Report from Dagstuhl Seminar 18482

Network Visualization in the Humanities

Edited by

Katy Börner[1], Oyvind Eide[2], Tamara Mchedlidze[3], Malte Rehbein[4], and Gerik Scheuermann[5]

1 **Indiana University – Bloomington, US,** katy@indiana.edu
2 **Universität zu Köln, DE,** oeide@uni-koeln.de
3 **KIT – Karlsruher Institut für Technologie, DE,** mched@iti.uka.de
4 **Universität Passau, DE,** malte.rehbein@uni-passau.de
5 **Universität Leipzig, DE,** scheuermann@informatik.uni-leipzig.de

Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 18482 "Network Visualization in the Humanities", which took place November 25–30, 2019. The seminar brought together 27 researchers from Network Visualization and Digital Humanities communities. During the seminar the participants shared knowledge on the existing methods of network visualization and on network visualization challenges present in the Humanities through the introductory talks, the abstracts of which are included in this report. Multiple innovative research challenges for Network Visualisation in the Humanities have been identified and according to those four working groups have been set up that discussed the topics in detail. The summary of the discussions of the working groups is given in this report.

Seminar November 25–30, 2018 – http://www.dagstuhl.de/18482
2012 ACM Subject Classification Human-centered computing → Visualization, Applied computing → Arts and humanities, Theory of computation → Design and analysis of algorithms
Keywords and phrases digital humanities, network visualization, graph drawing, human computer interaction, topic modelling, cyberinfrastructures, distant reading
Digital Object Identifier 10.4230/DagRep.8.11.139

1 Executive Summary

Katy Börner (Indiana University – Bloomington, US)
Oyvind Eide (Universität zu Köln, DE)
Tamara Mchedlidze (KIT – Karlsruher Institut für Technologie, DE)
Malte Rehbein (Universität Passau, DE)
Gerik Scheuermann (Universität Leipzig, DE)

Seminar Goals

The application of computer-based methods by scholars of the Humanities has a tradition that goes back to the mid 20th century. Labelled "Digital Humanities" some 15 years ago, it has seen a significant growth since then [1]. An important part of Digital Humanities methodology is to establish data sets [2] based on cultural artefacts such as fiction texts, paintings, musical scores and recordings, and historical sources in all media. This is done in a number of different ways and includes some sort of extraction of data from sources

Network Visualization in the Humanities, *Dagstuhl Reports*, Vol. 8, Issue 11, pp. 139–153
Editors: Katy Börner, Oyvind Eide, Tamara Mchedlidze, Malte Rehbein, and Gerik Scheuermann
Dagstuhl Reports
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

structured in different, less explicit ways than what is needed for operationalisation and computer assisted analysis and visualization. When this process works well, it supports scholars' endeavours to answer existing research questions and to generate new insights and novel research questions. A significant part of the data collected can be modeled as networks.

Existing network analysis and visualization techniques have already proven themselves immensely useful in analyzing Digital Humanities data and providing new discoveries [3]. The central goal of research on network visualization for digital humanities scholars is to develop visualization techniques and algorithms that empower scholars to use those effectively as part of their research process and for communicating study results to readers. While network science approaches are widely used in other research areas, the power of a network mindset and approach has not yet been fully exploited within the Humanities.

The seminar aimed to enhance the development of network visualization algorithms and tools centered around humanities research. In particular, its goals were as follows:

- Interdisciplinary Exchange: to discuss existing network visualization methods and algorithms in perspective of their potential application within the Humanities;
- Terminology Gap: Bridging the gap in terminology between Digital Humanities on one side and computer scientists in Network Visualization and Graph Drawing on the other side;
- Data: to discuss Humanities' data sources and their nature, research questions, use cases, and specific application profiles in perspective of their potential support by network visualisation.
- Reserch Agenda: Formulation of research agenda on "Network Visualization in the Digital Humanities". Creation of interdisciplinary teams of researchers that address specific scientific challenges of the agenda;

Seminar Program

The seminar brought together 27 researchers from Network Visualization and Digital Humanities communities. The initial two days of the weeklong event were devoted to bring together the different communities and to develop a mutual understanding. Researchers informed each other about their scholarly background through short, five-minute talks. In addition, there were eight long, 45 minutes, presentations in which digital humanities scholars discussed network and network visualization challenges and opportunities in their field of expertise. This was complemented by surveys on network visualization and successful examples of cooperation between visualization and digital humanities researchers.

During both days the participants were asked to post questions and issues they would like to discuss in the remaining three days of the seminar. After a voting, four research areas most interesting the participant were identified. All four met the guiding principles in that they describe both: highly relevant applications within the Humanities as well as innovative research challenges for Network Visualisation. They are as follows:

- Complex networks, in particular multivariate, multilayered, and multilevel networks;
- Linked networks;
- Temporal networks;
- Uncertainty, incompleteness, and ambiguity of data.

Four groups were formed to work on those four topics over the remaining three days. There were several opportunities for joint discussions and progress reports across the groups. Summaries of the group discussions can be found in Section 4.

Future Plans

During the seminar the participants decided to proceed with a publication of a manifesto, outlining a research agenda for "network visualisation in the Humanities". It was also planned to publish an edited volume on specific aspects of the overarching topic, possibly along the four major research areas identified by the seminar. The volume will be submitted as a special issue to "Historical Network Research", an Open Access Journal.

Evaluation

The feedback provided by the participants in form of a survey collected by Schloss Dagstuhl was highly positive and in most aspects above the average collected over the last 60 seminars. The participants agreed that the seminar inspired new ideas, collaborations, joint publications and brought insight from neighboring fields. There was a number of positive comments by the participants on the structure and organization of the seminar as well as several useful suggestions for the future seminars.

Acknowledgment

As an organizing committee of the seminar we would like to thank the scientific and administration staff of Schloss Dagstuhl for the excellent support they provided, both in the preparation phase and during the seminar. On behalf of all participants, we would also like to thank Dagstuhl for the high quality facilities provided, for excellent rooms for work and socializing, for the tasty meals, and of course also for the excellent wine cellar. The organizers of the seminar would also like to thank Ray Siemens and Dan Edelstein for their contributions to the initial Dagstuhl proposal. Finally, we thank Christina Gillmann for taking the responsibility for this report.

References

1 Thaller, Manfred (2017), *Geschichte der Digital Humanities.* In: Fotis Jannidis, Hubertus Kohle und Malte Rehbein (Hg.): *Digital Humanities. Eine Einführung.* Stuttgart, J.B. Metzler Verlag, S. 3–12.

2 Schöch, Christof (2013) *Big? Smart? Clean? Messy? Data in the Humanities.* In: Journal of Digital Humanities 2 (3).

3 Rehbein, Malte (forthcoming) *Historical Network Research, Digital History, and Digital Humanities.* In: Marten Düring, Florian Kerschbaumer, Linda von Keyserlingk und Martin Stark (Hg.): *The Power of Networks. Prospects of Historical Network Research*, Routledge.

2 Contents

3 Overview of Talks

3.1 Almost a Theory

Ulrik Brandes (ETH Zürich, CH)

Data, at least in the statistical sense, are the values of variables, which we conceive of as mappings of entities from some domain to values from some range. The distinct characteristic of network variables is that their domain consists of overlapping dyads. While the representational theory of measurement is concerned with the preservation of empirical structures on the domain in numerical structures on the range, a particular challenge in the Humanities is the conceptualization of the domain. Many relevant aspects of the phenomena or artifacts under scrutiny are difficult to represent in variables because abstraction is necessary to make them commensurable. With increasing levels of abstraction, comparability is widened at the expense of potentially relevant characteristics. This is reflected, for instance, in the notions of close and distant reading, and suggests interesting problems for network visualization research.

3.2 Actionable Data Visualizations

Katy Börner (Indiana University – Bloomington, US)

Main reference Katy Börner: "Atlas of Knowledge: Anyone Can Map", Cambridge, MA: The MIT Press, 2015.
Main reference Katy Börner, David E. Polley: "Visual Insights: A Practical Guide to Making Sense of Data", Cambridge, MA: The MIT Press, 2014.
Main reference Katy Börner: "Atlas of Science: Visualizing What We Know", Cambridge, MA: The MIT Press, 2010.

In the information age, the ability to read and make data visualizations is as important as the ability to read and write. This talk explains and exemplifies the power of data visualizations not only to help locate us in physical space but also to help us understand the extent and structure of our collective knowledge, to identify bursts of activity, pathways of ideas, or emerging areas of research. It introduces a theoretical visualization framework meant to empower anyone to systematically render data into insights together with tools that support temporal, geospatial, topical, and network analyses and visualizations. Materials from the Information Visualization MOOC (http://ivmooc.cns.iu.edu) and science maps from the Places & Spaces: Mapping Science exhibit (http://scimaps.org) will be used to illustrate key concepts and to inspire participants to visualize their very own data.

3.3 Modeling Data to Develop Intuitions

Nicole Coleman (Stanford University, US)

Main reference Johanna Drucker: "Humanities Approaches to Graphical Display", Digital Humanities Quarterly, Vol. 5(1), 2011.
URL http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html
Main reference Dan Edelstein, Paula Findlen, Giovanna Ceserani, Caroline Winterer, Nicole Coleman: "Historical Research in a Digital Age: Reflections from the Mapping the Republic of Letters Project", The American Historical Review, Volume 122, Issue 2, pp. 400–424, 2017.
URL https://doi.org/10.1093/ahr/122.2.400
Main reference Stefano Franchi: "The Past, Present, and Future Encounters between Computation and the Humanities", pp. 349–364, Springer, 2013.
URL http://dx.doi.org/10.1007/978-3-642-31674-6_26

A decade of problem-driven experiments with data visualization for humanities research at Humanities + Design research lab at the Center for Spatial and Textual analysis at Stanford University exposed challenges and opportunities at the intersection of humanistic inquiry and data visualization. The challenges result from the fact that humanistic inquiry tends to be a very internal thought process that does not explicitly reference external models. To externalize the research process it was necessary to learn to think procedurally, to think visually, and to capture the complexity of multiple perspectives. This talk elaborates on a number of preliminary visualization experiments, which eventually led to three opensource applications: Palladio, Breve, and Data Pen. There were mistakes and successes along the way to understanding how to produce visualization tools that reflect the needs of humanistic research. The result is a set of tools that reveal the incompleteness of data, that move seamlessly from abstraction back to rich contextual sources, that allow reflection through interaction, and that allow visual data modeling, including direct editing and enrichment of a data set. The principles underlying the design of the interaction environments we built are analogous to those within the artificial intelligence community that support the augmentation of human ability rather than autonomous systems. Computation manages the complex calculation and querying of multi-dimensional data; the visual interface renders the data concisely and intuitively; and the interaction makes the visualization an exploratory tool. This combination, with an even emphasis on missing and extant data, provides an instrument for modeling data to develop intuitions. It is a fundamental functional design of data-driven environments that can be applied to far more sophisticated underlying computational and visual techniques to augment humanistic inquiry. Finally, to legitimize this new form of scholarship means to make it publishable. The question remains, how can we share the interactive experience of these works in a way that will persist and contribute to future knowledge production?

3.4 Visualization of networks: A cartographer's view

Sara Irina Fabrikant (Universität Zürich, CH)

Maps have been used for at least 5000 years by humans to communicate about tangible networks (i.e., transportation, water, electricity, etc.) and intangible linear features of the environment (i.e., flows of goods, people, air masses, etc.). Networks can represent geographic, linear features in the environment, or metaphorical, relational information spatialized from

databases that are not necessarily spatial or geographic (i.e., relationships extracted from text archives, biological databases, financial records, etc.). Spatialized displays can be designed and explored as if they represented geographic information, considering the user, the use context, and the design characteristics of the spatialized views. Empirical studies on spatialized views depicting points, networks, and regions suggest that cartographically informed design guidelines allow information seekers to more effectively and efficiently explore relational information, and gain knowledge from large spatialized text databases.

3.5 Visualization & Digital Humanities

Stefan Jänicke (Universität Leipzig, DE)

Visualization as a research instrument for Digital Humanities scholarship has gained more and more importance in recent years, yielding mutual benefits for the involved research domains. On the one hand, visualizations aim at providing intuitive access to vast amounts of digitized data, on the other hand, the nature of digital humanities data, which is usually incomplete, inhomogeneous and uncertain, provokes new challenges for visualization research. The best practice to develop a visualization in a digital humanities project is a problem-driven user-centered design approach as it forces scholars from both fields to intensely engage with each others research interests, tasks and problems. Moreover, the strong interdisciplinary exchange increases the likelihood that the resultant visualization will serve the intended purpose, and that it prepares the ground for future research.

3.6 Graphs in Computational Literary Studies

Fotis Jannidis (Universität Würzburg, DE)

Graphs are used for very different purposes in Computational Literary Studies and the talk discusses three different examples. 1) In a complex digital edition of Goethe's Faust the statements of researchers about the chronological relations (edges) between manuscripts (nodes) allow us to find contradicting statements. 2) In Stylometry the stylistic similarity (edges) between texts (nodes) can be used to determine groups beyond authorship based on period or other factors. 3) In the analysis of plot events (nodes) are related chronologically (edges). In the last case more complex graphs, which would allow to render narratological concepts like the difference between discours/histoire, are discussed. Modeling novels like that would allow us to find similar texts based on the similarity of the graphs representing their plot structure.

3.7 Putting Networks on the Map

Stephen Kobourov (University of Arizona – Tucson, US)

Main reference Randy Burd, Kimberly Andrews Espy, Md. Iqbal Hossain, Stephen G. Kobourov, Nirav Merchant, Helen C. Purchase: "GRAM: global research activity map", in Proc. of the 2018 International Conference on Advanced Visual Interfaces, AVI 2018, Castiglione della Pescaia, Italy, May 29 – June 01, 2018, pp. 31:1–31:9, ACM, 2018.
URL http://dx.doi.org/10.1145/3206505.3206531

Relational data sets are often visualized with graphs: objects become the graph vertices and relations become the graph edges. Graph drawing algorithms aim to present such data in an effective and aesthetically appealing way. We describe map representations, which provide a way to visualize relational data with the help of conceptual maps as a data representation metaphor. While graphs often require considerable effort to comprehend, a map representation is more intuitive, as most people are familiar with maps and standard map interactions via zooming and panning. Map-based visualization allows us to explicitly group related vertices into "countries" and to present additional information with contour and heatmap overlays. We discuss the graph-to-map (GMap) algorithmic framework, including applications, such as Maps of Computer Science (MoCS) and the Global Research Activity Map (GRAM), as well as experimental results on the effectiveness of the approach.

3.8 Networks in the Humanities: Challenges & Opportunities

Scott Weingart (Carnegie Mellon University – Pittsburgh, US)

Networks have a long history in the Humanities, going back to the earliest sociometry research in the 1930s, with contributions flowing both directions. By 2010, networks became a pillar of digital humanities research. The renewed interest brought to light several broad challenges needing to be addressed over the next decade, including toolkits that account for uncertain data, that support the entire research workflow, and that prioritize readability. The Humanities ought to also contribute new theoretical and historical understandings of network visualizations and analyses.

4 Working groups

4.1 Network Taxonomies

Oyvind Eide (Universität zu Köln, DE), Francis Harvey (Leibniz Institut für Länderkunde – Leipzig, DE), Andreas Kerren (Linnaeus University – Växjö, SE), Tamara Mchedlidze (KIT – Karlsruher Institut für Technologie, DE), and Florian Windhager (Donau-Universität Krems, AT)

During the meeting, our group met and discussed network taxonomies in the Humanities, a broad and intriguing topic because of its relevance for situating network research and development in the field. Correspondingly, the scope of the discussion ranged over the breadth and depth of network-based analysis in the Digital Humanities as well as in the Humanities at large. We recognized that taxonomies of networks in the Humanities are challenging to develop because of their heterogeneity. From this starting point we made progress by considering the disciplinary roots of networks and identified differences between types of networks. These points were then taken up when we focused on different networks used in the Digital Humanities, the data they were based on, and visualization techniques. We recognized how important terminological standardization is, and suggested it can be accomplished to some degree through taxonomies. Specific topics the group took up include: scale, interaction, hierarchy of networks, types of networks, including multivariate networks, multilayer networks, multimodal networks, and types of visualisations. Uncertainty in models and data as well as the challenge of accounting for the open nature of digital humanities research entered into the discussion as key points to be considered when developing taxonomies. The support of multiple analytical perspectives on data representations and interactive visual representations is necessary in order to promote and support the potential of networks for digital humanities research. Addressing the heterogeneity of networks and research theories, methodologies and context is necessary to account for the richness and unique composition in humanities research. This reminded us that our taxonomies can be helpful for a number of reasons and can fulfill different needs. The potential for standardization of terms is considerable, but taxonomies should not end with attempts at canonical specification. This could constrain scientific concept development. Taxonomies need to be dynamic and develop over time in order to keep them helpful and useful in the development of research.

4.2 Linked Networks Perspectives for Humanities (beyond Nodes, Links and Clusters)

Gregor Betz (KIT – Karlsruher Institut für Technologie, DE), Stephen G. Kobourov (University of Arizona – Tucson, US), Martin Nöllenburg (TU Wien, AT), Gerik Scheuermann (Universität Leipzig, DE), Timothy Tangherlini (University of California at Los Angeles, US), and Christopher Warren (Carnegie Mellon University – Pittsburgh, US)

The members of the linked networks perspectives for Humanities working group were Gregor Betz, Stephen Kobourov, Martin Nöllenburg, Gerik Scheuermann, Timothy Tangherlini, and Christopher Warren. We discussed three main questions: different networks based on the same data due to varying aspects of interest, good visualizations for bipartite/k-partite networks and how to draw linked networks. We see a clear need for research on these questions based on the needs in the Humanities and lack of approaches on the computer science side.

With respect to different networks based on the same data, we discussed several cases where there are good reasons to create different networks based on the same data, even for the same question. For example, one might consider a play with several scenes. For each scene, one can construct a network of the interactions of the characters. Then, the humanities scholar is left with the task of comparing these networks with respect to differences or consensus. Also, a very important task is the critical review of the network construction that questions all definitions of nodes or edges by looking at the original text. Especially the last task is not supported by current network visualization tools.

Regarding bipartite and k-partite graphs, we detected a high need of such graphs in the Humanities. E.g. any analysis of a novel that is based on the relationship between e.g. characters and scenes, character groups, characters and locations, etc. leads to such graphs. Surprisingly, there is only very little work in the literature on drawing bipartite graphs besides the common notion of drawing each set on one side of the screen. For k-partite graphs, there is even less work. This defines a clear need on the Network Visualization and Graph Drawing side.

As third part, we discussed the drawing of linked networks. Humanities scholars create usually several networks to describe relations of interest in their study. Quite often, they also create or note links between nodes (sometimes edges) in these networks. An example is korean pop music. One might look at the performers and their grouping into bands. One can also study the relations to a (fairly small) group of music production companies. In addition, one may study the role of auxiliary contributors such as choreographers, songwriters or promoters. This creates interlinked networks. The drawing and visual analysis of such linked networks is not well researched so far and presents a lot of challenges to the Graph Drawing and Network Visualization communities.

Overall, we found in all three parts substaniell potential for new solutions from computer science and high research interest from the Humanities. We expect fruitful cooperation among the participants of the seminar and beyond on this topic and encourage everyone interested to start research. Of course, the group is willing to share further thoughts on request.

4.3 Visualizing Networks and Temporality

Melanie Conroy (University of Memphis, US), Kimmo Elo (University of Turku, FI), Gerhard Heyer (Universität Leipzig, DE), Fotis Jannidis (Universität Würzburg, DE), Malte Rehbein (Universität Passau, DE), Antonis Symvonis (National TU – Athens, GR), and Scott Weingart (Carnegie Mellon University – Pittsburgh, US)

The members of the temporality working group were Melanie Conroy, Kimmo Elo, Gerhard Heyer, Fotis Jannadis, Malte Rehbein, Antonic Symvonis, and Scott Weingart. We discussed both ways to graph temporal networks and how to visualize data in the Humanities that include networks. While we discussed problems which could be addressed using temporal graphs, we found that temporal graphs could often not be constructed from the datasets with which we were familiar. Our discussion focused on how to incorporate non-linear time sequences and ways of perceiving time that differ from chronological time into network diagrams and other visualisations. After attempting to construct network graphs for various use cases, we discovered that many problems in the Humanities do not permit the construction of a temporal network graph due to multiple perspectives on the network and variable or uncertain time sequences. While we discussed ways to reduce the number of perspectives and series, however, we rejected the idea of reducing the complexity of data models. We decided that starting with the visualizations that we needed for a number of case studies would be more valuable than attempting to reduce the complexity of humanities research questions to make them graphable as a single network. For this reason, we decided to work backwards from the types of visualisations needed for individual projects to the data model that would be necessary to produce such a visualisation.

Problems that appeared repeatedly in our discussions of the temporality of networks in the Humanities included a mismatch between methodology (data models and metrics) and available technology, different data models and collection practices in various humanities fields, and project-specific data models. We also discussed incomplete or uncertain data and shifting or incommensurable perspectives related to time. We decided that no one visualisation or set of visualisations would be adequate to deal with all of these issues.

We discussed four main use cases for networks in the Humanities:

1. Story vs. Discourse – Literary character networks, in which the nodes are literary characters and the edges are co-occurrences in a series of scenes.
2. Word Co-Occurrence – Evolution of word use over time (word careers), in which the nodes are words and co-occurence in a text is represented by the edges.
3. Republic of Letters – Correspondence networks, in which the nodes are correspondents and the edges are letters.
4. Reports of Secret Police – Network model of the evolving knowledge of investigators into the relations of conspirators, in which each agent has a different view on the network of possible conspirators and nodes in the network appear and disappear as the police discover more about the network.

One idea that recurred frequently in our discussion was “snapshots” of a network which could be arranged into series by linking them to produce sequences instead of graphing a single temporal network. We designed and refined visualizations which could be used in each of these cases. Solutions included a stream graph of centrality and centralization, dyad visualization, temporal / witness matrix, collation networks, and a discourse/story/perspective model of

networks. Our solution to the problems presented by the variety and complexity of humanities data models was to combine network visualizations with representations of how the data was modeled–for example, placing network diagrams in a matrix that shows both the state of the network over time and how the network appears according to various perspectives which are made explicit in witness reports. By using a matrix, for example, we can show the state of a network across time according to various perspectives, such as witnesses to a series of events, or changes in the network.

For all four of our use cases, the combination of multiple visualisations was necessary to convey the most significant information about how the network was structured and how it developed over time; these visualisations could include a timeline or scatterplot to show the place in the temporal sequence of the network currently being visualized.

4.4 Uncertainty Visualization in Digital Humanities (DH) Network Data

Katy Börner (Indiana University – Bloomington, US), Nicole Coleman (Stanford University, US), Marten Düring (University of Luxembourg, LU), Tim Dwyer (Monash University – Caulfield, AU), Sara Irina Fabrikant (Universität Zürich, CH), Christina Gillmann (TU Kaiserslautern, DE), and Stefan Jänicke (Universität Leipzig, DE)

Uncertainty visualisation has been studied extensively in computer science, cartography, geography, information science, and related disciplines [1, 2, 3, 4]. Today, many individual solutions exist (e.g. [1, 2, 3]) within Digital Humanities (DH) projects but these have not yet been translated into generalizable solutions specific to either DH network data or the graph drawing community in general. Group members have diverse disciplinary backgrounds in computer science/visual analytics, information science, library studies, history, and GIScience/geovisual analytics. Such diversity within a group can be considered an advantage as it promises synergies between different conceptual frameworks. At the same time, it requires a shared terminology to avoid misunderstandings. Our team therefore drafted definitions of related key concepts such as error, uncertainty, graphs/network, sample bias, data quality, (in)homogeneity, probabilistic networks, data operationalization, and (in)compatibility with the goal to have them validated by the rest of the group. The representation of uncertainty in network visualisations in general and with regard to works in DH in particular, remains an open practical and research challenge. While specific problems such as the positioning of nodes due to uncertain node [5] or edge attributes [4] has received some attention in the past, to date there is no comprehensive overview of problems and recommended solutions. We expect that visual representations of uncertainty which were developed in other disciplines like information science, visual analytics, cartography, or bioinformatics can be adapted. This, however, raises the question to which extent data and research interests in DH differ from those in other domains. Is DH data special? The organizers of the Vis4DH workshop series [5] point to three distinctions: – differences in rhetorics of proof and discovery (and so differences in data culture and use) in the Humanities, – the difficulty of performing task analysis and evaluation for many humanities questions, that may have no ground truth, and finally, – in text visualization specifically, the difference between the needs of digital humanists (who perform close readings and critical engagements with texts) as opposed to

more standard text visualization scenarios (e.g. text analytics on datasets media analysis).

Here we seek to expand these observations further:

1. The data model itself is rhetorical in humanities research. Data modeling is part of an argument to be debated within the field. Data models therefore tend to be project-specific. So far, there are no generic data models which find general acceptance and are used for research across the DH.
2. Humanities data can be characterised by missing data, inhomogeneous representations of the available information and hard-to-resolve ambiguities. These problems appear as known unknowns and unknown unknowns.
3. Data visualisation is considered a (complementary) part of a research workflow alongside more traditional practices in DH.
4. Because historical datasets are inherently constructed, research data may include information from multiple sources, including attributes and values generated by the scholars. Data is seen to only partially represent scholarly knowledge. Interpretation requires enrichment with external information which defies representation in data.
5. Data analyses focus on data visualization as heuristic rather than as proof.
6. Data is often extracted or enriched manually which also explains the typically limited size of research datasets. Scholars seek to preserve the ability to manually edit, annotate and manage versions of their datasets.

This emphasis on often but not exclusively qualitative research practices, interpretation and personalised data models which also require representations of uncertainty [12] stands in contrast to the more empirical, probabilistic models of uncertainty which are common in other disciplines. A preliminary survey of interaction models developed in visual analytics [5, 6, 7, 8, 9, 10, 11] reveals that points 1 and 2 are not sufficiently captured in the existing approaches.

In order to properly define uncertainty visualisation challenges, we identified a preliminary taxonomy of different types of uncertainty in network visualisation. Existing taxonomies developed in other domains (cite) were reviewed and considered to be promising starting points. We identified four types of uncertainty (and their combinations) which we encounter frequently in network data:

- Time: When was it?
- Location: Where was it?
- Topic: What was it?
- Relation: What is the relation?

Our forthcoming survey paper will describe the state-of-the-art in depicting uncertainty in network visualisations, identify key challenges for Digital Humanities applications, and point towards best practices in cognate disciplines to resolve them.

References

1 G. Caviglia and N. Coleman, *Idiographic Network Visualizations* Leonardo, vol. 49, no. 5, pp. 447-447, Mar. 2016.

2 S. Jänicke and D. J. Wrisley, *Visualizing Uncertainty: How to Use the Fuzzy Data of 550 Medieval Texts?*, in Conference Abstracts of the Digital Humanities 2013, 2013.

3 P. Girard et al., *RICardo Project?: Exploring XIX Century International Trade*, HAL, Post-Print hal-01835245, Jul. 2016.

4 J. Schwank, S. Schöffel, J. Stärz, and A. Ebert, *Visualizing Uncertainty of Edge Attributes in Node-Link Diagrams*, in 2016 20th International Conference Information Visualisation (IV), 2016, pp. 45-50.

5 D. Sacha, H. Senaratne, B. C. Kwon, G. Ellis, and D. A. Keim, *The Role of Uncertainty, Awareness, and Trust in Visual Analytics*, IEEE Trans. Vis. Comput. Graph., vol. 22, no. 1, pp. 240-249, Jan. 2016.

6 T. Zuk and S. Carpendale, *Theoretical analysis of uncertainty visualizations*, presented at the Electronic Imaging 2006, San Jose, CA, 2006,

7 Torre Zuk, *Visualizing Uncertainty* Unpublished, 2008.

8 N. Andrienko et al., *Viewing Visual Analytics as Model Building* Comput. Graph. Forum, vol. 37, no. 6, pp. 275-299, 2018.

9 *(PDF) The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis* Online available: https://www.researchgate.net/publication/215439203_The_sensemaking_process_and_leverage_points_for_analyst_technology_as_identified_through_cognitive_task_analysis. Accessed: 29-Jan-2019.

10 *(PDF) Beyond Tools: Visual Support for the Entire Process of GIScience*, Online available: https://www.researchgate.net/publication/200043275_Beyond_Tools_Visual_Support_for_the_Entire_Process_of_GIScience. Accessed: 29-Jan-2019.

11 M. El-Assady et al., *Visual Text Analytics in Context of Digital Humanities*, Published at the 1st IEEE VIS Workshop on Visualization for the Digital Humanities as part of the IEEE VIS 2016, 2016.

12 *nodegoat | Blog | Formulating Ambiguity in a Database*, nodegoat, Online available: https://nodegoat.net/blog.s/21/formulating-ambiguity-in-a-database. Accessed: 19-Feb-2019.

Participants

- Gregor Betz
KIT – Karlsruher Institut für Technologie, DE
- Katy Börner
Indiana University – Bloomington, US
- Ulrik Brandes
ETH Zürich, CH
- Nicole Coleman
Stanford University, US
- Melanie Conroy
University of Memphis, US
- Marten Düring
University of Luxembourg, LU
- Tim Dwyer
Monash University – Caulfield, AU
- Oyvind Eide
Universität zu Köln, DE
- Kimmo Elo
University of Turku, FI
- Sara Irina Fabrikant
Universität Zürich, CH
- Christina Gillmann
TU Kaiserslautern, DE
- Hans Hagen
TU Kaiserslautern, DE
- Francis Harvey
Leibniz Institut für Länderkunde – Leipzig, DE
- Gerhard Heyer
Universität Leipzig, DE
- Stefan Jänicke
Universität Leipzig, DE
- Fotis Jannidis
Universität Würzburg, DE
- Andreas Kerren
Linnaeus University – Växjö, SE
- Stephen G. Kobourov
University of Arizona – Tucson, US
- Tamara Mchedlidze
KIT – Karlsruher Institut für Technologie, DE
- Martin Nöllenburg
TU Wien, AT
- Malte Rehbein
Universität Passau, DE
- Gerik Scheuermann
Universität Leipzig, DE
- Antonis Symvonis
National TU – Athens, GR
- Timothy Tangherlini
University of California at Los Angeles, US
- Christopher Warren
Carnegie Mellon University – Pittsburgh, US
- Scott Weingart
Carnegie Mellon University – Pittsburgh, US
- Florian Windhager
Donau-Universität Krems, AT