics, information integration, mental model.

1 INTRODUCTION

Across different historically oriented academic domains, the study of biographies is a common and essential investigative perspective since centuries [1]. The results though – i.e. hundreds of thousands biographical texts and their accumulations into biographical and prosopographical libraries and lexica – have recently been transformed by digital humanities initiatives: Methods of digitization and natural language processing enable the extraction of structured entities and of their relations from these text collections, which largely remained invisible raw data to digital research methods before [2], [3].

After extraction and disambiguation procedures, the resulting databases comprise named entities like *actors*, *places*, *works*, and *institutions*, all in *relations* to one another, and all interwoven by lifelines as sequences of *events* (actions and activities) [4]. The data in these biographical collections thus are commonly

- i) *multidimensional* (i.e. have spatial dimension, relational dimension, categorial dimensions), and data in all these dimensions are commonly
- ii) *time-oriented*, i.e. entities are linked to events or connected to (multiple) timestamps (see Fig. 1).

Aggregated data collections bring along new options of searching and querying, combining and comparing, but also novel *challenges*, among which the cognitive efforts, required for reasoning and sensemaking with these large amounts of multidimensional and abstract data, rank high (cp. [5]).

Basic methods for the *visualization* of data (like maps, networks, treemaps, and timelines) already support historians and humanities researchers with their well-known strengths: They

make abstract data and data structures visible for the human eye, and utilize the unique abilities of human perception to visually identify patterns, clusters, or trends, and to visually reason on them [6]. Also in face of complex humanities data, people thus should be enabled to “use visual analytics techniques, to synthesize information and derive insight from massive, dynamic, ambiguous, and often conflicting data” [7] to detect expected constellations, but also to discover novel and unexpected connections and to use these visual means to communicate their insights and assessments effectively.

Yet, especially the *synthesis* of information on a higher level (e.g. in face of a visualization system utilizing multiple views) is largely left to the unaided cognitive capabilities of analysts, resulting in high cognitive load, and likely in reductionist, fragmented, distorted or even faulty mental models of the data [8].

In accordance with this challenge, the leading question of this paper is: How can we support the visual reasoning and sensemaking of history and humanities scholars with biographical data by an integrated information visualization environment? How can we offer multiple standard visualization perspectives, but also decidedly support the cognitive integration of multidimensional insights and information? After discussing related work in the humanities and visualization context in section 2, we will propose a synoptic visualization framework illustrated by a case study in section 3, and discuss related challenges in section 4.

2 RELATED WORK

Looking at the prior art, we find various developments to support the visual analysis of selected dimensions of biographical or prosopographical data. Due to their prominence and availability, map-based visualization have already been adapted for the visualization of biography data [6], and methods for the geotemporal visualization of individual movement patterns are

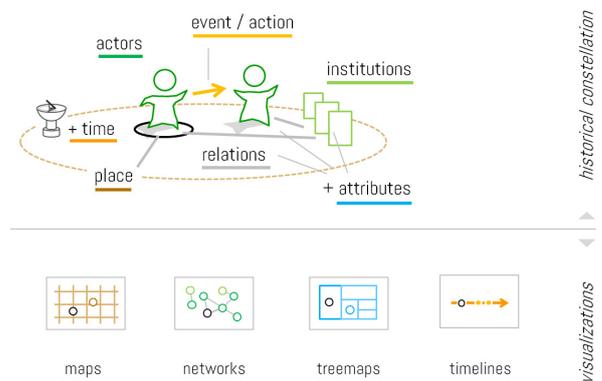


Fig. 1. Investigation into the life and work of historic individuals (top) can analytically distinguish various entities (like actors, places, institutions, events), and their interrelations, including relation to time. Visualization methods (bottom) map these data by different layouts like maps, networks, treemaps, or timelines.

* 1 firstname.lastname@donau-uni.ac.at

2 lastname@ifs.tuwien.ac.at

3 firstname.lastname@oeaw.ac.at

under constant development [9] [10] [11] [12]. From the relational perspective, network frameworks [13] and mixed method approaches [14] for biographical network visualization have been proposed. Attributes of actors or institutions, like professions or fields of cultural production have been visualized by treemaps [15] [16], while other approaches propose multi-method visualizations [17]. As for the visualization of time, various approaches to the linear mapping of time as *timelines* have been discussed [18] [19] [20]. For other methods to visually encode time (e.g. animation, layer juxtaposition, layer superimposition, and space-time cube representations) see section 3.2 and figure 2.

Given this growing amount of visualization approaches to complex phenomena of biographies and prosopographical life pattern analysis, the challenge of integrating different data dimensions have not been addressed systematically so far. We consider a more systematic support of cognitive synthesis, of insight and information integration to be a relevant research and development gap, which we want to address further down.

3 MULTIPLE VIEWS FOR MULTIDIMENSIONAL DATA ANALYSIS

For the interpretation of multidimensional biographies, data complexity means that there is only so much meaning, that a single visual-analytical perspective can reveal. Different visualization methods have their strengths – for certain data and tasks – but also their limitations or even drawbacks with regard to others. Advanced visual interfaces forego these limitations by utilizing multiple views or perspectives, to cover multiple data dimensions and aspects either in a parallel (often as coordinated multiple views) or to be chosen in a serial manner. In the following - and due to the relevance that time is playing in biography data - we distinguish *temporal* (diachronic, longitudinal, dynamic, time-oriented, temporal) from *cross-sectional* (synchronic, structural, spatial, non-temporal) data perspectives or views (see Fig. 2), and we argue that advanced visual-analytical interfaces to biography data are well-advised to integrate multiple perspectives from both categories.

3.1 Multiple Cross-Sectional Views

With the term cross-sectional, we refer to all visualizations which are not encoding temporal aspects of biography data, but represent data from a synchronic or aggregated perspective. Example of these views are basic visualizations of geographic maps, network diagrams or treemaps, as shown in figure 2 (outer left column). Implemented as multiple views, they can combine their analytical features, but commonly have to be complemented by at least one analytical perspective on temporal aspects of data organization.

3.2 Multiple Temporal Views

Temporal aspects of biography data could be represented by linear encodings (e.g. timelines, see Fig. 1, bottom right), or by various hybrid techniques to encode time as joint projections together with cross-sectional representations (see Fig. 2, second to fifth column). According to Kerracher et al. [21], the importance of offering multiple views on the data "in order to maximise insight, balance the strengths and weaknesses of individual views, and avoid misinterpretation" is a design principle that also applies to combinations of multiple temporal views. For example, in temporal graph visualization it is increasingly common for systems to combine a number of visual approaches to temporal aspects, so to "allow the user to select and switch between the most appropriate representations for the data and task at hand" (ibid.). Given the consistent and oftentimes complex time-orientation of biography data, we want to argue that the implementation of multiple temporal views is a design strategy of

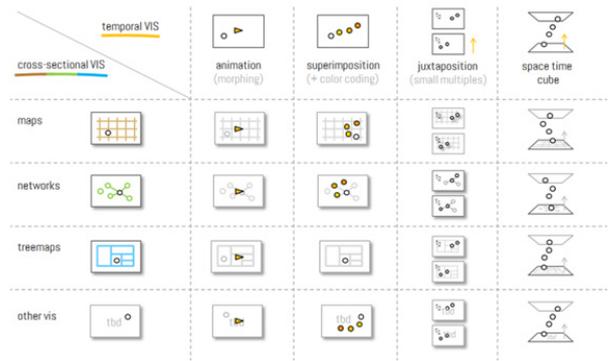


Fig. 2. The matrix of temporal and cross-sectional VIS methods for visual biography data analytics, including animation, superimposition, juxtaposition, and space-time cube, as well as geographic, relational, and distributional visualizations.

high relevance for this field of application too, and lay out a range of options in figure 2.

This matrix systematically assembles prominent cells of a design space for visual-analytical interfaces to biography or prosopography data. It points out well-established options for the visual representation of time-oriented data - as orthogonal combinations of various temporal (columns) and cross-sectional visualization methods (rows). While single methods have already been implemented in independent fashion by interfaces to biographical data collections (see sec. 2), we consider it to be a major design challenge for future approaches to find well-composed *combinations* of these multiple views. Thus designer have to assemble data and task-tailored multi-perspective interfaces, which combine analytical strengths of multiple views, and which allow to choose the most suitable ones for different tasks.

3.3 From 'Split Attention' Interfaces to a Systematic Support of Cognitive Information Integration

Yet, especially for interfaces with multiple views, a new challenge arises, which we consider a second-order problem of visual reasoning and sensemaking. When research questions encompass multiple data dimensions (like "How did the migration of an individual affect her/his social network, institutional affiliations, or means and motivations of cultural production?"), researchers have to combine information from multiple views and build up a complex mental model that integrates the different data dimensions. This task is high in cognitive effort, as attention is commonly split between the multiple views and linked data have to be identified and related, before they can be integrated into one mental model. Yet different visualization techniques (which we refer to as "coherence techniques", see [8]) can support researchers in assembling their local insights into a bigger picture. Some well-established techniques for such a support derive from the visual integration of different data dimension into one multidimensional visualization, and among those, *space-time cube representations* show a largely untapped potential to mediate across different splits and ruptures of usually separated particularistic perspectives.

In the following we introduce a framework revolving around space-time cube representations. While it firstly shows what one specific diachronic perspective can add to the visual analysis of biography data (i.e. the outer right column in figure 2), we contend that this perspective has a specific role to play for further integration due to its abilities to conceptually integrate all the other temporal perspectives [22].

4 AN INTEGRATED VISUALIZATION FRAMEWORK BASED ON SPACE-TIME CUBE REPRESENTATIONS AND OPERATIONS

While the following framework has been set up to support data and insight integration with regard to cultural collection data [23], we extend it to the specifics of biography data, as we are convinced that it provides similarly rich options for interface design in historical and biographical studies, and to also decidedly support information integration between multiple views there. We unfold the framework by starting from its geotemporal origins and move on to point out its potential to other than geographic aspects of biography data. For this demonstration we combine prototype visualizations developed across three different research projects [24] [25] [23], and a case study exploration conducted with biography data [26].

4.1 Setting up Multiple Space-Time Cubes for Biography Research

4.1.1 Geographic Space-Time Cube

The concept and technique of the space-time cube has been developed in human geography to support the visual analysis of human movement patterns and the spatial diffusion of innovation. The operating principle of this method is to orthogonally blend cross-sectional views (horizontal plane) and a temporal view (vertical axis) together, which allows to map data points (like the spatiotemporal presence of various entities) as a three-dimensional shape. Every behavior thus translates into the unique shape of a spatiotemporal trajectory and enables analysts to interpret biographic movements as visual patterns [11][26].

We illustrate this option for biography research by displaying a selected geotemporal lifeline extracted from the APIS project [4] [6] (see Fig. 3). The trajectory shows the life and travels of the Austro-Hungarian singer and actor Jozsef Szabo, including a tour to Paris, Brussels, and Italy in the middle of his career. Figure 4 shows eight further examples of biographies extracted from the APIS database, to hint towards the strengths of a comparative

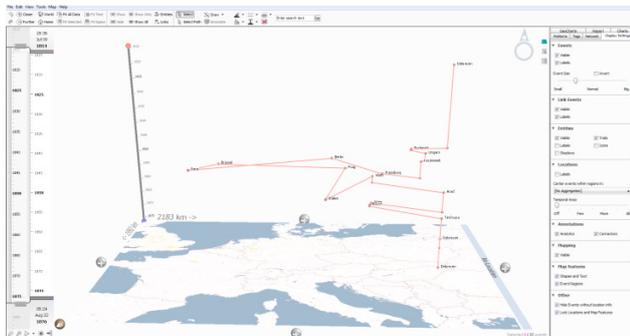


Fig. 3. Visualization of an individual biographical trajectory, from a geotemporal perspective, created with [11] as a case study [26].

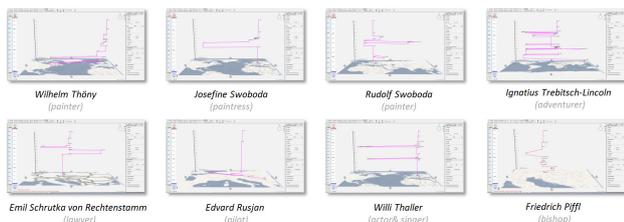


Fig. 4. Small multiples enabling the comparison of eight different biographical datasets from the APIS project [11] [25].

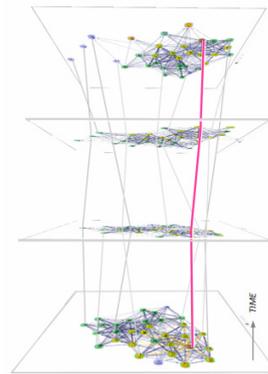


Fig. 5. Visualization of an individual movement through relational-temporal space, as demonstrated by Federico et al. [24].

approach, where small multiples directly enable comparisons of biographical space-time paths, including similarities and differences of patterns among different entities [26].

4.1.2 Relational Space-Time Cube

Beyond the geotemporal data domain, space-time cube representations can visualize different other data dimensions - in combination with time. The resulting trajectories thus can disclose the movements of historic individual through further (analytically distinguishable) spacetimes, like movement through the social-relational space of interindividual collaboration or conflict. Figure 5 conceptually illustrates this option by the highlighted movement of an actor through relational spacetime [24]. Depending on the richness of relational-temporal data, this enables to frequently follow the movements of historic actors from socio-cultural peripheries to structural cores, thus showing macro patterns but also detailed interactions of individuals, including the development of their network centrality measures.

4.1.3 Categorical Space-Time Cube

Given possible categorical spaces in which historic individuals have been active (i.e. differentiated fields of activities, professions, cultural domains, or knowledge areas), visualizations like treemaps can provide a valuable synchronic perspective [16]. Thus, by implementing treemaps into categorical-temporal space-time cubes (see Fig. 6), another analytical perspective onto historic spacetime opens up, which discloses novel behavioral movement patterns.

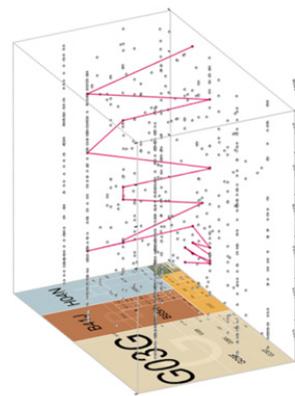


Fig. 6. Visualization of an individual movement through categorical-temporal spacetime, as demonstrated with regard to mobilities in the knowledge space of patent classification by Smuc et al. [25].

4.2 Linking Multiple Space-Time Cubes

In analogy to multiple linked views [28], we promote the connection of multiple space-time cubes to synoptic ensembles, to enable the visual exploration of biographies in multiple relevant spacetimes (Fig. 7). The specific line up of cubes – which could include various further methods – naturally depends on available data (dimensions) and intended analytical tasks. We consider such a synoptic setup to provide an effective visualization environment, which could be explored by the means of different interaction techniques (like brushing or combined navigation), but also serve as a versatile scaffold for the selection of further analytical perspectives, including well-established methods of flat visualization design (Fig 8).

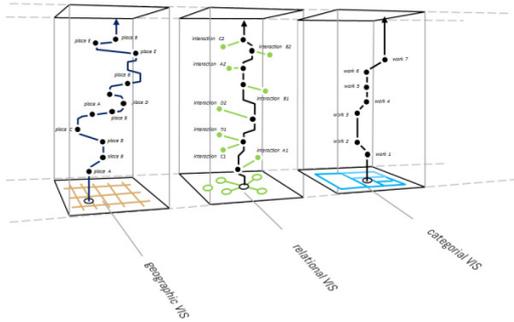


Fig. 7. A visualization environment for biography data using multiple linked space-time cubes.

4.3 Mediating Multiple Cross-Sectional and Temporal Views

As Bach et al. [22] have shown, space-time cube representations can support the cognitive translation and mediation of multiple temporal and cross-sectional views by utilizing seamless canvas transitions or seamless adaptation of a visualization’s field of view [29]. Given the outlined (linked) visualizations of the outer right column of figure 2, the other temporal visualizations (i.e. layer juxtaposition, layer superimposition, or animation – as well as all possible “space-flattened or time-flattened” standard perspectives – could be seamlessly generated out of the three-dimensional arrangements. It is our conjecture, that such seamless translations will have a strong easing effect on the preservation of mental models of complex time-oriented data, and as such for the navigation and visual reasoning – especially as an initial technique for non-expert users.

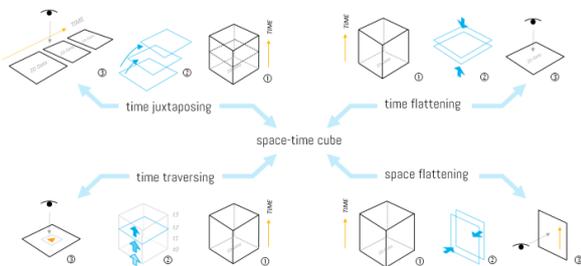


Fig. 8. Space-time cube representations as means for visual and conceptual transitions (blue), preserving and mediating mental models about temporal and cross-sectional perspectives [22] [8].

4.4 From Biographical to Prosopographical Data

Going beyond single trajectories, the outlined framework is open for more complex analyses with bigger prosopographical datasets. *Prosopography* is the domain for studying biographies as seen from a collective perspective [30]. For that purpose, groups can be visualized as *sets*, and figure 9 delineates basic flow patterns of sets, which in combination can map most complex temporal developments of historical groups or collective entities (like organizations, religions, art schools, political bodies, fashions, disciplines, or any other innovation). As a method for aggregated visualization it can complement the display of line-like, individual trajectories in geographic or relational spacetimes.

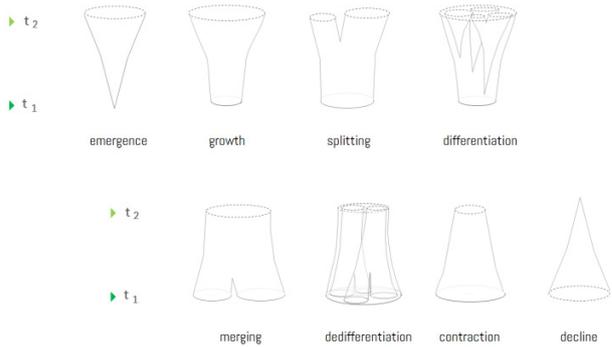


Fig. 9. Visualization of temporal developments of sets, showing the basic evolutionary patterns of groups to be utilized for the prosopographical research perspective.

5 CONCLUSION

With this paper we argued the need for a visualization framework tailored to support the visual analysis of biographical and prosopographical data. Going beyond the use of well-established multiple but separated views, its main objective is the support of insight and information integration for scholars and data analysts from a wide range of humanities and historical sciences. We consider its added value to consist in the provision of

- visual-analytical access to rich and multidimensional data, supporting visual reasoning and sensemaking,
- multiple perspectives to generate richer and non-reductionist mental models of the available data, and ultimately in
- the considerate support of scholar’s information integration, and of their required macrocognitive syntheses by the means of space-time-cube representations and operations.

Accompanying the challenges arising from the ongoing implementation and evaluation, we see future challenges with regard to the reflected handling of *data quality* and *uncertainty*, which arise from historically fragmented and often disputed data sources. Following discussions of data provenance, concepts of critical visualization design [31] could help to bring transparency into the collective interpretation process. Finally we consider this framework as a device not only to communicate results, but also to motivate and support the *collective critical editing* of biographical data as co-created trajectory graphs. By the means of shared repositories for different scholarly interpretations and annotations, variations and controversies could be made productive, and lead to a better integrated and more multi-faceted understanding of historical knowledge also with regard to shared mental models in prosopographical and biographical research.

ACKNOWLEDGMENTS

The authors wish to thank the team of the APIS project at the Austrian Center of Digital Humanities (ACDH), as well as Michael Smuc and Murat Sari (Figure 6). This research was supported in part by a grant from the Austrian Science Fund (FWF), project number P28363.

REFERENCES

- [1] B. Roberts, *Biographical research*. Open University Press Buckingham, 2002.
- [2] M. Reinert, M. Schrott, B. Ebnet, and others, "From Biographies to Data Curation-The Making of [www. deutsche-biographie. de.](http://www.deutsche-biographie.de)," in *BD*, 2015, pp. 13–19.
- [3] Reznik, I. and Shatalov, V. "Hidden revolution of human priorities: An analysis of biographical data from Wikipedia". In: *Journal of informetrics*. 2016
- [4] Bernád, Á. Z. Kaiser, M. Lejtovicz, K. Schlögl, M. "Mapping historical networks. Working with biographical data" in: *Entangled Worlds: Network analysis and complexity theory in historical and archeological research*, Wien Holzhausen, 2017 (forthcoming).
- [5] J. G. Trafton, S. S. Kirschenbaum, T. L. Tsui, R. T. Miyamoto, J. A. Ballas, and P. D. Raymond, "Turning pictures into numbers: extracting and generating information from complex visualizations," *International Journal of Human-Computer Studies*, vol. 53, no. 5, pp. 827–850, 2000.
- [6] APIS project, available at <https://apis.acdh.oeaw.ac.at>
- [7] J. Thomas and K. Cook, *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. National Visualization and Analytics Center, 2005.
- [8] G. Schreder, F. Windhager, M. Smuc, and E. Mayr, "A Mental Models Perspective on Designing Information Visualizations for Political Communication," *JeDEM - eJournal of eDemocracy and Open Government*, vol. 8, no. 3, pp. 80–99, Dec. 2016.
- [9] T. Gonçalves, A. P. Afonso, and B. Martins, "Cartographic visualization of human trajectory data: overview and analysis," *Journal of Location Based Services*, vol. 9, no. 2, pp. 138–166, Apr. 2015.
- [10] M.-P. Kwan and G. Ding, "Geo-narrative: Extending geographic information systems for narrative analysis in qualitative and mixed-method research*," *The Professional Geographer*, vol. 60, no. 4, pp. 443–465, 2008.
- [11] T. Kapler and W. Wright, "GeoTime information visualization," in *INFOVIS '04 Proceedings of the IEEE Symposium on Information Visualization*, 2004, pp. 25–32.
- [12] K. Ellegaard, M. Cooper, and others, "Complexity in daily life—a 3D-visualization showing activity patterns in their contexts," *electronic International Journal of Time Use Research*, vol. 1, no. 1, pp. 37–59, 2004.
- [13] M. Schich et al., "A network framework of cultural history," *Science*, vol. 345, no. 6196, pp. 558–562, Aug. 2014.
- [14] N. Armitage, "The Biographical Network Method," *SRO*, vol. 21, no. 2, p. 16, 2016.
- [15] A. Z. Yu, S. Ronen, K. Hu, T. Lu, and C. A. Hidalgo, "Pantheon 1.0, a manually verified dataset of globally famous biographies," *Scientific Data*, vol. 3, sdata2015.75, 2016.
- [16] Pantheon - Mapping Historical Cultural Production. Available: <http://pantheon.media.mit.edu/>
- [17] O. Gergaud, M. Laouenan, and E. Wasmer, "A Brief History of Human Time. Exploring a database of notable people," 2017
- [18] P. T. Hiller, "Visualizing the Intersection of the Personal and the Social Context-The Use of Multi-Layered Chronological Charts in Biographical Studies," *The Qualitative Report*, vol. 16, no. 4, p. 1018, 2011.
- [19] S. B. Davis, E. Bevan, and A. Kudikov, "Just in time: defining historical chronographics," in *Electronic Visualisation in Arts and Culture*, Springer, 2013, pp. 243–257.
- [20] M. Champagne, "Diagrams of the past: How timelines can aid the growth of historical knowledge," *Cognitive Semiotics*, vol. 9, no. 1, pp. 11–44, 2016.
- [21] N. Kerracher, J. Kennedy, and K. Chalmers, "The design space of temporal graph visualisation," in *Proceedings of the 18th Eurographics Conference on Visualization (EuroVis '14)*, vol. Short, N. Elmquist, M. Hlawitschka, and J. Kennedy, Eds. Swansea: Eurographics Association, 2014.
- [22] B. Bach, P. Dragicevic, D. Archambault, C. Hurter, and S. Carpendale, "A Descriptive Framework for Temporal Data Visualizations Based on Generalized Space-Time Cubes," in *Computer Graphics Forum*, 2016.
- [23] F. Windhager, E. Mayr, G. Schreder, M. Smuc, P. Federico, and S. Miksch, "Reframing Cultural Heritage Collections in a Visualization Framework of Space-Time Cubes," in *Proceedings of the 3rd Histoinformatics Workshop*. <http://ceur-ws.org>, Krakow, 2016, vol. 1632, pp. 20–24.
- [24] P. Federico, W. Aigner, S. Miksch, F. Windhager, and L. Zenk, "A visual analytics approach to dynamic social networks," in *Proceedings of the 11th International Conference on Knowledge Management and Knowledge Technologies*, New York, NY, USA, 2011, p. 47:1–47:8
- [25] M. Smuc, F. Windhager, M. Sari, P. Federico, A. Amor-Amorós, & S. Miksch, "Interweaving Pathways of Innovation. Visualizing the R&D Dynamics of Companies Provided by Patent Data", 35th Sunbelt Conference of the International Network for Social Network Analysis (INSNA 2015), Brighton, 24.06.2015.
- [26] F. Windhager, E. Mayr, G. Schreder, M. Schlögl, Á. Z. Bernád, E. Wandl-Vogt, & C. Gruber, "Zur polykubistischen Informationsvisualisierung von Biographiedaten." *Digitale Nachhaltigkeit. Digital Humanities im deutschsprachigen Raum (DHD'17)*, Bern, 15. 2. 2017.
- [27] R. Eccles, T. Kapler, R. Harper, and W. Wright, "Stories in GeoTime," *IEEE Visual Analytics Science and Technology*, 2007.
- [28] M. Scherr, "Multiple and coordinated views in information visualization," *Trends in Information Visualization*, vol. 38, 2008.
- [29] P. Federico, W. Aigner, S. Miksch, F. Windhager, and M. Smuc, "Vertigo zoom: combining relational and temporal perspectives on dynamic networks," in *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI2012)*, Capri Island, 2012, pp. 437–440.
- [30] Verboven, K. Carlier, M., and Dumolyn, J. "A short manual to the art of prosopography". In: *Prosopography Approaches and Applications. A Handbook*, Oxford, Unit for Prosopographical Research (Linacre College), 2007, pp. 35-69.
- [31] M. Dörk, P. Feng, C. Collins, and S. Carpendale, "Critical InfoVis: exploring the politics of visualization," in *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 2013, pp. 2189–2198.