

Augmenting Network Analysis with Linked Data for Humanities Research



Jörn Kreutel

Abstract We will outline the motivation of a web based software tool for social network analysis that particularly aims at enriching the exploration of network visualisations by linking them back to the original data from which networks have been created. We will identify three aspects of enrichment, present the architecture of an existing prototypical implementation and compare its scope with the one of an exemplary tool for SNA that is widely used within the digital humanities. With respect to the latter, we see evident advantages of our solution when it comes to supporting a seamless workflow of network creation, analysis and exploration based on parameterisable queries of the underlying data sources.

Keywords Social network analysis · Digital humanities · Semantic web

1 Introduction

Social network analysis (SNA, see, e.g., [6, 22]) has been found to be a promising quantitative approach offering various fields of application in the area of humanities and cultural studies [10, 16]. Complementing concepts like, e.g. Bourdieu's idea of *field* [8] SNA can be employed for reconstructing the mechanisms that underlie particular actors' careers within a given discipline [7, 11], and thus may contribute to understanding processes of reception like an actor's canonisation or decanonisation, which appear relevant from the point of view of the emergence of cultural heritage. In fact, applications of SNA range beyond its original focus on relationships between social actors towards a coverage of any phenomena that can be represented by graphs and analysed in terms of the formal properties of a graph's nodes and edges [21].

However, actually employing SNA for particular research objectives requires appropriate technical tools for creating and analysing networks and visualising analysis results, which should allow for a seamless workflow at the level of the employed user interfaces. Here, we will present and discuss the prototype of a software sys-

J. Kreutel (✉)

Beuth University of Applied Sciences, Luxemburger Strae 10, 13353 Berlin, Germany
e-mail: joern.kreutel@beuth-hochschule.de

© Springer Nature Switzerland AG 2020
H. Kremers (ed.), *Digital Cultural Heritage*,
https://doi.org/10.1007/978-3-030-15200-0_1

tem und user interface supporting SNA and associated statistical analyses on the underlying data, which has been developed in the course of ongoing research on bibliographical data, dealing with GDR literature, on the one hand, and the history of German literature studies, on the other. Considering the current state of our system, the objective of this paper is to raise the question of how SNA can be integrated within a wider scope of applying quantitative methods in the humanities and what tools are required for this purpose.¹

In particular, our prototyping efforts addressed the question how an integrated workflow can be supported for SNA and the exploration of analysis results on the basis of various types of linkable data, including databases and web pages. As we will argue below, this requires appropriate expressive means that go beyond the scope of tools for mere network analysis and visualisation (see [1] for an overview and [2] as a prominent example) and provide enhancements both at the level of data evaluation and the user interface. Adhering to the idea of a widely accessible platform for supporting augmented SNA as a *Software as a Service* solution [17], our prototype is based on standard web technologies and can be accessed by any state-of-the-art web browser, rather than requiring individual installations.

In the following sections, we will outline the rationale for augmenting network analysis and enhancing analysis tools, looking at how networks to be analysed may be created as abstractions from arbitrarily rich underlying data. We will then describe three aspects of enhancements that are currently supported by our prototype and exemplify their usage with examples from bibliographical and co-occurrence networks. After a brief outline of the software architecture of the system, we will compare its functionality and expressive means with what could be achieved using *Gephi* [2] as one of the most prominent and established SNA tools in the area of digital humanities. [13, 24].

¹Starting point of system development was the attempt to create a proof-of-concept implementation for integrating network creation based on SPARQL [12] queries, network analysis and network visualisation, using co-occurrence networks of drama characters as an example case (see Sect. 3). The requirements of the system then have been iteratively extended based on the author's cooperation with domain experts, and functionality has been verified by intense system usage of the author himself and within joint workshops with the experts. However, the system is prototypical in the sense that it has not been used by the latter experts independently of the author. In addition to DCH 2017, the ideas outlined here have been discussed with SNA experts from the digital humanities community, see *A Web Based Workbench for Augmenting Network Analysis with Linked Data Evaluations*, presented by Jörn Kreutel at *Potsdamer Arbeitstreffen zur digitalen Literaturwissenschaft No. 2*, December 4, 2017 (<https://www.uni-potsdam.de/lit-19-jhd/digitale-literaturwissenschaft/potsdamer-arbeitstreffen/no2-2017.html>). First findings from co-publication analyses on a corpus of works on German literature studies have been presented by Steffen Martus and Fotis Jannidis in *Was verändert sich eigentlich? Korpusanalytisch basierte Wissenschaftsgeschichte der germanistischen Praxis* at *DFG-Symposium 'Digitale Literaturwissenschaft'* on October 10, 2017 (<http://www.germanistik.uni-wuerzburg.de/en/lehrstuehle/computerphilologie/aktuelles/veranstaltungen/dfg-symposium-digitale-literaturwissenschaft/>).

2 Network Creation as Data Abstraction

As outlined in, e.g. [6, 22], SNA is usually done on networks the nodes of which either represent entities of a single type—e.g. persons—or, in the case of *bipartite* networks, belong to two different types of entities, whose affiliation with each other is modelled by the network—as in a network that models people’s attendance to events or their association with institutions. In any case, the connecting edges of the network represent a single type of, possibly weighted and either directed or undirected, relation with respect to which the quantitative properties of the network may be determined.² These relations might be concrete relations or abstractions from the latter, e.g. the two fairly concrete relations of persons being friends or colleagues might be abstracted to the relation of persons knowing each other which could be assigned a weight depending on the number and type of its underlying concrete relations.

Hence, in order to employ SNA for some given domain, the networks to be analysed may not necessarily be represented by the primary data available, but might need to be constructed from the latter, which might involve the assumption of abstract relations like the ones mentioned above. In particular, this is the case if SNA shall be applied on data the model of which has not been designed with the idea of applying SNA in mind, but that, nevertheless, can be interpreted in terms of relations between associated entities from a given research perspective. An example for this is bibliographical data, a possible structure of which is exemplified in Fig. 1. In order to analyse, e.g., the relationships between authors that are constituted on the basis of the common appearance of authors’ contributions in collections or anthologies, the actual relations that will be represented by the network’s edges need to be extracted from the bibliographical data and be weighted, e.g., depending on the number of joint appearances. At the level of the network, the detail information on the associated entities and the grounds of their associations will then not be present anymore. However, both for exploring the network at the level of an interactive visualisation and for quantitatively analysing the network, additional information on the associated and associating entities—like an author’s name, age and their affiliation with institutions, or the collections’ titles, publishers and years of appearance, respectively—might be necessary for a holistic understanding of the phenomena to be investigated. For bibliographical analyses, this holds, for example, if different networks of the same type, different partitions of an overall network or networks for different time intervals shall be compared to each other in terms of domain specific ‘key performance indi-

²For example, the relevance of a node within a network may be measured applying different concepts of *centrality* like *degree*, *eigenvector* or *betweenness* centrality. Using methods of *community detection*, on the other hand, nodes can be classified as belonging to different groups or ‘cliques’ [6, 22].

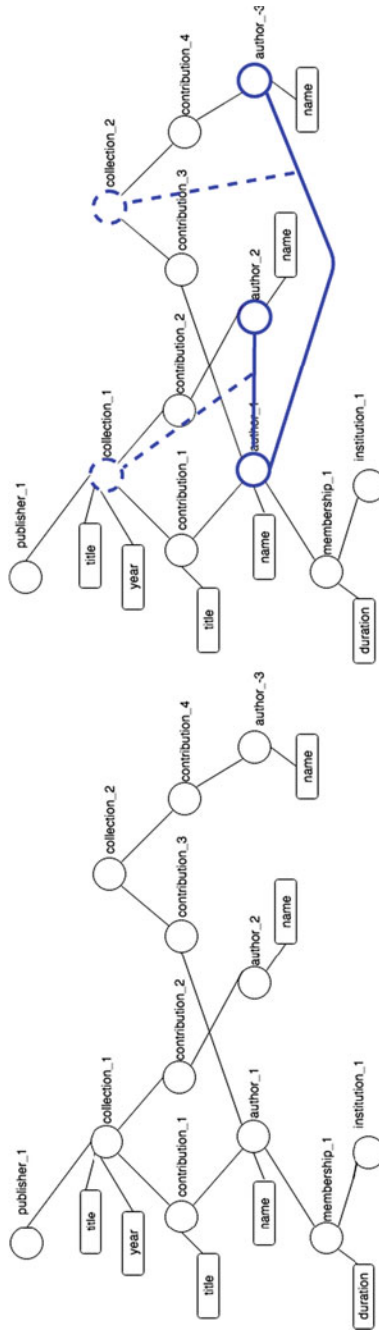


Fig. 1 Bibliographic sample data (left) and abstraction of network relations based on authors' contributions to collections (right). Here, co-publication relations can be assumed between *author_1* and *author_2*, and *author_2*, and *author_3*, on the basis of their contributions to *collection_1* and *collection_2*, respectively

cators' like the network authors' overall publication output or their association with publishers and other publishing media than the ones that constitute the network.³

As for the user interfaces employed for SNA studies in some particular domain and the expressive means for quantitative analyses, tools therefore need to allow for a certain degree of re-concretisation of abstracted network data, the extent of which will depend on the given overall research interests and the possibly iteratively refined proceeding of the investigation. For the particular case of bibliographic analyses, Klink et al. [14] propose a powerful tool that provides advanced browsing capabilities for the DBLP computer science bibliography (<http://dblp.uni-trier.de/>), including the creation and visualisation of network analyses on the basis of the bibliographic data. However, the *DBL Browser* is a standalone software solution to be run locally on each user's machine, and it is restricted to the domain of bibliographic analyses, which is a relevant, but yet not the only domain where advanced capabilities for network analyses are desirable. Hence, we find it worthwhile to conceive of a domain independent tool, which rather than being a complete, but monolithic solution provides a comparably thin integration layer that mediates between the primary data sources for SNA, the execution and visualisation of SNA and enhanced statistical analyses, and the exploration of analysis results linking back to the original data and other linkable data sources. The latter does not imply the provision of full-fledged browsing capabilities on the part of the integration layer, but may at least partially rely on existing web based solutions for browsing content like the viewers provided by the databases themselves.⁴

3 Enhancements for Network Exploration and Analysis

For the desired re-integration of primary data and the further integration of associated data sources, we identified three major 'extension points' at the level of a user interface for SNA, further described below. The extension points *network manipulation* and *network browsing* are related to the level of the visualisation of a single network as a graph and the interactive exploration of its content. *network statistics*, on its part, may apply to the creation of statistical analyses to sets of networks, which may represent, e.g., sequences of stages of a network at different time intervals, different partitions of a single network, or different networks representing different relations of actors, which shall be compared to each other:

1. **Network Manipulation:** When exploring a network starting from a visual representation, it is desirable to be able to manipulate this visualisation not only on the basis of genuine SNA measures like the *centrality* of nodes, the partitioning

³These requirements are based on a synopsis of work on scientific networks and the types of analyses employed therein, e.g. [9, 20, 26], as well as on desiderata that were raised in the course of carrying out our own analyses on bibliographical data.

⁴See, e.g., the RDF content browsing facilities of the *GraphDB* database (<http://graphdb.ontotext.com/>), which we employ for data storage.

of nodes in different *communities* or the strength of edges [6, 22], but also with respect to other attributes of nodes and edges that might not have been themselves subject of SNA measurements. For example, in a co-publication network, the size of nodes might be modified, as shown in Fig. 2, depending on the overall publication output of the author represented by the node, and edges might be subject to applying measures considering, e.g., the seniority of related authors in terms of age or status [20]. Hence, the user interface for graph visualisation should allow for domain specific extensions⁵ that support the consideration of such additional criteria in order for domain experts to gain a more wholistic view of a network, and to be able to interactively identify sub-networks, which may then, on their part, be subject to dedicated SNA measurements.

2. **Network Browsing:** In addition to network manipulation, which operates with domain specific measurements as quantitative abstractions over concrete data associated with nodes and edges, network exploration might also involve the insight into the concrete and more complete data underlying the latter. For example, co-publication networks might be linked to structured or unstructured data related to their nodes and edges such that, e.g., the available database content on authors or content from additional data sources like online encyclopaediae will be displayed when interacting with their visualisations. As for another prominent field of applying SNA in the humanities, edges of co-occurrence networks of actors extractable from literary works [15, 25] might be linked to the actual text fragments within the respective works that are constitutive for the edges. For this purpose, the user interface must support the domain specific creation of views that will display the associated information and integrate them within the overall navigation structure for network exploration. Figures 2 and 3 show how these requirements are currently dealt with by our prototype.
3. **Network Statistics:** Particularly when dealing not only with a single network, but with several networks or several partitions of a network, it is desirable to gain an overview both of the SNA core measures, in the sense of [22], of the respective networks or sub-networks, and of associated domain specific measurements applicable to the networks' nodes and edges. For example, if bibliographical data is enriched by additional biographical data, age profiles or seniority profiles for the authors of co-publication networks can be calculated and analysed, e.g., in terms of temporal development or with respect to different sub-networks of authors, as shown in Fig. 4. Likewise, the overall publication output of network authors beyond a given network of co-publications can be made subject to comparative evaluation by applying appropriate calculations. Considering not only aspects like the quantity of output, but also, e.g., publishers or journals in which contributions appear, might help to identify network partitions as, e.g., 'mainstream' versus 'elite' versus 'niche' sub-networks. Hence, the tools employed for network analysis need to support the creation and visualisation of such additional

⁵As we will argue in Sect. 5, this does not require that the additional information on nodes and edges is represented by the network itself, but only that it is retrievable based on the given representation of the network.

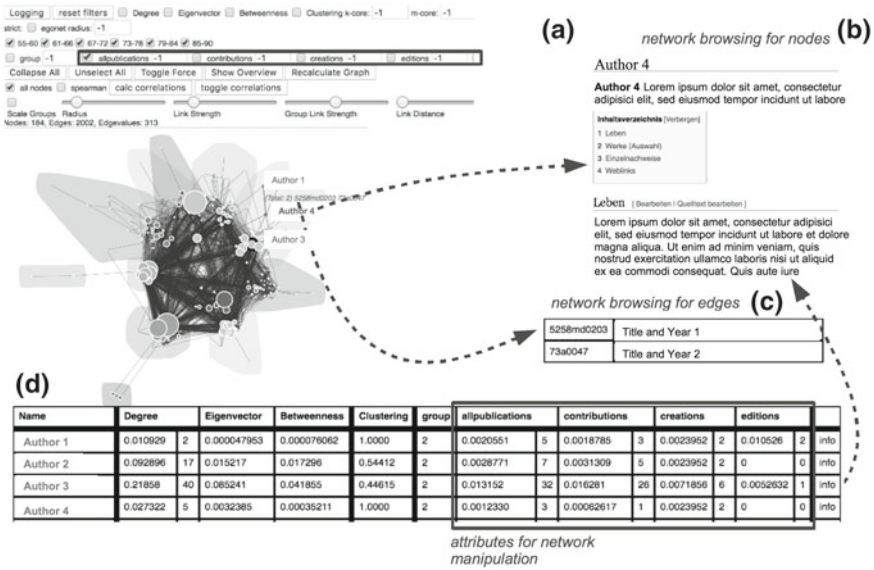


Fig. 2 Examples for domain specific *network manipulation* and *network browsing* on the basis of a co-publication network. Size of network nodes can be manipulated, and nodes can be filtered on the basis of the highlighted domain specific attributes, where control elements are provided in addition to the elements for base SNA measures (a, d). Selecting a node will navigate the user to an online encyclopedia entry of the respective author, if available (b). On selection of an edge, the list of collections to which the associated authors have contributed will be displayed (c)

measurements, which like the other two extension cases enrich the means for core network analysis, in order for researchers to gain wider quantitative insight into a given domain.

In addition to these functional requirements on enhanced network analysis, which are formulated from a domain expert’s perspective, both for sustainable continuous work within particular research projects and for transparently disclosing the methods underlying their outcomes, analysis tools should support a modular organisation of the technical artifacts, e.g. queries and processing scripts, to be created for domain specific analyses, aiming at seamless reusability and extendibility. The following sections will give a brief overview of how these technical requirements are addressed by our current prototype.

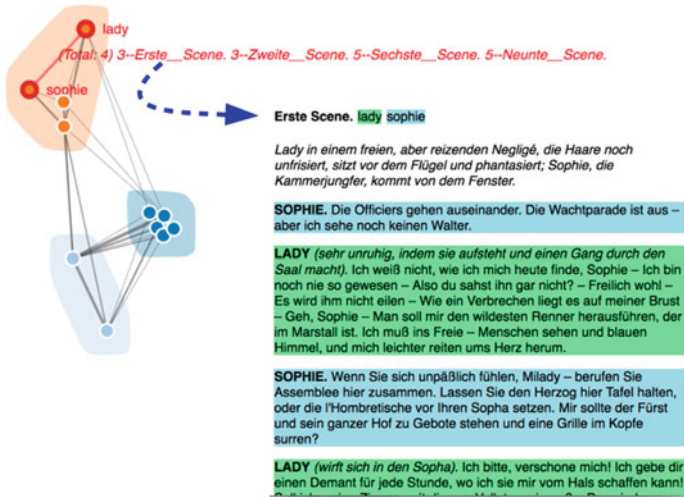


Fig. 3 Example for *network browsing* of a co-occurrence network. Selection of an edge value will navigate the user to the selected scene in which the two characters appear. Texts are taken from an online corpus made available through the *Project Gutenberg* initiative (e.g. <http://gutenberg.spiegel.de/>), which also serves as the basis for identifying co-occurrence

4 System Architecture

The realisation of the overall system architecture shown in Fig. 5 comprises a rich client user interface using state-of-the-art web technologies⁶ and a JavaScript backend provided by a *Node.js* (<https://nodejs.org>) server for managing domain specific configuration settings and extensions, and for executing the overall workflows for network analysis and enhanced network statistics in the above sense. Adhering to the overall idea of the semantic web [3] and the support of technologies compatible with linked open data [5] infrastructures, primary data from which analysable networks are extracted are stored in an RDF [19] repository, read and write access to which is mediated by a web service that uses the Java *rdf2j* library (<http://rdf4j.org/>) for processing SPARQL [12] queries. For the core network analysis employing SNA measurements, we run the *NetworKit* toolkit [23] within a thin server layer written in Python.

The support for domain specific enhancements of network exploration and analysis is based on the employment of dynamic JavaScript execution of custom implementation artifacts both within the browser and the Node.js backend. Both environments provide support for executing queries on the underlying data and ex-

⁶In particular, we use the *D3.js* framework (<https://d3js.org/>) for visualising graphs and statistical analyses, and *CodeMirror* (<https://codemirror.net/>) for displaying and editing implementation artifacts within the web browser. Starting point for the concrete graph visualisation was the open source solution proposed on <http://bl.ocks.org/GerHobbelt/3071239>.

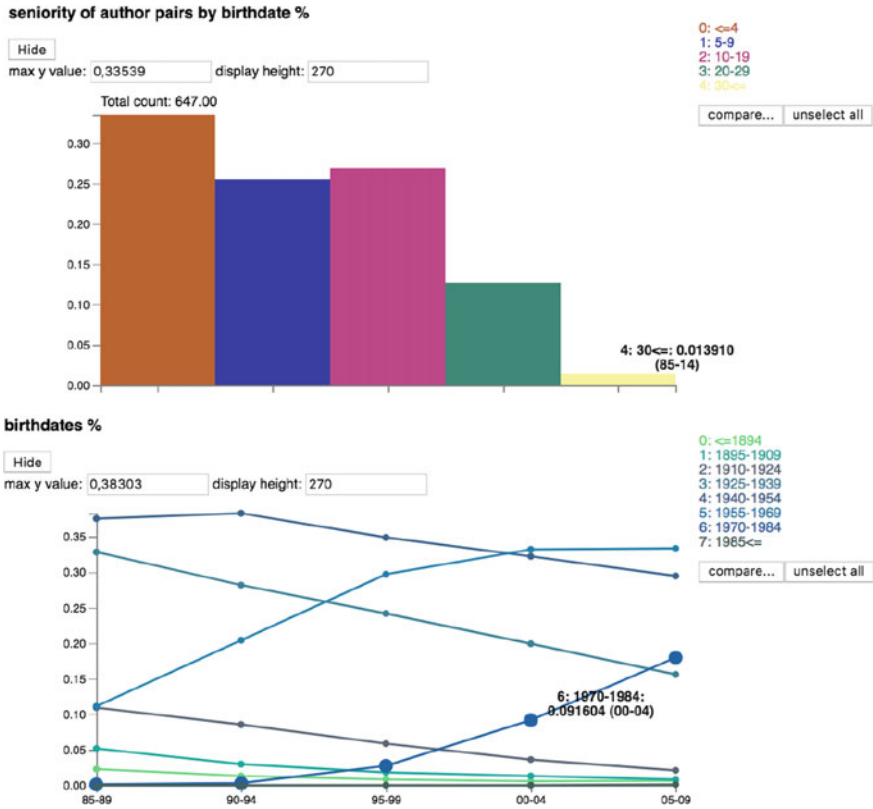


Fig. 4 Examples for *network statistics*. The bar chart (top) shows the age difference, as one aspect of seniority, of author pairs in a co-publication network, whose underlying data has been enriched with biographical information. The line chart (bottom) displays, for a complete bibliography, the shares of authors’ age groups with respect to all publishing authors over a sequence of time intervals. Here, one can observe, e.g., a fairly parallel development of the age groups 1955–1969 and 1970–1984 with regard to their positioning within the overall field

ternal data sources, including access to web pages, and offer a library of data types and operations on the latter for processing query results. In particular, results may be represented as sets of tuples or interpreted as graphs, and operations comprise, among others, set operations like intersections and unions, as well as filtering and partitioning operations on these types, and the application of statistical operations like correlation analysis.⁷

For implementing domain specific artifacts that realise one or more of the extension points outlined in the previous section, users create JavaScript components,

⁷Using these means, for example, a set of publications can be filtered on the basis of the nodes of a co-publication network in order to determine publication output, publishers, journals, etc. for the authors of the particular network. Additional statistical analyses can be done using the JavaScript *jstat* library (<https://github.com/jstat/jstat>).

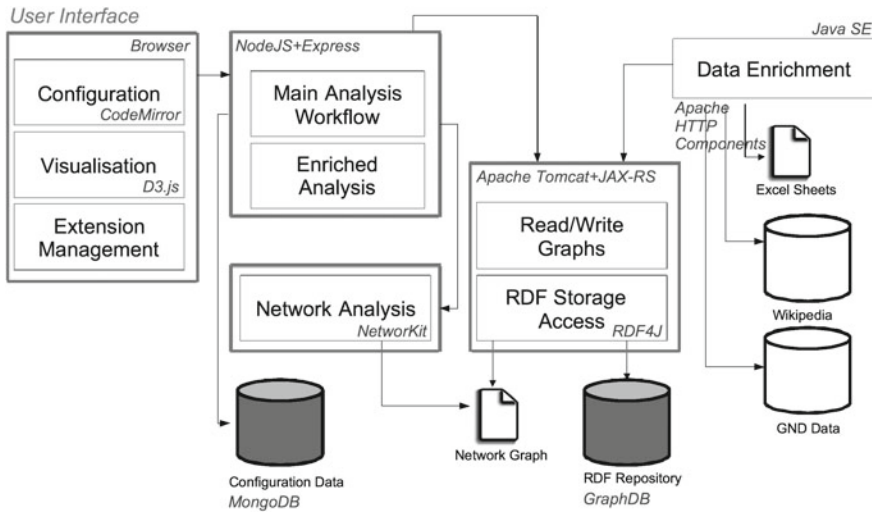


Fig. 5 System architecture and technologies used for realisation. On the right hand, different examples for enriching primary data using content from linkable structured or unstructured data sources like the GND authority file (http://www.dnb.de/EN/Standardisierung/GND/gnd_node.html) or Wikipedia, respectively, are shown

which can be either uploaded and imported via the user interface or edited within the latter. These components are named and can be reused for analysing any networks that have a common structure and are based on structurally compatible primary data, for example co-publication networks that are created from data sources with shared RDF vocabulary. Additionally, components are subject to an inheritance mechanism inspired by object oriented programming languages and can extend and override each other's functionality.

Apart from these possibilities to create domain specific extensions for network analyses, the prototype allows for the, even more fundamental, creation of networks based on 'query patterns' for the underlying primary data. In our particular case, we use SPARQL queries that contain variables which are filled by the system before executing the query. This way, it is possible, e.g., to create a set of temporal 'snapshots' of a network for different subsequent or overlapping time intervals, each of which will be subject to core SNA and possibly further measurements. At the level of network visualisation, snapshots can be integrated into a single view. Filtering capabilities will then allow to visualise the dynamic evolution of a network over time, including the application of enhancements for network manipulation and browsing mentioned above.

Given this overall functionality, the system does not only support a seamless workflow for network creation, analysis, visualisation and exploration, and the creation of network statistics from a single web based user interface. It also allows for a clear division of labour between domain experts using the system for getting insight into

a particular data set, and research engineers who create domain specific implementation resources based on the experts' requirements and make them available using the system's extension points.⁸ Here, modularisation and extension capabilities at the level of the domain specific programming artifacts themselves are important features for ensuring maintenance and extendability of artifacts for longer running research with, expectedly, iteratively refined perspectives on the data to be investigated.

5 Related Work

Particularly within the digital humanities community as a our primary field of application, Gephi [2] seems a widely used tool for network analysis, visualisation and exploration.⁹ For comparing the existing and envisaged functionality of our prototype, Gephi therefore appears to be an appropriate benchmark. See [1] for a wider comparison of existing tools, which, however, only focusses on the mere SNA related functionality, rather than the enhancements to SNA in the sense of the re-concretisation of abstract networks that has been a main motivation of our own development efforts.¹⁰

To mention, at the beginning, a major drawback of our own solution, our web based prototype currently is by far less powerful than Gephi and the other tools analysed in [1] with regard to the performance of rendering network visualisations and interactively exploring a network. It runs into trouble if networks with more than 10,000 nodes and/or 30,000 edges shall be visualised in an expanded way, i.e. not only at the macroscopic level of a network's communities.¹¹ In such cases, filters need to be applied in order to obtain a sub-network with handleable size. For improving performance it appears feasible to change the underlying framework for graph visualisation from SVG based *D3.js* (<https://d3js.org/>) to a more powerful alterna-

⁸A systematic reflection on these roles and their respective contribution to the joint research effort of digital humanities projects has only started recently. See, e.g., the working group on *research software engineering in the digital humanities* (<http://dig-hum.de/ag-research-software-engineering-digital-humanities>) that dedicates itself to this purpose.

⁹This judgement is based on the author's impression gained from research literature and blog posts on concrete SNA usage (see, e.g. <https://dlina.github.io/Mayakovsky-Klop/>) and has been confirmed by members of the digital humanities community. For a public, yet informal, source see <https://en.wikipedia.org/wiki/Gephi>.

¹⁰As for SNA itself, our architecture is designed in a modular way such that *NetworKit*, which is currently used, could be replaced with any other tool offering an API that can be integrated and exposed by a web service, see, e.g., the SNA package for the R statistical computing environment (<https://cran.r-project.org/web/packages/sna/index.html>). Likewise, at the level of the user interface, the setting of parameters for SNA has been designed as an extension point on its own and, hence, would be replaceable by tool-specific settings. Finally, access to the actual network data and the results of network analysis is mediated by the Java backend service, which could be extended to cover alternative graph representations if necessary, or be replaced with an alternative service adhering to the latter's API.

¹¹In contrast, the comparative analysis in [1] uses a network with 77,317 nodes and 982,787 edges.

tive, e.g. *sigma.js* (<http://sigmajobs.org/>) that uses WebGL.¹² However, particularly if both quantitative and qualitative perspectives are meant to be applied, network visualisation in the humanities might not necessarily deal with large amounts of data,¹³ but strongly requires user interfaces that are able to mediate between the two perspectives.

Gephi, on its part, is designed as an extendable infrastructure that offers a plugin mechanism for domain specific customisation and extensions of the tool's core functionality. As for the three aspects of re-concretisation mentioned in Sect. 3, it mainly supports domain specific extension points for network manipulation and network statistics. For example, it allows to implement customised filters that run on the graph data itself or on external data and modify the visualisation (see, e.g., <https://github.com/gephi/gephi/wiki/How-to-use-filters>). It is also possible to execute customised statistic evaluations, primarily on the graph data itself, but also here, external data could be included (see <https://github.com/gephi/gephi/wiki/Statistics>). As Gephi is built on top of the Java *NetBeans* platform (<https://netbeans.org>) and allows to capture interaction events at the level of graph visualisation, extensions for *network navigation* in the sense described above, also appear possible. However, even though installation and update of custom plugins can use an integrated distribution mechanism, usage of plugins will always be bound to concrete Gephi execution environments, rather than being independent of any workplace as in the case of a web based platform like ours.

Principally, Gephi supports domain specific data enrichment either at the level of the network graphs themselves at the time of graph creation, by allowing to enhance a graph's nodes and edges with domain specific attributes, or at the time of graph exploration by appropriate data access implementations. However, this flexibility may result in monolithic solutions that try to deal with any extensions at the level of the network graph, outside of the working environment provided by Gephi. In this case, repeated graph creation would be required as research proceeds and new perspectives on the network's actors and relations arise without the network itself being changed.¹⁴ In contrast to this, our approach to network enrichment, be it for the purpose of manipulation, navigation or statistics, is consequently modular by separating network data from data used for network enrichment and integrating the two aspects dynamically at runtime, where caching mechanisms are employed to avoid processing overhead. This way, enrichments of any type can be reused for analysing and exploring all networks whose nodes and edges have a common type, regardless of the initial research perspective and interests at the time of their creation.

¹²See <https://www.w3.org/TR/SVG2/> and <https://www.khronos.org/registry/webgl/specs/latest/2.0/> for the specifications of SVG and WebGL, respectively.

¹³See, e.g., the case of co-occurrence analysis [25] or the analysis of particular circles of actors and their interrelations [4, 20].

¹⁴This assumption is not based on empirical grounds, but on the author's view of the range of possibilities to employ the expressive means provided by the Gephi infrastructure. However, a random overview of respective tutorials and forums on the latter does at least not strongly hint at a wide spread usage of exploration time data enrichment.

Considering network graph creation, finally, our prototype originated from the attempt to integrate graph creation from primary data in RDF, accessed via SPARQL queries that can be formulated as query patterns, as mentioned above.¹⁵ This way, it is not only possible to easily create sequences of network ‘snapshots’, e.g. for subsequent time intervals that can be integrated to a joined graph, but also to make the snapshots’ parameters available to all extensions that need to take them into consideration. Hence, there is a built-in mechanism for parameterising data enrichment to the given scope from which a network is being viewed and analysed. For Gephi, there also exists a plugin for creating Graphs using SPARQL (<https://github.com/gephi/gephi/wiki/SemanticWebImport>). Yet, it does not come with dedicated expressive means for pattern based dynamic graph creation and enrichment. Hence, for network graph creation itself external tools would need to be employed. However, from the point of view of our own research and other research interests in the humanities, the evolution of networks will always be an important aspect of the domains to be analysed and should therefore be supported in a seamless way by an integrated platform for network creation, analysis and exploration.

6 Summary

Starting from the observation that the application of SNA on a given data set may require the creation of network graphs that abstract away from many details of that data, we identified three generalisable aspects of re-concretisation that may be provided by software tools for graph analysis and exploration. We then outlined the software architecture of our own prototypical tool, which particularly addresses these requirements, and compared its functionality and overall approach with the one of Gephi as a prominent solution in the area of digital humanities. We see clear advantages of our solution not only by providing a web based environment, but also considering its support of a seamless workflow for graph creation, analysis and exploration based on a modular organisation of reusable domain specific extensions. Further work will focus on improving graph visualisation performance, on the one hand, and on further enhancing the functionality of the tool’s user interface, on the other. In terms of [18], we thus aim at further supporting a seamless mediation between the aspects of *distant reading* and *close reading* with regard to a particular research objective, by linking networks and statistical analyses as abstractions to the concrete data and content from which they originated.

¹⁵As data access via SPARQL is encapsulated into dedicated components, alternative data sources and query languages like SQL could be integrated into the system infrastructure without any changes to the overall system architecture.

References

1. Akhtar, Nadeem: Social network analysis tools. In: 2014 Fourth International Conference on Communication Systems and Network Technologies (CSNT), pp. 388–392. IEEE (2014)
2. Bastian, M., Heymann, S., Jacomy, M., et al.: Gephi: an open source software for exploring and manipulating networks. In: International AAAI Conference on Weblogs and Social Media. (2009)
3. Berners-Lee, T., Hendler, J., Lassila, O., et al.: The semantic web. *Scientific american* **284**(5), 28–37 (2001)
4. Biehl, T., Lorenz, A., Osierenski, D.: Exilnetz33. Ein Forschungsportal als Such- und Visualisierungsinstrument. In: C. Baum, T. Stäcker (eds.) *Grenzen und Möglichkeiten der Digital Humanities* (Zeitschrift für digitale Geisteswissenschaften, Sonderband 1) (2015)
5. Bizer, C., Heath, T., Tim, B.L.: Linked data – the story so far. *International Journal on Semantic Web and Information Systems* **5**, 1–22 (2009)
6. Borgatti, S.P., Everett, M.G., Johnson, J.C.: *Analyzing social networks*. SAGE Publications Limited (2013)
7. Bottero, W., Crossley, N.: Worlds, fields and networks: Becker, Bourdieu and the structures of social relations. *Cultural Sociology* **5**(1) (2011)
8. Bourdieu, P.: *Die Regeln der Kunst. Genese und Struktur des literarischen Feldes*. Suhrkamp (1999)
9. Burckhardt, D.: Comparing disciplinary patterns: Exploring the humanities through the lens of scholarly communication. *digital humanities quarterly* **11**(2) (2017)
10. Comsa, M.T., Conroy, M., Edelstein, D., Edmondson, C.S., Willan, C.: The french enlightenment network. *The Journal of Modern History* **88**(3), 495–534 (2016)
11. De Nooy, W.: Fields and networks: correspondence analysis and social network analysis in the framework of field theory. *Poetics* **31**(5-6), 305–327 (2003)
12. Harris, Steve and Seaborne, Andy: SPARQL 1.1 Query Language. <https://www.w3.org/TR/sparql11-query/> (last accessed on March 21, 2018) (2013)
13. Jannidis, F., Kohle, H., Rehbein, M. (eds.): *Digital Humanities*. Metzler (2017)
14. Klink, S., Reuther, P., Weber, A., Walter, B., Ley, M.: Analysing social networks within bibliographical data. In: S. Bressan, J. Küng, R. Wagner (eds.) *Database and Expert Systems Applications*, pp. 234–243. Springer Berlin Heidelberg, Berlin, Heidelberg (2006)
15. Knuth, D.E.: *The Stanford GraphBase: a platform for combinatorial computing*, vol. 37. Addison-Wesley Reading (1993)
16. Meirelles, I., Schich, M., Malina, R.: Arts, humanities, and complex networks. In: R. Alhaji, J. Rokne (eds.) *Encyclopedia of Social Network Analysis and Mining*. Springer (2014)
17. Mell, P., Grance, T.: *The NIST definition of cloud computing*. Special Publication 800-145, National Institute of Standards and Technology (2011)
18. Moretti, F.: *Graphs, maps, trees: abstract models for a literary history*. Verso (2005)
19. RDF Working Group: *Resource Description Framework (RDF)*. <http://www.w3.org/RDF/> (last accessed on March 21, 2018) (2104)
20. Rehr, M., Palonen, T., Gruber, H.: Professionelle Vernetzung in der Wissenschaft. *Unterrichtswissenschaft* **38**(4), 334–353 (2010)
21. Sarv, Mari and Laineste, Liisi: *Network Analysis in Humanities – Session at the Second European Conference on Social Networks*. <https://eusn2016.sciencesconf.org/89668/document> (last accessed on March 21, 2018) (2016)
22. Scott, J.: *Social Network Analysis*, 3 edn. Sage (2012)
23. Staudt, C., Sazonovs, A., Meyerhenke, H.: *NetworKit: An interactive tool suite for high-performance network analysis*. CoRR, abs/1403.3005 (2014)
24. Svensson, P.: The landscape of digital humanities. *digital humanities quarterly* **4**(1) (2010)
25. Trilcke, Peer and Fischer, Frank and Göbel, Mathias and Kampkaspar, Dario: *Digital Literary Network Analysis*. <http://dlina.github.io/> (last accessed on March 21, 2018) (2015)
26. Wellmon, C., Piper, A.: *Publication, power, and patronage: On inequality and academic publishing*. *Critical Enquiry* (forthcoming) (2017)