Aristoteles multimodal - with *ediarum* to the graph

by Martin Fechner and Andreas Kuczera

Based on the research data of the academy project "Commentaria in Aristotelem Graeca et Byzantina" (CAGB), this contribution shows the possibilities of using a graph database for the analysis of research data and why interfaces of digital editions are useful for graph databases.

At the Berlin-Brandenburg Academy of Sciences and Humanities (BBAW) the research software *ediarum* has been increasingly used in recent years for the input, management, and presentation of digital scholarly editions. *ediarum* allows transcribing, and adding mark-up and comments on historical texts, as well as linking them to other texts and register entries. The success of recent years is proven in the high demand and the growing number of editions created and operated with *ediarum*.[1]

The *ediarum.WEB* module is a web framework that simplifies the creation of digital scholarly editions. The basic idea is to build a sustainable program library that bundles the shared functionalities and strictly separates backend and frontend (Fechner2018, Fechner2018a). By using this web framework, the digital scholarly editions benefit from all available and future developed features. Among the already existing functionalities are displaying and filtering of individual texts and register entries of a digital scholarly edition, as well as accessing to data and metadata via interfaces. A predefined interface allows a uniform access to the relationships of objects contained in the database. This has the advantage that one can connect a graph database to the interface, allowing the information to be visualized as a graph. This type of visualization enables new analytical approaches.

In recent years, graph databases have developed into a flexible tool for the management of highly linked research data. In contrast to relational databases, where data is stored in linked tables, graph databases have nodes and edges (Kuczera2017 and Kuczera2018).

Research data as graph structure

The academy project "Commentaria in Aristotelem Graeca et Byzantina" (CAGB) produces such highly linked research data with the help of *ediarum,* including a detailed manuscript catalogue containing ancient and medieval commentaries on works by Aristotle. In addition to the descriptions of manuscripts and persons, especially the relationships between the data objects contain important information for Aristotle research. These ultimately form a graph and can be analyzed with the help of graph-based tools. The possibility of directing precise queries to such a graph allows answering research-relevant questions, which previously required additional work. The simultaneous contextualization of manuscripts,

---

1 Currently, about 20 editions use the core components of ediarum or individual modules from the ediarum family (vgl. Fechner2017). URL: http://www.bbaw.de/telota/software/ediarum.

persons and Aristotle's works in a common graph allows the revaluation of the known relations.

The advantages of a graph-based analysis are that the data can be displayed as a coherent network graph. Specifically, a graph database enables the search and display of relationships across multiple nodes.

As an example, some use cases are listed which are improved by using a graph database. Research is interested in texts to which a certain person had access. However, in printed catalogues as well as in the database this information is only implicitly available: The manuscript descriptions contain information about the owners of manuscripts and the texts contained. In a graph-based database, this is a simple query and the results can be displayed not only as a list of texts for one person, but as a list or network for all persons and texts. Graph databases can handle queries that go beyond this. In the context of the CAGB project, for example, it is of interest which owners and copyists of the "Analytica Priora" there were in the 14th century or which persons had simultaneous access to a particular text corpus. A good overview of the research situation regarding the reception of certain texts is available by searching for individual texts, the manuscripts they contain and the corresponding owners.

Through the analyses, it becomes clear that the data model used to date is still inadequate to answer further questions. In concrete terms, the project shows the importance of the temporal component, which is more complex than the current data model. The change of perspective from the observation of a single manuscript to the combined analysis of many manuscripts leads to a feedback on the data modelling and can sharpen the view on the contexts of manuscripts.

Conclusion

The enrichment of texts with context, for example via registers and comments, is a central aspect of scholarly editions and digital scholarly editions. The data itself is based on accurate research at the sources, which requires a great deal of background knowledge. Exporting the objects and relations of a digital scholarly edition via an interface like *ediarum.WEB* opens the possibility to make the linked data accessible for new analyses. In the graph one can see all information simultanously and special queries show relations across long sequences of nodes. Of course, the graph-based analyses only use the information available in the data and in the best case the contexts shown in the sources is reproduced. By presenting these contexts in a common and linked way, graph-based approaches provide access to a better understanding of the historical sources. By providing clues as to where the data sets are incomplete or ambivalent, they contribute to a more complete critique of data and thus help to improve scientific interpretation. Thus, they help to evaluate, contextualize and classify the individual historical sources and the corpus as a whole.

Bibliography

- Fechner2017 = Martin Fechner, "ediarum: eine digitale Arbeitsumgebung für Editionsvorhaben" (Presentation at the workshop of the union of the German

Academies of Sciences and Humanities, 20.10.2017 in Mainz) 2017, URL: https://nbn-resolving.org/urn:nbn:de:kobv:b4-opus4-28169.

- Fechner2018 = Martin Fechner, Eine nachhaltige Präsentationsschicht für digitale Editionen, in: DHd 2018. Kritik der digitalen Vernunft. Konferenzabstracts, Universität zu Köln 26. Februar bis 2. März 2018. URL: http://dhd2018.uni-koeln.de/wp-content/uploads/boa-DHd2018-web-ISBN.pdf
- Fechner2018a = Martin Fechner, A Standardized Interface for Digital Scholarly Editions, 2018. URL: https://nbn-resolving.org/urn:nbn:de:kobv:b4-opus4-33277
- Kuczera2017 = Andreas Kuczera, Graphentechnologien in den Digitalen Geisteswissenschaften, in: ABI Technik 2017; 37(3): 179–196 (https://doi.org/10.1515/abitech-2017-0042). URL: https://www.degruyter.com/downloadpdf/j/abitech.2017.37.issue-3/abitech-2017-0042/abitech-2017-0042.pdf
- Kuczera2018 = Andreas Kuczera, Graphentechnologien in den digitalen Geisteswissenschaften. Modellierung – Import – Exploration (Online publication at https://kuczera.github.io/Graphentechnologien/). (2018)