# Advancing RDM: From Immersion to Argumentation in Science

Ralf Möller

*University of Hamburg, Institute for Humanities-Centered AI (CHAI), Warburgstraße 28, 20354 Hamburg, Germany*

**Abstract**

In the age of digital science more and more information systems based on curated data for a certain research field become available. Scholars can use information systems to analyze data and draw conclusions, which are then written down in scientific publications. Arguments contained in the publications can, however, usually be tied to data with a lot of effort only. Visualizations for scholars, for instance, in particular 3D visualizations with immersion effects, can hardly be made available with current technology such that other researchers can indeed verify the arguments that a scholar derived from 3D display elements (and more conventional display elements) shown in information systems. Pictures and videos that are usually referred to are no longer enough for scientific argumentation in the digital age. This essay discusses the current state of the are in scientific information systems and provided new ideas for combining data display, citation, and verification of arguments using formal means.

As many people use support systems based on large language models for summarizing, rewriting, or even writing texts or generating pictures or videos, school activities as well as science activities have changed considerably, and many people are concerned that human abilities will change for the worse or will no longer develop. However, text work is not the only part of school-related or scholarly activity. Usually, first comes the development of new hypotheses about effects, phenomena, correlations, and causalities in a domain of investigation. Hypotheses, so the goal should be, are to be supported by arguments, that is to say, by a systematic analysis of already available or newly acquired data. Hypotheses might be combined into scientific theories. A hypothesis or even a theory, though well-supported with arguments initially, might be falsified later as new data become available. On the one hand, there exists a well-developed methodology of arguing with scientific hypothesis tests in a quantitative way, and on the other hand, for qualitative research, scientific arguments are based on propositions tied together with causal relations. In both contexts, arguments are written down in natural language together with citations to publications that have influenced the development of the arguments. Since falsification of hypotheses (or theories) drives science, access to data supporting arguments must be made as easy as possible, such that counter-arguments can systematically connected to initial arguments and respective data. This way, hypotheses are explicitly modeled, such that negative effects of generative AI technology (e.g., generating erroneous statements) are at least mitigated.

In this paper we argue that for this purpose citation of data as known today might not be enough. Data must be turned into explicit scientific arguments used to strengthen or weaken scientific hypotheses. We sketch how this can be achieved even if researchers do not have an information technology background. Starting with citation in the digital age, we then extrapolate existing technology to be able to explicitly model scientific argumentation. Example scenarios are taken from research on written artifacts as part of the Cluster of Excellence "Understanding Written Artefacts."[1]

## Scientific Citation

Usually, citations involved in supporting or refuting arguments refer to papers, articles, or even books. Sometimes, also data, e.g., stored in research data management systems, are cited, possibly together

*Large Language Models for Research Data Management!? 2025 (LLMs4RDM 2025), Workshop at the INFORMATIK Festival 2025 (55th Annual Conference of the German Informatics Society), September 18, 2025, Potsdam, Germany*
ralf.moeller@uni-hamburg.de (R. Möller)
https://www.philosophie.uni-hamburg.de/chai/personen/moeller.html (R. Möller)
0000-0002-3228-316X (R. Möller)


[1] https://www.csmc.uni-hamburg.de/research/cluster-projects.html

with respective data processing programs, such that the tenability of hypotheses under consideration is supported or weakened depending on the circumstances. In other cases, information systems are used to display hypothesis-related data (see Figure 1 for an example information system called Epigraphic Database of Ancient Asia Minor (EDAK)[2]).

In a text which, e.g., uses the EDAK dataset for argumentation support, citing a huge dataset with a digital object identifier (DOI) might often not be appropriate. Even if the concrete data item the argumentation referred to was additionally specified in the citation, downloading the data, getting the information system to display exactly this data item is usually much too much work. Furthermore, if the web-based information system was driven by a database, then the respective data might even be changed at all times, and thus, citations might be dangling references as the citation refers to a no-longer-existing version of the data, which is never displayed again in the version that gave rise to the arguments behind the citation. This problem is solved by information systems that obtain data directly from a research data archive because citations refer to persistent data [1, 2]. With recent developments even single data items shown in an information system can be easily cited. In Figure 2, a detailed view of a persistent data item for an inscription from EDAK is shown [3].

[2]https://www.csmc.uni-hamburg.de/edak.html

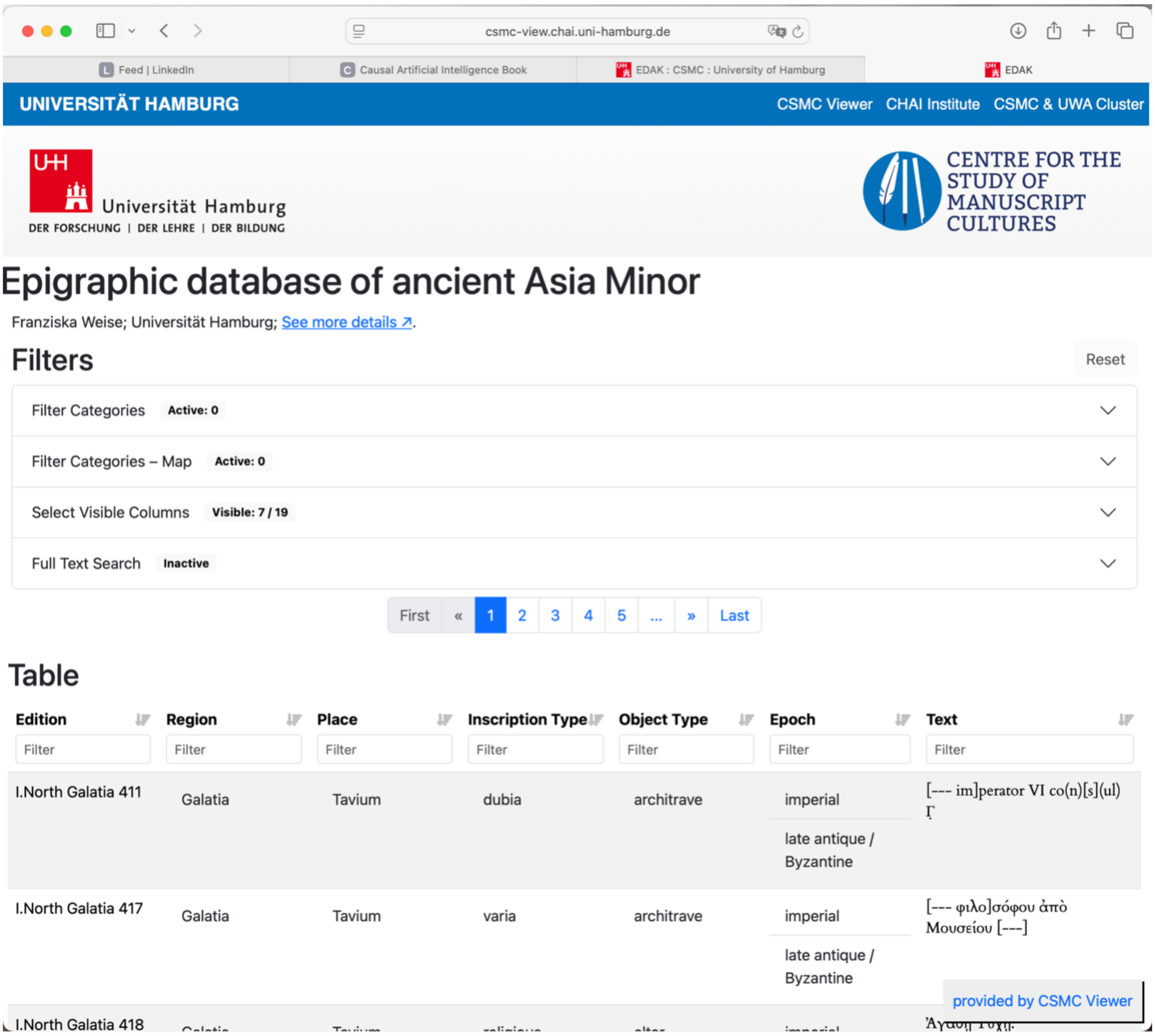



**Figure 1:** Example for an information system (EDAK) on epigraphic data (inscriptions) of ancient Asia minor. The information system provides filtering capabilities. In total 6500 data items for describing inscriptions are shown.

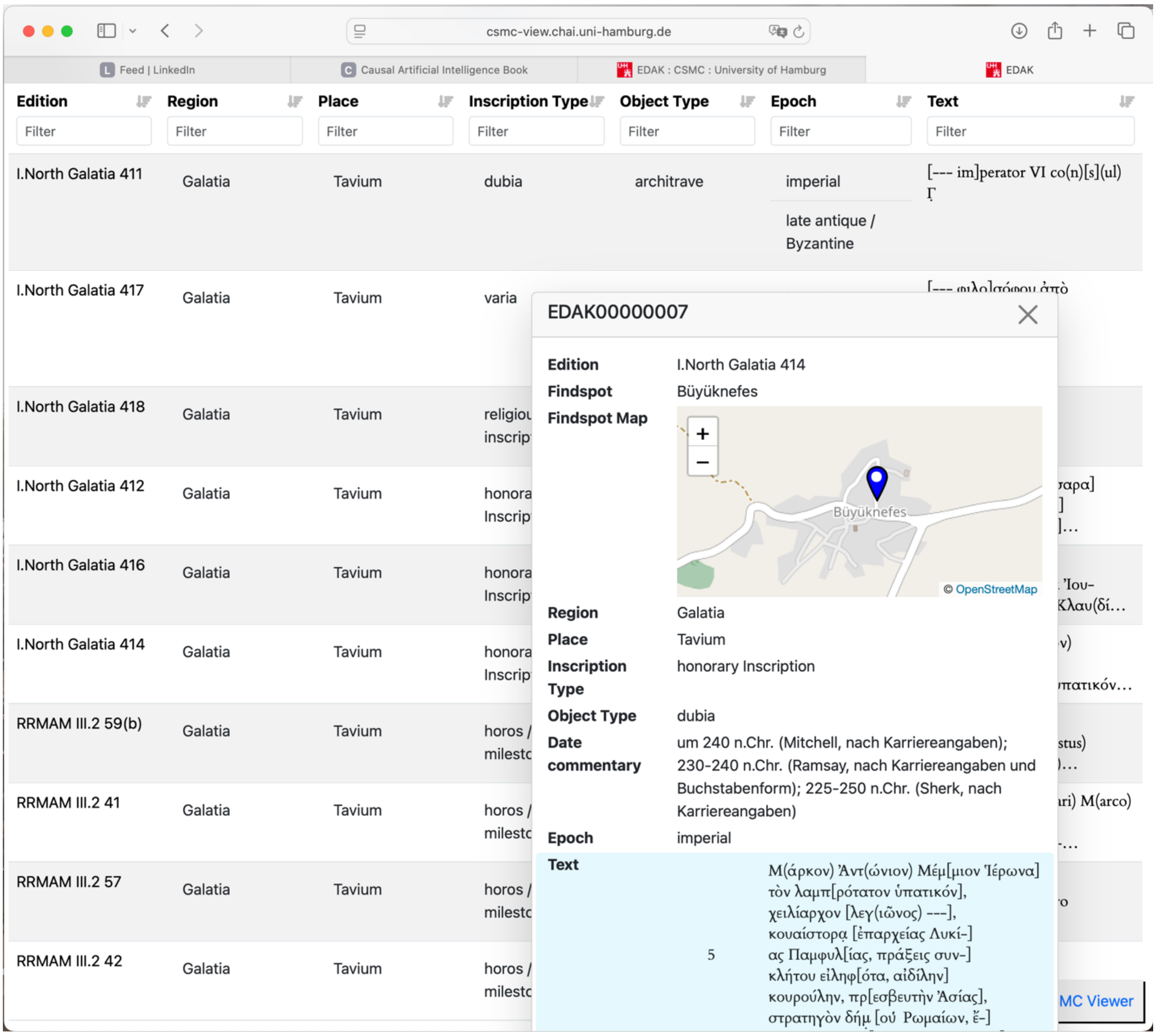


**Figure 2:** Detailed view of a certain data item.

From the detail view in Figure 2, besides a verbal citation also a DOI-based citation link can be obtained such that the detailed display to which the citation refers can be redisplayed with a single mouse-click via the underlying research data repository from which the interface is generated. In Figure 3 it is shown how a citation link can be obtained. The link can be copied into a new scientific document, like the one you are reading. The example link used here is: https://staging-rdm.fdr.uni-hamburg.de/records/0aevs-xp230#7. Now let us assume that, with the filtering facilities shown in Figure 4, some number of datasets with certain properties are selected for display.

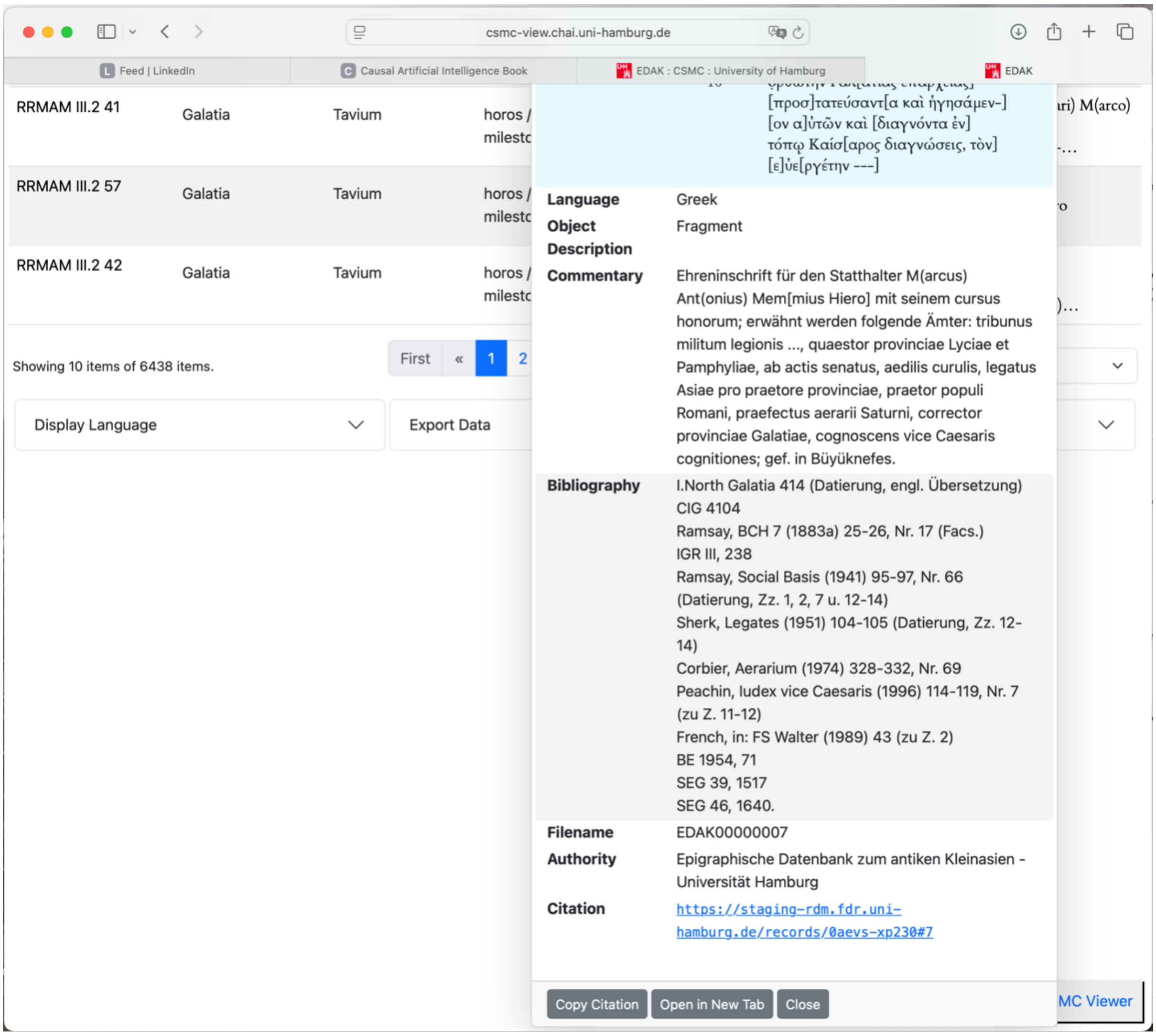



**Figure 3:** Getting a citation link for a single data item.

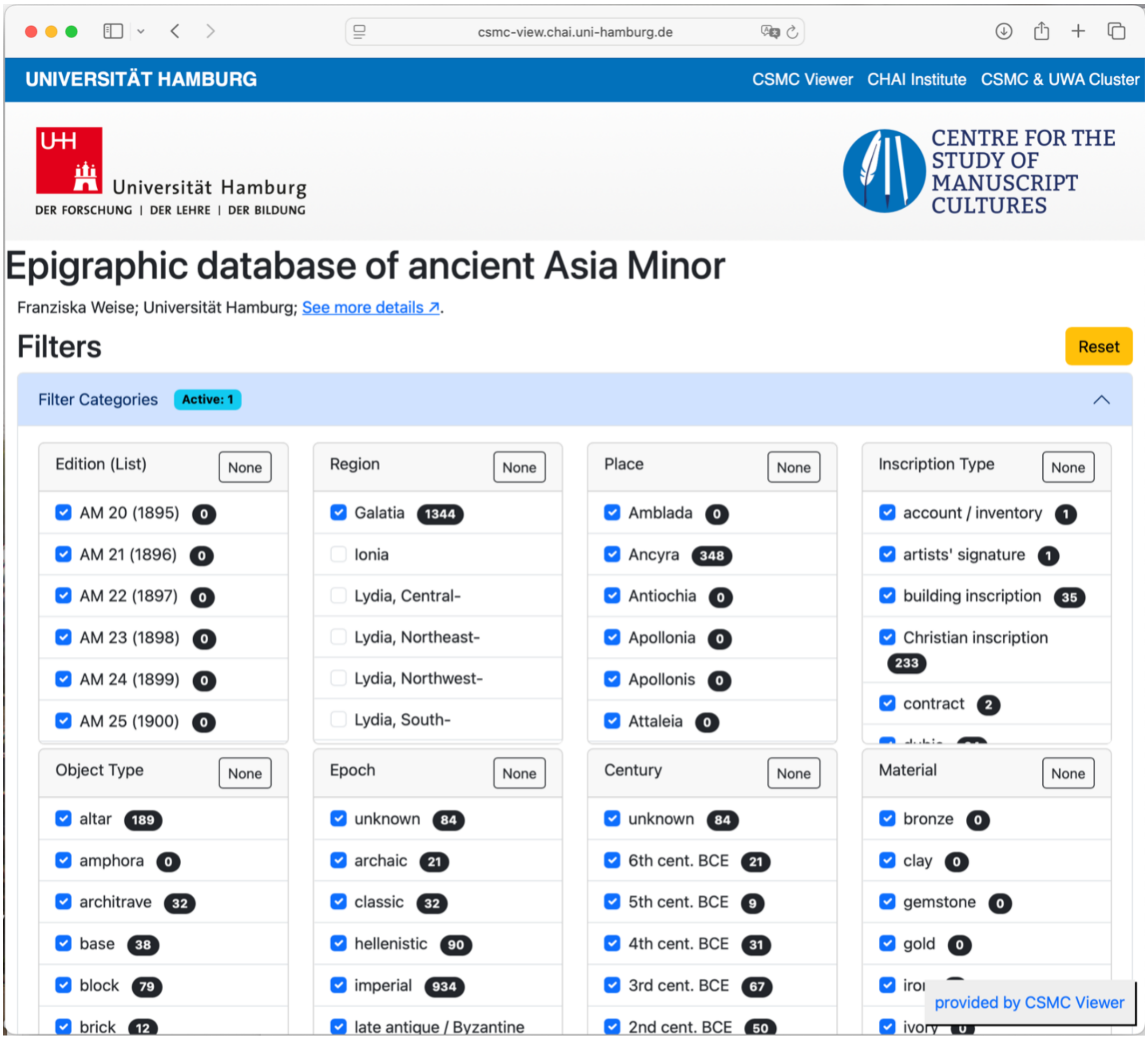
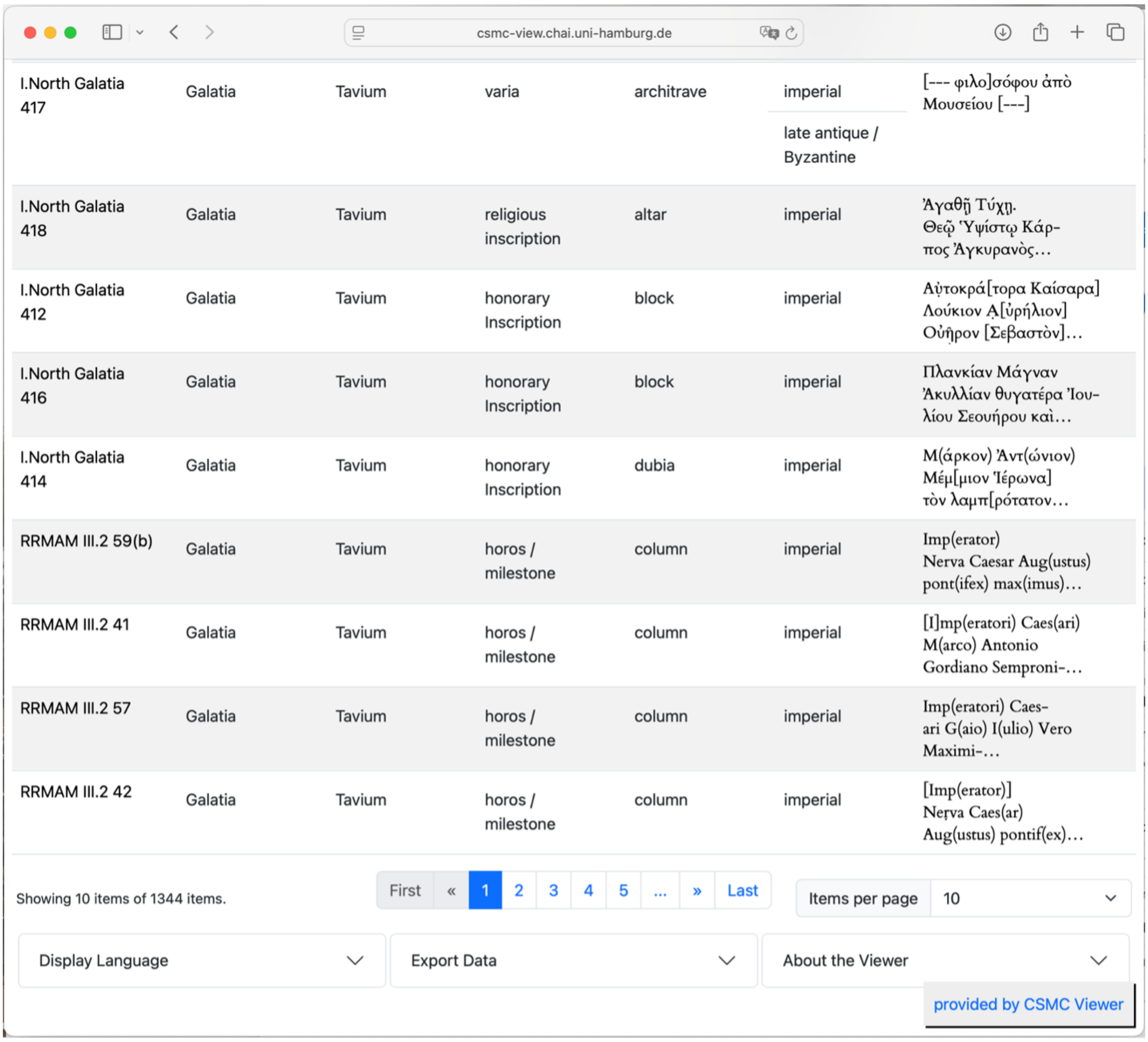



**Figure 4:** Filtering allows for selecting data items relevant for a certain hypothesis.



csmc-view.chai.uni-hamburg.de

| | | | | | | |
|---|---|---|---|---|---|---|
| I.North Galatia 417 | Galatia | Tavium | varia | architrave | imperial<br>late antique / Byzantine | [--- φιλο]σόφου ἀπὸ Μουσείου [---] |
| I.North Galatia 418 | Galatia | Tavium | religious inscription | altar | imperial | Ἀγαθῇ Τύχῃ. Θεῷ Ὑψίστῳ Κάρ-πος Ἀγκυρανὸς… |
| I.North Galatia 412 | Galatia | Tavium | honorary Inscription | block | imperial | Αὐτοκρά[τορα Καίσαρα] Λούκιον Α[ὐρήλιον] Οὐῆρον [Σεβαστὸν]… |
| I.North Galatia 416 | Galatia | Tavium | honorary Inscription | block | imperial | Πλανκίαν Μάγναν Ἀκυλλίαν θυγατέρα Ἰου-λίου Σεουήρου καὶ… |
| I.North Galatia 414 | Galatia | Tavium | honorary Inscription | dubia | imperial | Μ(άρκον) Ἀντ(ώνιον) Μέμ[μιον Ἱέρωνα] τὸν λαμπ[ρότατον… |
| RRMAM III.2 59(b) | Galatia | Tavium | horos / milestone | column | imperial | Imp(erator) Nerva Caesar Aug(ustus) pont(ifex) max(imus)… |
| RRMAM III.2 41 | Galatia | Tavium | horos / milestone | column | imperial | [I]mp(eratori) Caes(ari) M(arco) Antonio Gordiano Semproni-… |
| RRMAM III.2 57 | Galatia | Tavium | horos / milestone | column | imperial | Imp(eratori) Caes-ari G(aio) I(ulio) Vero Maximi-… |
| RRMAM III.2 42 | Galatia | Tavium | horos / milestone | column | imperial | [Imp(erator)] Nerva Caes(ar) Aug(ustus) pontif(ex)… |

Showing 10 items of 1344 items. First « 1 2 3 4 5 … » Last Items per page 10

Display Language | Export Data | About the Viewer

provided by CSMC Viewer

**Figure 5:** Selected data items (filtering results) can be exported (see the Export Data button at the bottom in the middle).

Let us further assume that scientific arguments written down in a scientific document are based on exactly the selected items (see Figure 5 in which 1344 data items are selected). Then, copying DOIs for 1344 single data items is hardly possible, not to speak about inserting them into a document as citation links. Technically it is possible to also offer a copy citation facility that takes into the filtering settings.[3] Alternatively, selected data items could be exported and inserted as a new contribution into a research data management system such that a citable DOI would indeed be available. Then, a reader of the scientific publication in question could review the hypothesis with references to the relevant underlying data. Despite unnecessary data duplication, it is important to understand that just showing respective data items to the reader (possibly on demand) only implicitly reveals how the items contribute to the scientific argumentation behind a scientific hypothesis. In other words, aggregated data systematically supports arguments for scientific hypotheses, and exactly those arguments must be made explicit. We conjecture that developing reliable AI tools to summarize, rewrite, or deal with scientific texts is made easier if argumentation structures are represented explicitly (and automatically verified). Whereas for data items shown with tables as discussed above, at least the citation problem has been solved [4], there remains the problem of systematically and explicitly referencing (sets of) data items in scientific argumentation structures. Before we tackle this problem, we broaden our view on data and also consider information systems for 3D data.

[3]The software system behind the display in Figure 5 will be extended accordingly.

Visually displaying information has to follow different rules. As an example, we also consider inscription data, but this time data about inscriptions that are found, e.g., on the seats in a theater at Miletus.[4] The theater was scanned, and the 3D data is shown in Figure 6. Note that in the example we discuss, data from 3D scans (dark gray) are augmented with fictitious data (shown in light gray).

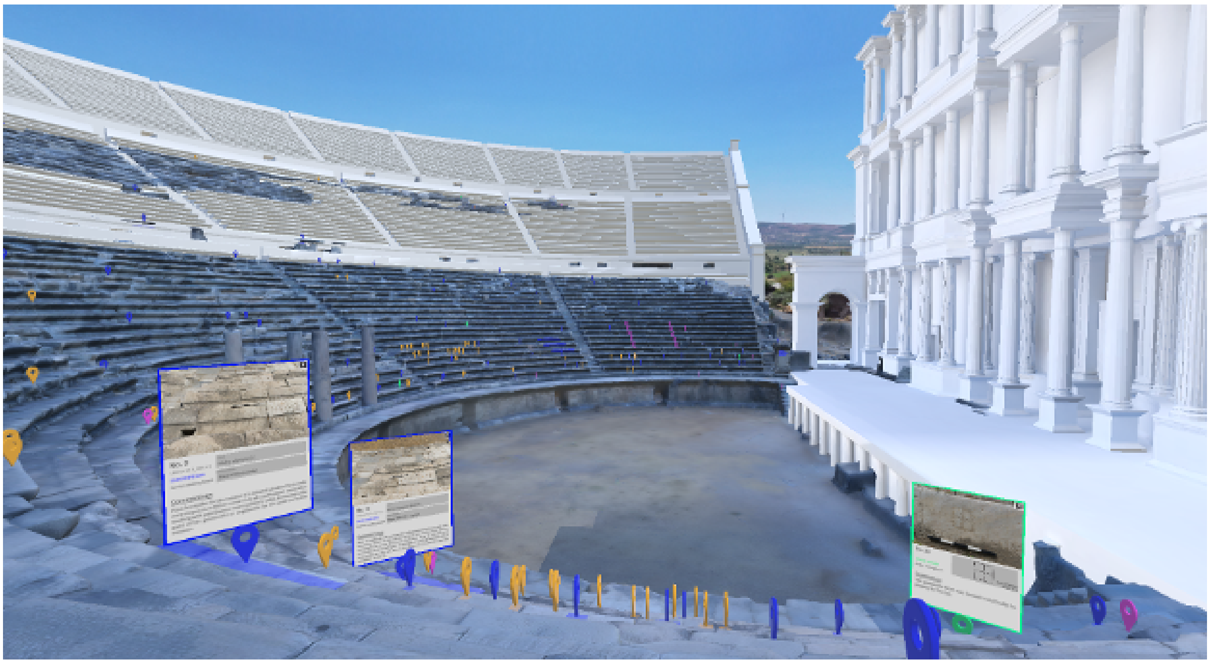


**Figure 6:** Miletus theater (Roman empire) (see https://www.csmc.uni-hamburg.de/research/cluster-projects/completed-cluster-projects/rfb02.html).

Rather than just providing pictures or videos of the scene[5], the display of 3D data could very well be integrated into a fully-fledged research data information system, such that citation of single elements of the display (e.g., inscriptions indicated by small blue and yellow markers) as shown in Figure 6 would become possible, which already would be a huge benefit. This way, scientific conclusion could really be grounded on data made available in a research data management system because hypotheses derived from seeing the seen on the writer side (and written down in an article, say) can be more easily verified by the reader of the article.

Given 3D data, possibly augmented with relational and textual data, researchers can even be embedded into the scene using cage technology or virtual reality technology. This way, some kind of immersion of scholars into a scene becomes possible to better stimulate new scientific insights. However, how can we make sure that arguments derived, with hypotheses about fictitious or real data, are supported by evidence for other researchers? In Figure 7, scholars are immersed into the Miletus theater scene, with a building that is tentatively included into the amphitheater scene. Some conclusions might be drawn from the immersion experiment (data display in a cave). The conclusions might be written down in a paper. However, as stated above, how can we make sure that the conclusions can be verified by a third-party researcher afterwards? Otherwise, without evidence being provide to third party researchers, there are just claims, not well-supported hypotheses. Questioning claims is important for science. Speaking with Popper, without support for questioning hypotheses (and therefore conclusions), there is hardly any science at all. In order to support questioning of hypotheses at least a facility is required to make the scene effortlessly available to third-party researchers with a URL as has been shown with the information system EDAK discussed above that is based on persistent data a research data repository.

[4] https://www.theatrum.de/633.html

[5] https://www.csmc.uni-hamburg.de/research/cluster-projects/completed-cluster-projects/rfb02.html or https://l2gdownload.rrz.uni-hamburg.de/abo/00.000_video-59872_2022-01-12_08-52.mp4

**Figure 7:** Immersion of scholars into a 3D scene of an ancient theater. Ficticious data is shown in gray color (see https://www.csmc.uni-hamburg.de/artefact-lab/computer-science-lab/human-interaction.html)).

## From Scientific Citation to Explicit Argumentation

The fact that the tentative display of the building in Figure 7 is indeed sensible is supported by very concrete foundation structures shown in Figure 8.

**Figure 8:** Foundations as evidence for a building (see https://www.csmc.uni-hamburg.de/research/cluster-projects/completed-cluster-projects/rfb02.html).

The data underlying the displays in Figures 6-8 were manually created with 3D scanning, imaging, and CAD modeling techniques. Just providing the picture is not enough. The question is how can we make the argument explicit *for machine processing* that there is evidence for a building based on the foundation structures shown in Figure 8? Just maneuvering the reader to a respective display at reading time of the paper is not enough. The argumentation must be made available in an explicit way to machines questioning the conclusions drawn and published from hypothesized objects.

Even additional techniques concerning so-called generative AI technology come into play these days. With current AI technology even fictitious graphics can be generated easily based on textual descriptions to augment visualizations generated from scanned data with some sensible context. In Figure 9 a reconstruction of a 17th century fortress is shown, and – based on 3D data of the fortress – a visualization is generated to show the surroundings. Today, with the advent of generative AI technology, such visualizations can be quite easily generated by scholars based on textual descriptions about assumptions of the ancient reality. Some arguments are directly supported by data, whereas other graphical elements are fictitious to emphasize interpretation of available data. For some assumptions, there are supporting data, other assumptions are deductively, inductively, or abductively derived from underlying data. Not in all cases, fictitious elements are indicated in gray as done in Figures 6-8, and thus, there is a lurking danger of over-interpreting augmented visual displays. Understanding written artifacts, for instance, means suggesting conclusions as tenable hypotheses about human objectives or activities with and around ancient artifacts. If conclusions about human objectives or activities were derived from generated visualizations, exactly those hypotheses and assumptions had to be made explicit in order to guarantee scientific research in accordance with good scientific practice. The same holds for the immersive techniques exemplified in Figures 6-8.

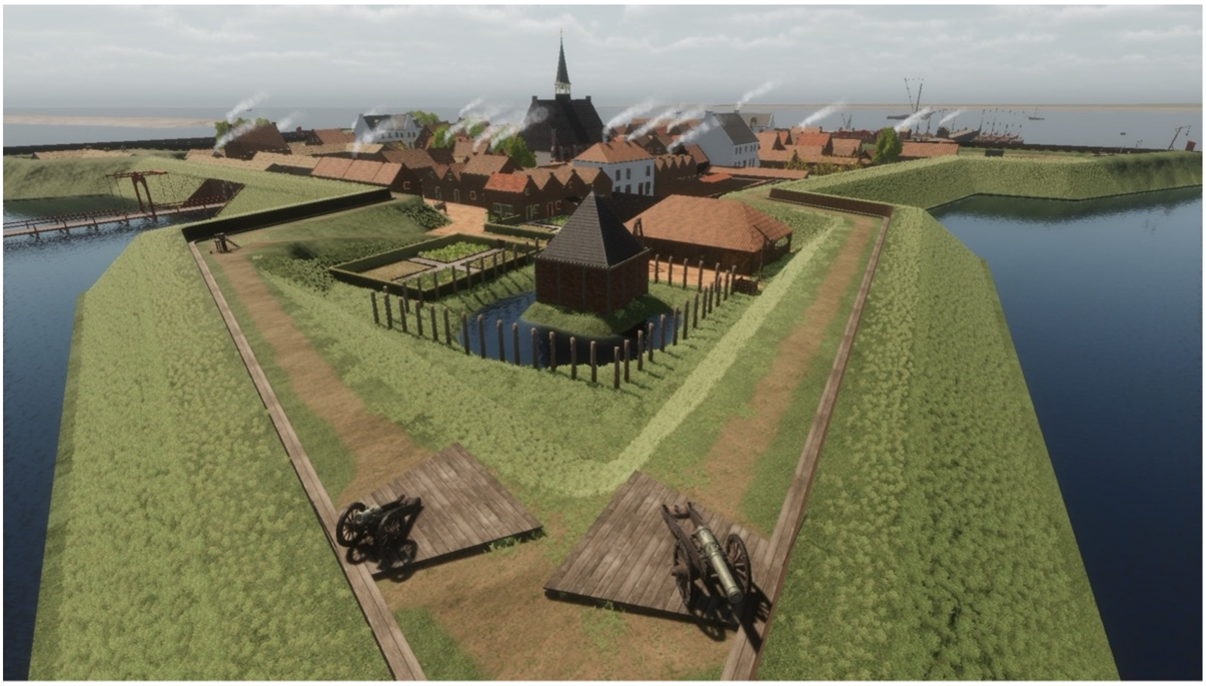


**Figure 9:** 17th century fortress of Lillo, Belgium (see https://www.timemachine.eu/events/lillo-1640-methodology-and-workflow-of-virtual-reconstruction/).

It becomes clear that human objectives and motivations in ancient times are only indirectly accessible based on evidence for human behavior concerning ancient objects involved, i.e., production of written artifacts, buildings (and their foundations). More concretely, hypotheses about human objectives in ancient times must be supported with models about human behavior, which, in turn, is explained by evidence derived from objects (artifacts) and their relations available today. In this context, there is no absolute truth, i.e., scientific argumentation in this context is inherently tied to uncertainty or plausibility. Thus, for making scientific arguments explicit, we conjecture that probabilistic models are beneficial, in particular those supporting clausal and dynamic (temporal or longitudinal) influences. To realize the sketched vision of scientific argumentation (linking data to argumentation), research data management must be extended such that data stored in a research data repository can be systematically used in *model-based scientific argumentation*. Model-based scientific argumentation must be made available to non-data-literate scholars (not to speak of AI literacy). As we have argued, model-based scientific argumentation must be based on formalisms for systematically handling causality and uncertainty with machines.

## A Formalism for Scientific Argumentation in the Humanities

To illustrate the main ideas of formalizing argumentation in the humanities, we consider an example from the field of understanding written artifacts. For a book manuscript, say, certain data describing book-binding information might become available, and the dataset in a research data repository is considered to be reliable. Data might suggest that the leather binding is worn. It is highly likely that this is causes by frequent use by humans in a certain context, Due to the text transcription there is a lot of evidence that the book was used in specific services. With further information (possibly from data found in other research repositories) a relation to other artifacts, a christening candle with holder, say, with the same finding location is very likely. These arguments might be written down in a publication, together with pointer to the respective datasets. As we have argued, the argumentation structure might be there, but it is all but easy for computer systems (e.g., based on large language models) to capture those argument structures in a stable and computationally effective way. Arguments are given as natural language text, albeit with pointers to (single) data items. With generative AI techniques even a graphical visualization of the manuscript as it is used in a service might be generated as illustrations to appeal to the reader's imagination (see the theater example with the building as discussed above).

Appropriate ontological and epistemological commitments must be made to appropriately deal with argumentation structures in this context:

1. Objects and relations (first-order structures are required) and uncertainty about object existences, their attribute values, and their relations,
2. Temporal behavior of objects and uncertain influences over time,
3. Causality and respective uncertainty.

One candidate formalism that can handle the representation requirements for argumentation are parametric factor graphs [5, 6] for which so-called lifted inference algorithms provide a first step to scalable probabilistic relational representation and reasoning. Since the advent of [5], several dissertations contribute to probabilistic relational reasoning now even with causal and dynamic information (see the following citations for newer references to conference and journal papers).

In order to be able to support probabilistic argumentation the basis was laid by Tanya Braun [7] with her work on lifting for efficient query answering algorithms concerning multiple queries based on the transfer of the idea of so-called junction trees to the context of relational probabilistic models (or parametric factor graphs, PFGs, as the models are usually called). Lifting supports efficient query answering w.r.t. models based groups represented by variables (parameters). The number of placeholder is determined as evidence becomes available. Evidence destroys lifting possibilities in principle as information about specific objects in a group required to treat objects indivudally rather than with placeholders (shattering of models). In order to be able to handle temporal aspects of reasoning in case of evidence collected over (discrete) time and hence the lifted structure of models being shattered, Marcel Gehrke's research [8] provided results on systematically restoring groups after collecting evidence over time such that lifting becomes effective again. Error bounds on accuracy of query answering could be provided. Nils Finke avoided shattering while evidence is collected in certain cases over time. He also investigated different ways to regroup objects in models if shattered by evidence [9]. For prediction, an important topic in reasoning about temporal behavior, the work of Marisa Mohr becomes relevant, in particular because she dealt with periodic events [10]. For continuous time, lifting was investigated in the dissertation of Mattis Hartwig [11]. Furthermore, as argumentation structures refer to parts of text and images or videos, Felix Kuhr and Magnus Bender have investigated so-called subjective context descriptions, which provide the basis for anchoring arguments in textual and other media [12, 13] using language modeling and embedding technology. In the dissertation of Sylvia Melzer, the combination of embedding and logical approaches was first investigated [14]. Malte Luttermann contributed to incorporating cause knowledge into parametric factor graphs and also investigated learning algorithms for these models [15]. Automatic model acquisition for the resulting causal dynamic parametric factor graphs (CD-PFGs) from data is quite well understood, whereas CD-PFG model acquisition from natural language text is only partially investigated (but see [16]).

How formal probabilistic models are used for scientific argumentation is investigated , for instance, in [17, 18, 19, 20]. Further references can be found in the literature. Causal structures and argumentation are discussed in [21, 22]. Again, further references can be found in the literature. Given the research results already known, we are now in the position to tackle the following research goals for advanced research data management:

1. How to link data to scientific arguments using causal dynamic parametric factor graphs?
2. How can formalism be hidden from non-IT-literate scholars such that the idea of linking data to arguments actually works?
3. How can we provide scalable inference to determine whether certain argumentation structures might apply in a new investigation context (e.g., to automatically close some apparent gaps in argumentation).

## Conclusion

In this essay, we have argued that time is ripe for moving from linking data to data toward linking data to arguments. While data provision in a research data repository, together with interfaces to understand data available right from a research data repository, is now a solved problem, linking data to argumentation is considered to be an extremely interesting field of research. Constructing explicit structures of argumentation is seen as the new task for scholars. The new view also provides a way to escape the fear that the use of language models in schools or science areas change the way people think to the worse. On the contrary, scientific thinking is leveraged by linking data arguments via LLMs and also multi-modal models as we have discussed with the Miletus example. We have argued that just providing immersion for a scholar generating ideas is not enough. Generated scientific insights and hypotheses in publications must be tied directly to the appropriate 3D displays so that external researchers can effectively verify scientific hypotheses in this context.


## Acknowledgments

In this essay I present joint work with my CHAI team at University of Hamburg: Thomas Asselborn, Marcel Gehrke, Malte Luttermann, Florian Marwitz, Sylvia Melzer, and Simon Schiff. In addition, my former team member Magnus Bender, now at Aarhus University, contributed a lot. I would like to also thank Kaja Harter and Franziska Weise for their collaboration concerning the activities behind developing the EDAK information system.

## Funding Information

This contribution was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC 2176 "Understanding Written Artefacts: Material, Interaction and Transmission in Manuscript Cultures," project no. 390893796. The research was mainly conducted within the scope of the Centre for the Study of Manuscript Cultures (CSMC) at University of Hamburg.

## Declaration on Generative AI

No generative AI was used to produce even parts of the text of this essay.

## References


[1] S. Melzer, H. Peukert, S. Thiemann, E. Radisch, Humanities in the Center of Data Usability: Data Visualization in Institutional Research Repositories, volume 3814, CEUR-WS.org, 2024, pp. 67–74.

[2] T. Asselborn, R. Möller, S. Melzer, Implementation of information systems for the long-term reuse of data in humanities research, volume 256, Elsevier, Netherlands, 2025, pp. 86–92. doi:10.1016/j.procs.2025.02.099.

[3] K. Harter-Uibopuu, F. Weise, S. Melzer, R. Möller, T. Asselborn, M. Bender, F. Marwitz, K. Woerner, S. Thiemann, Epigraphic database for ancient asia minor, 2025.

[4] S. Schiff, Enabling Research Data Management for Non IT-Professionals, Ph.D. thesis, Computer Science Department, 2024.

[5] L. D. Raedt, K. Kersting, S. Natarajan, Statistical Relational Artificial Intelligence: Logic, Probability, and Computation, Morgan & Claypool Publishers, 2016.

[6] G. Van den Broeck, K. Kersting, S. Natarajan, D. Poole, An Introduction to Lifted Probabilistic Inference, MIT Press, 2021. URL: https://mitpress.mit.edu/books/introduction-lifted-probabilistic-inference.

[7] T. Braun, Rescued from a Sea of Queries: Exact Inference in Probabilistic Relational Models, Ph.D. thesis, Universität zu Lübeck, 2020.

[8] M. Gehrke, Taming Exact Inference in Temporal Probabilistic Relational Models, Ph.D. thesis, Universität zu Lübeck, 2021.

[9] N. Finke, Navigate Through Troubled Waters: Adapt to Environmental Changes in Dynamic Probabilistic Relational Models, Ph.D. thesis, Universität zu Lübeck, 2023.

[10] M. Mohr, Multivariate Ordinal Pattern Representations for Time Series: Learning from Ups and Downs:, Ph.D. thesis, Universität zu Lübeck, 2022.

[11] M. Hartwig, New Methods for Efficient Query Answering in Gaussian Probabilistic Graphical Models, Ph.D. thesis, Universität zu Lübeck, 2022.

[12] F. Kuhr, Context-aware Corpus Annotation using Subjective Content Descriptions: Context is the Key, Ph.D. thesis, Universität zu Lübeck, 2022.

[13] M. Bender, Toward better Subjective Content Descriptions: A Spiral of Estimation, Enrichment, Usage, and Improvement, Ph.D. thesis, Universität zu Lübeck, 2024.

[14] S. Melzer, Semantic Assets: Latent Structures for Knowledge Management, Ph.D. thesis, Universität zu Lübeck, 2018.

[15] M. Luttermannn, Causal Dynamic Parametric Factor Graphs: Reasoning and Machine Learning, Ph.D. thesis, Universität zu Lübeck, 2026. Forthcoming.

[16] A. Nafar, K. B. Venable, Z. Cui, P. Kordjamshidi, Extracting probabilistic knowledge from large language models for bayesian network parameterization, 2025. URL: https://arxiv.org/abs/2505.15918. arXiv:2505.15918.

[17] S. T. Timmer, J.-J. C. Meyer, H. Prakken, S. Renooij, B. Verheij, A two-phase method for extracting explanatory arguments from bayesian networks, International Journal of Approximate Reasoning 80 (2017) 475–494. URL: https://www.sciencedirect.com/science/article/pii/S0888613X16301402. doi:https://doi.org/10.1016/j.ijar.2016.09.002.

[18] P. Grim, F. Seidl, C. McNamara, H. Rago, I. Astor, C. Diaso, P. Ryner, Scientific theories as bayesian nets: Structure and evidence sensitivity, 2021. URL: https://philsci-archive.pitt.edu/18705/, forthcoming in Philosophy of Science.

[19] Y. Cao, R. Fuchs, A. Keshmirian, Enhancing argument generation using bayesian networks, in: P. Cimiano, A. Frank, M. Kohlhase, B. Stein (Eds.), Robust Argumentation Machines, Springer Nature Switzerland, Cham, 2024, pp. 253–265.

[20] T. Kondo, K. Washio, K. Hayashi, Y. Miyao, Bayesian argumentation-scheme networks: A probabilistic model of argument validity facilitated by argumentation schemes, in: K. Al-Khatib, Y. Hou, M. Stede (Eds.), Proceedings of the 8th Workshop on Argument Mining, Association for Computational Linguistics, Punta Cana, Dominican Republic, 2021, pp. 112–124. URL: https://aclanthology.org/2021.argmining-1.11/. doi:10.18653/v1/2021.argmining-1.11.
[21] P. Besnard, M. Cordier, Y. Moinard, Arguments using ontological and causal knowledge, CoRR abs/1401.4144 (2014). URL: http://arxiv.org/abs/1401.4144. arXiv:1401.4144.
[22] L. Bengel, L. Blümel, T. Rienstra, M. Thimm, Argumentation-based probabilistic causal reasoning, in: P. Cimiano, A. Frank, M. Kohlhase, B. Stein (Eds.), Robust Argumentation Machines, Springer Nature Switzerland, Cham, 2024, pp. 221–236.