Giulia Giannini *

TOWARDS A COMPUTATIONAL HISTORY OF SCIENCE: LIMITATIONS AND PERSPECTIVES OF AN EMERGING RESEARCH APPROACH **

Abstract – Data science has recently emerged as a multi-disciplinary field of research where statistics, data analysis, machine learning and their related techniques are combined in a systematic way to support understanding of actual phenomena concerned with data. The growing power of storage infrastructures and the consequent availability of large amount of data opened up unprecedented opportunities to support the specification of ad-hoc data-driven approaches and tools for a number of application fields, such as biology, medicine, economy, politics, and history. In historical studies of science and knowledge, the use of data-science solutions is gaining more and more attention and the scientific debate is more topical than ever.

Collecting and Preserving the Past: Digital History and Digital History of Sciences

Since at least the beginning of the millennium, a new wave of studies has gathered under the umbrella of 'Digital History.' This new arena of scholarship has involved various academics experimenting in the field of history with new tools made available by information technology and computer science. Efforts focused mainly on using new media and technology to collect, preserve and present the past.

The most tangible effect of these almost two decades of research is undoubtedly the immense amount of digitization, virtual libraries, and ar-

* Dipartimento di Studi Storici, Università degli Studi di Milano – Via Festa del Perdono 7, 20122 Milan, Italy, e-mail: Giulia.Giannini1@unimi.it

** This work is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (TACITROOTS, PI: Giulia Giannini, Grant agreement No. 818098).

chives accessible from one's own computer. Over the past 15-20 years, the historical community has benefited enormously from the digitization of manuscripts and printed texts undertaken by major repositories around the world.

A number of international initiatives have been launched with the explicit aim of collecting digital content in mass quantities for access via the internet. From Google Books to HathiTrust, from Internet Archive to Jstor (just to name the best known and most widely used), these actions have impressively increased the amount of digitized content now available online. Significant and large archives have been made accessible by computer also thanks to substantial support from major European and international research funding agencies. Billions of euros in grants have enabled the creation of virtual archives, digital corpora of manuscripts, online editions of texts and transcriptions.

The electronic environment provides the historian with access to texts and documents that are geographically distant or hardly accessible due to conservation and preservation issues; it allows for repeated consultation and not infrequently offers documents that can be sorted out and searched, creating a mode of reading and analysis that is radically different from other forms of reproduction and consultation. This 'digital revolution' has definitely influenced, and continues to influence, the work of the historian (not only of science) – who finds enormous volumes of material readily and differently available – but it has been, above all, a revolution in the field of Library and archival science which can no longer fail to confront digital technologies.

In recent years, however, the interaction between history (of science) and data science has entered a more mature phase. The dramatic growth of available data and documents, the entirely new format in which they are presented to the historian, and the rapid and recent developments in data science, machine learning, and artificial intelligence have posed new problems, along with a sense of new possibilities.

Historians have begun to approach the massive amount of data at their disposal with more awareness, to question the potentials derived from their format and from the intersection/intertwining with the data science community, as well as to experiment with new approaches to analyzing the historical record. The need to keep humans 'in-the-loop' led to crowdsourcing solutions, to help in large-scale, human-intensive processes such as text tagging, commenting, rating, and reviewing, along with the creation and upload of content in a methodical, task-based fashion. Furthermore, the rapid development and diffusion of artificial intelligence techniques and Data Science approaches enable research in the field of humanities to become more and more computational.

Even in the field of history of science the focus is progressively shifting from large-scale digitization projects and tools for producing, curating, and exploiting historical data, to the attempt to rely on data science techniques and new hybrid methodologies for generating new historical knowledge.

The design and exploitation of digital repositories has been prompted in the last few decades by the need of scholarly communities in the humanities (including history) for better and more effective access to scholarly material and is commonly known as *Digital Humanities* (or Digital History). The use of coding, machine learning techniques, and cutting-edge technologies for processing and analyzing these repositories implies a novel and different interaction between the humanities and computer science. The computational aspect acquires a new and decisive role for the research itself. For this reason, this new field of disciplinary interaction is now normally referred to as *Computational Humanities* (or Computational History).

The 'Big Data revolution' and the 'Computational turn:' towards a Computational History of Science

As we move from books, folders, and manuscripts to databases, pdfs, and jpegs, the question is what to do, as historians, with all this data. How can the historian extract, articulate, and structure information from pixels, numbers, and records? How can and should computational tools and technologies be adapted to apply to the historical record? And what kinds of new research questions, methodologies, and perspectives can they open up?

Over the past 5 years, the 'big data revolution' and the 'computational turn' within the history of science studies have been the subject of debate in some of the leading journals in the field.[1] The challenge of increasingly powerful technologies, the quantitative multiplication of data, and the qualitative changes necessarily associated with these processes open up the possibility of considering new research practices and technologies. A number of scholars and projects have begun to experiment with new methodological approaches to historical data, and conversations have arisen among scientific communities that were until now far apart.

On the one hand, we are faced with an extraordinary abundance of information. Large repositories of correspondences, books, images, and

---

[1] Aronova, Von Oertzen, Sepkosky (eds.), 2017; Gibson, Laublicher, Maienschein (eds.), 2019.

documents of various kinds are now available to the historian offering enormous opportunities for studies with long spatial and temporal perspectives. Big data implies not only huge masses of data in terms of volume, but also an exponential multiplication of interconnections between the data themselves. The risk is to fall into an 'information overload,' into 'too much to know.' [2] On the other hand, developments in data science are providing an increasing set of methods and technologies that can support the historian in extracting, structuring, analyzing and visualizing data. The explosive growth of available historical resources definitely allows historians to know them better, but – more importantly – to know and study them differently.

1. *Integrating scales and factors in the history of Knowledge*

Over the past forty years, the boundaries and contours of the history of science have been redrawn by a wave of social, cultural, and historical work, some of which fits within the parameters of 'connected' and global histories.[3] New attention has been brought to the practical aspects of science and its materiality, to the political, social, institutional, and economic dynamics of 'making science.' In these meticulous works, strictly anchored to the sources, spatially and temporally circumscribed worlds acquire size through the exploration of a 'third dimension:' the analysis of the interaction of multiple factors. New actors and new objects came into the focus of the historian of science, and a new dialogue has begun with disciplines such as sociology, anthropology, and economics.

On the other side, there is an increasing call for approaches based on a global and *longue-durée* perspective, capable of relating different contexts and formulating hypotheses on the evolution of knowledge by covering long temporal and spatial distances.[4] A global and/or *longue-durée* perspective can highlight continuities and disruptions, repeated patterns, and winning or losing strategies in knowledge dissemination mechanisms. The need to integrate and reconcile the so-called 'micro' (in a broad sense) and macro-history has recently animated the debate not only in the particular field of the history of science.[5] One of the potentials of the burgeoning computational history is precisely the ability to merge all these scales.

---

[2] Rosenberg, 2003; Blair, 2010; Burke, 2000.

[3] See, for instance, Romano, 2016.

[4] See, for instance, Renn, 2020.

[5] Ghobrial (ed.), 2019.

Due to the impossibility of a 'close reading'[6] of a large mass of documents, long-term and global historical reconstructions have so far been based on the analysis of selected case studies. Even the collaboration of multiple scholars in a research group cannot reasonably result in covering long temporal and spatial distances without making important selections. Computational techniques potentially offer the historian the capacity to analyze large datasets in their entirety, including materials that would have been set aside in an inescapable source selection process. The scholar can analyze a network (e.g. of correspondence) as a whole, measure its centrality, and represent it through maps and graphs. A wide historical dataset can provide insight into patterns of knowledge spread and evolution at large scales. Through computational methods, multiple factors (theoretical, social, political, institutional, economic, etc.) can be investigated simultaneously over long temporal and spatial distances.

2. *The Network Beyond Metaphor: From a Descriptive Tool for Narrative to an Analytical Tool for Inquiry*

The application of computational methods to historical research is still in its infancy and the first results are only now beginning to emerge. However, the analysis of historical networks is certainly one of the most explored and promising approaches in the field.

The network metaphor has long and widely been used to describe the complexity of relationships that characterize social and intellectual systems. Developments in Social Network Analysis (SNA) have shifted the focus from the network as a metaphor to the network as an analytical mathematical model able not only to describe, but also to analyze and explain structures and processes. In recent decades, SNA has profoundly integrated study methodologies across a wide variety of fields, from the sciences to the humanities and has become one of the core methods of the digital and computational humanities.

Network analysis focuses on the study of relationships between 'nodes.' These can be people, books, instruments, ideas, places, institutions, etc. The analysis aims to capture, describe and measure the behavior of the entire network, highlighting its density, the centrality of its nodes and the nature of the relationships between them.

A central perspective in the computational history of science is the attempt to apply the network lens in order to understand the process of

---

[6] As opposed to 'distant reading.' On the utility of a 'distant reading' approach in the history of knowledge, see, among others, Pomata, 2014.

developing knowledge networks, their changes over time, and the implications of the network's embeddedness on the behavior and strategies of scholars and institutions. The most immediate and intuitive application of network theory to the history of knowledge is the analysis of the relationships between individuals. The case of exchanges between scholars, typically through correspondence, is striking. Tons and tons of letters crowd the archives all over the world and a growing portion of them is now available in electronic format by various means. Digitized letters, transcriptions or substantial sets of metadata can be now consulted through digital repositories. Within this perspective, the mathematical model of the network is applied to a real-world network. A social network of correspondents is translated and conveyed within a mathematical model whose 'nodes' consist of individuals and 'edges' are marked by letters.[7]

The study of the past as a thick array of social relationships between individuals is certainly not a novelty, but the development of new digital technologies and social network studies has made historical networks a reality that is no longer merely metaphorical. Through the application of social network analysis to historical data corpora, several projects have begun to explore the emergence and development of knowledge networks.

Special attention has been paid to correspondence and in particular to epistolary exchanges within the so-called 'Republic of Letters.' This self-proclaimed community of scholars has attracted the attention of historians since at least the 1970s. The *Republica Literaria* was above all a virtual community connecting scholars across national borders through epistolary exchanges. Instruments, ideas, work-in-progress, printed books, political and personal information were shared by means of letters, and people joined the 'Republic' through the act of writing and receiving them.[8] Hundreds of thousands of letters have linked scholars from all over Europe for nearly three centuries in a veritable network of knowledge. Thus, it is not surprising that from the earliest studies on the Republic of Letters, the need to apply the methods of social analysis and to engage in research involving laboratories and electronic memories came forward.[9]

---

[7] See, for instance, Gingras, 2010, p. 330.

[8] Van Miert, 2014.

[9] As early as 1971, Hans Bots, in the list of theses (*Stellingen*) included at the end of his edition of Dupuy and Heinsius' correspondence stated: "The availability of an electronic memory, which also presupposes the presence of a research laboratory with sufficient financial resources, would facilitate the optimal use of seventeenth-century historical sources, such as correspondences, and would also give us a more accurate knowledge of intellectual life in the 18th century" (My translation), Bots, 1971, p. 236. On the Republic of Letters, the need for

Among the most well-known projects in the emerging field of computational history of knowledge are *The Circulation of Knowledge and Learned Practices in the 17th-century Dutch Republic* (CKCC)[10] based at the Huygens Institute, *Mapping the Republic of letters*[11] at Stanford University, *Cultures of Knowledge*[12] at Oxford University, *Six Degrees of Francis Bacon*[13] at the Carnegie Mellon University. These are projects established between 2008 and 2012, primarily developed and led by interdisciplinary consortia of universities, research institutes, and cultural heritage institutions, and supported by significant external funding. They all focus, with different tools and approaches, on the digital analysis/representation of connections among scholars within the Republic of Letters.

One of the main products of these projects has been the creation of web applications for collecting, browsing, and analyzing historical documents, primarily of epistolary nature. With the collaboration of the Bodleian Library, the *Cultures of Knowledge*'s Project created EMLO,[14] a web-based collaborative open source union catalogue of 16th-17th-18th Century correspondence. EMLO, which currently collects the metadata of approximately 180,000 letters connecting some 32,000 individuals and 7,000 locations, is continually expanding and among its key goals is the development of maps and network diagrams for charting individual's movements, the study of the network's evolution in space and time, and the intersections between networks. Similarly, from the CKCC project, arose *ePistolarium*,[15] a web application gathering around 20,000 letters coming into and going out of the Dutch Republic during the 17th Century. In addition to metadata, *ePisto-*

---

computational studies of the documents related to it, and their connection to earlier literature, see also Van Vugt, 2018.

[10] *Circulation of knowledge and learned practices in the 17th-century Dutch Republic. A web-based Humanities' collaboratory on correspondences* (CKCC) is a collaborative project led by a Dutch consortium of universities, research institutes and cultural heritage institutions and sponsored by the Netherlands Organization for Scientific Research (NWO) and CLARIN-NL: http://ckcc.huygens.knaw.nl/, date accessed: 4 April 2022.

[11] *Mapping the Republic of Letters* is a project launched in 2008 which received funding from the National Endowment for the Humanities and the Stanford Presidential Fund for Innovation in the Humanities: http://republicofletters.stanford.edu/, date accessed: 4 April 2022.

[12] *Cultures of Knowledge: Networking the Republic of Letters, 1550-1750* is a collaborative, interdisciplinary research project based at the University of Oxford, set up in 2009 and funded by the Andrew W. Mellon Foundation and the Arts and Humanities Research Council: http://www.culturesofknowledge.org/, date accessed: 4 April 2022.

[13] http://www.sixdegreesoffrancisbacon.com/, date accessed: 4 April 2022.

[14] http://emlo.bodleian.ox.ac.uk, date accessed: 6 April 2022.

[15] http://ckcc.huygens.knaw.nl/epistolarium/, date accessed: 6 April 2022. See also Ravenek, Van den Heuvel, Gerritsen, 2017.

larium also pays attention to the content of the letters themselves through the use of Natural Language Processing techniques. Moreover, it allows for geographical and time-based visualizations of networks, as well as full-text searches and co-citation queries.[16]

The results of *Mapping the Republic of letters* gave rise to *Palladio*.[17] *Palladio* is a general suite of tools for filtering and visualizing data spatially and temporally. Unlike other software (e.g. *Gephi*)[18] it does not allow mathematical analysis of the network, but it is a useful tool for qualitative analysis of data and patterns.

Within the computational history of knowledge studies and the application of network theory methodologies to historical data, visualization is gaining an increasing role. The use of maps and visualization techniques allows for the creation of a synthetic picture of large collections of data. Far from being simply an eye-catching tool for presenting results to the public, representations are taking on a more interesting exploratory function.[19] Visualization offers a view at a glance of a set of data and relations, potentially opening up new ideas and perspectives. Visualizing is therefore not only a way to convey research data effectively, but also to understand them differently.

3. *Beyond the Real-World Network: The Network as a New Approach to Research*

The application of network theory to social networks of individuals, i.e. the mathematical translation of a real network of correspondents in the past, can highlight the strategic importance of figures that have so far remained in the background. For instance, Evan Bourke studied the flow of information within the Hartlib circle, the intellectual correspondence network formed in London in 1641 around the figures of Samuel Hartlib, John Dury and Jan Amos Kaminski (Comenius). Through computational methods, Bourke shows the integral role in the network of some female members such as Dorothy Moore Dury and Katherine Jones.[20]

---

[16] Other digital resources for researching and analyzing historical documents through network theory methodologies have been created and are being developed in recent years (e.g. Electronic Enlightmnet: https://www.e-enlightenment.com/). For co-citation inquires see, especially, Gingras, 2010.

[17] https://hdlab.stanford.edu/palladio/, date accessed: 6 April 2022. See also: Conroy, 2021.

[18] https://gephi.org, date accessed: 6 April 2022.

[19] On the epistemic role of visualizations see Valleriani, Giannini, Giannetto (eds.), 2022.

[20] Bourke, 2017.

However, the possibilities of the field go far beyond the mathematical representation of actual networks of relationships between individuals. In recent years, a number of works have focused on the application of network theory to texts and commentaries. Fernandez Riva[21] applied the network model to German medieval manuscript transmission. The nodes of the network consist, in this case, not of people, but of texts. In particular, the author links together the texts included in the same manuscript. This makes it possible to investigate the structure of the connected components (e.g. whether the same group of texts tends to always be transmitted together) or the number of links of a single component with the others (the so-called 'node degree's'). This is a preliminary study that has among its ambitions to provide clues about the effect of being part of a miscellany of manuscripts for the survival of a text. Similarly, Evina Steinova[22] is analyzing the corpus of early medieval annotations to Isidore of Seville's *Etymologiae*, the most important medieval Latin encyclopedia. Through the application of network theory to the glosses, she aims to reveal connections between manuscripts and trace the transmission patterns of particular annotations.

The greatest potential of the application of network theory concerns the analysis of complex systems. In recent years attempts have been made to use multilayer networks in the field of history of science. In 2015, Charles van den Heuvel[23] – who worked in the CKCC and *ePistolarium* projects – emphasized the desirability of applying what he called a 'deep network' approach in historical studies in order to better capture the complexity of knowledge systems.

'Deep,' 'multimodal' or 'multilayer' networks incorporate and structure relationships of different types within the same system. Their potential lies in the possibility of describing the different interactions between elements and of showing and measuring the interconnectivity between different types of connections. It is not a simple comparison between independent networks: in multilayer networks the nodes of one layer may coincide fully or partially with those of another layer. A characteristic of this type of network is the possibility of studying the connections between the different layers and therefore the relations between the different types of relations that characterize the system.

Deep or multimodal networks have captured attention in many disciplines such as sociology, transportation studies, ecology, neuroscience, and

---

21 Fernández Riva, 2019.

22 Steinova, 2019.

23 Van den Heuvel, 2015.

others. In historical research, however, the application of complex multilayer networks remains largely *terra incognita* at this time.

One of the very few attempts to explore historical and scientific resources through this approach can be identified in the work of the project *The Sphere*, led by Matteo Valleriani at the Max Planck Institute for the history of science in Berlin. The Project aims to analyze the process of transformation of cosmological knowledge through the history of one of the most important texts used in universities between 1472 and 1650: Sacrobosco's *Tractatus de sphaera*. Within *The Sphere*, more than 350 editions of the *Tractatus* are analyzed through the creation of a multilayer network capable of tracing several factors at the same time. This approach makes it possible, for instance, to analyze the structure of different editions, the circulation of specific comments on the text, the role of authors and printers, or to investigate what factors played a role in the consolidation or dispersion of specific types of knowledge.[24]

Conclusive Remarks: Hurdles, Challenges and Opportunities of the Computational History of Science

In the face of the growing amount of data now available online, a number of issues and challenges arise when we want to go beyond a simple interrogation and use them to apply algorithms. The scholar normally approaches the work with a set of hypotheses and driving questions. To test them, he needs not only a specific dataset, but also specific metadata and relations, which are not always already available. For this reason, instead of taking advantage of existing data collections, more often than not researchers design and populate their own databases. Already from the construction and design of the database, close collaboration between profiles with substantially different skills and competences is essential. This is a profound transformation in the working practice of the historian, traditionally accustomed to reading, studying, thinking and writing in blissful solitude. In addition, in order to set up a research group, one needs funding or the luck of working in an institution where all the competences and a shared interest in the topic are in place.

Simply assembling expertise is also not enough. In a field still in its early stages, a common language and vision needs to be developed. The comput-

---

[24] Kräutli, Valleriani, 2018; Valleriani, Kräutli, Zamani et al., 2019; Valleriani, 2020; Zamani, Tejedor, Vogl et al., 2020; Valleriani, Kräutli, 2022.

er scientist cannot be reduced to a mere external developer of supporting software. The historian cannot limit his work to the final interpretation of the results released by the machine. A deep dialogue, the sharing of knowledge and contexts need to accompany the work since the choice of data and their organization. The difficulties of finding common ground from scratch and the still scarce presence of a solid base of prior work in the field are often the cause of project delays and are increasingly the focus of debate.[25]

Once a dialogue is established, the two communities of scholars still run into an enormous first difficulty: no matter how gigantic a dataset may seem to the historian, to the computer scientist it is not Big Data. The data corpora are too small to lead to relevant and reliable results.

By multiplying the amount of relations among the data, research using a multilayer network approach somewhat alleviates this difficulty. But one of the major challenges in the field is the attempt to standardize data and methodologies. Data and metadata are often collected and stored in different ways by individual research groups and thus cannot be compared or reused. The possibility of combining datasets from different research projects would greatly expand the potential of computational history of knowledge. Although there are too many data available for the analysis capacities of a single researcher, they are often still too few to be considered representative for effectively testing hypotheses through mathematical models. This is amplified by the fact that the data the historian works with can never be considered complete. Whether books, information, letters, instruments… objects from the past often come to us with gaps. Sometimes it is possible to trace their absence, other times their past existence is completely ignored.

Another major obstacle in exploiting the potential associated with the use of computational methods in the history of science is the dissimilarity and limited availability of textual material. Machine learning and text analysis techniques, which have proven very useful in other disciplines, when addressing the history of knowledge run up against limited and extremely heterogeneous text sets.

Automated transcription of manuscripts through AI-based text recognition is still very imperfect and not always applicable. Achieving an adequate set of textual sources from handwritten documents still requires monumental work on the part of the historian. Even in the case of dig-

---

[25] See, for instance, Edelstein, Findlen, Ceserani et al., 2017; Ravenek, Van den Heuvel, Gerritsen, 2017; Hotson, Wallnig (eds.), 2019.

itized printed texts, applying OCR techniques to the historical document can introduce 'noise.' Errors may concern spelling, grammar, can follow a word-segmentation error, or a spurious character-insertion. The presence of multiple errors of different type in a certain text segment can introduce so much noise that the overall digitization process becomes useless.

Even once a sufficient number of transcriptions or readable digital copies have been collected, one is more often than not confronted with documents in different languages, historical linguistic variants, major variations in abbreviations or spelling.

Resolving these obstacles jointly and not within individual projects is one of the big challenges going forward. The potential of computational history of science is not only about the possibility of a different study of sources, which combines sometimes constraining historiographic needs, but about the very way researchers, groups and institutions work and share research.

References

Aronova Elena, Von Oertzen Christine, Sepkoski David (eds.), 2017. "Data Histories," *Osiris*, 32.

Bastian Mathieu, Heymann Sebastien, Jacomy Mathieu, 2009. "Gephi: An Open Source Software for Exploring and Manipulating Networks," *Proceedings of the International AAAI Conference on Web and Social Media*, 3/1 (Third International AAAI Conference on Weblogs and Social Media), pp. 361-362.

Blair Ann M., 2010. Too Much to Know: managing scholarly information before the modern age, New Haven, Yale University Press.

Bots Hans, 1971. *Correspondance de Jacques Dupuy et de Nicolas Heinsius (1646-1656)*, La Haye, Nijhoff.

Bourke Evan, 2017. "Female Involvement, Membership, and Centrality: A Social Network Analysis of the Hartlib Circle," *Literature Compass*, 14/4, pp. 1-17.

Burke Peter, 2000. *A Social History of Knowledge: From Gutenberg to Diderot*, Cambridge, Polity Press.

Conroy Melanie, 2021. "Networks, Maps, and Time: Visualizing Historical Networks Using Palladio," *DHQ: Digital Humanities Quarterly*, 15/1, http://www.digitalhumanities.org/dhq/vol/15/1/000534/000534.html

Edelstein Dan, Findlen Paula, Ceserani Giovanna et al., 2017. "Historical Research in a Digital Age: Reflections from the Mapping the Republic of Letters Project," *The American Historical Review*, 122/2, pp. 400-424.

Fernández Riva Gustavo, 2019. "Network Analysis of Medieval Manuscript Transmission. Basic Principles and Methods," *Journal of Historical Network Research*, 3, pp. 30-49.

Ghobrial John-Paul A. (ed.), 2019. "Global History and Microhistory," *Past and Present*, 242 (Supplement 14).

Gibson Abraham, Laubichler Manfred D., Maienschein Jane (eds.), 2019. "Computational History and Philosophy of Science," *Isis*, 110/3, pp. 497-501.

Gingras Yves, 2010. "Mapping the Structure of the Intellectual Field Using Citation and Co-Citation Analysis of Correspondences," *History of European Ideas*, 36/3, pp. 330-339.

Graham Shawn, Milligan Ian, Weingart Scott, 2016. *Exploring big historical data: the historian's macroscope*, London, Imperial College Press.

Gregory Ian L., Ell Paul S., 2007. *Historical GIS: Technologies, Methodologies, and Scholarship*, Cambridge, Cambridge University Press.

Hotson Howard, Wallnig Thomas (eds.), 2019. *Reassembling the Republic of Letters in the Digital Age*, Göttingen, Göttingen University Press.

Kräutli Florian, Valleriani Matteo, 2018. "CorpusTracer: A CIDOC database for tracing knowledge networks," *Digital Scholarship in the Humanities*, 33/2, pp. 336-346.

Pomata Gianna, 2014. "The medical case narrative: Distant reading of an epistemic genre," *Literature and Medicine*, 32/1, pp. 1-23.

Ravenek Walter, Van den Heuvel Charles, Gerritsen Guido, 2017. "The ePistolarium: Origins and Techniques," in Jan Odijk, Arjan van Hessen (eds.), *CLARIN in the Low Countries*, London, Ubiquity Press, pp. 317-323.

Renn Jürgen, 2020. *The Evolution of Knowledge: Rethinking Science for the Anthropocene*, Princeton, N.J., Princeton University Press.

Romano Antonella, 2016. *Impressions de Chine: L'Europe et l'englobement du monde (XVI^e^-XVII^e^ siècle)*, Paris, Fayard.

Rosenberg Daniel, 2003. "Early modern information overload," *Journal of the History of Ideas*, 64, pp. 1-9.

Steinova Evina, 2019. "Innovating Knowledge: Isidore's Etymologiae in the Carolingian Period," *Mittelalter. Interdisziplinäre Forschung und Rezeptionsgeschichte*, 2, pp. 12-15.

Valleriani Matteo, 2020. "Prolegomena to the Study of Early Modern Commentators on Johannes de Sacrobosco's *Tractatus de sphaera*," in Id. (ed.) *De sphaera of Johannes de Sacrobosco in the Early Modern Period*, Cham, Springer, pp. 1-23.

Valleriani Matteo, Giannini Giulia, Giannetto Enrico (eds.), 2022. *Scientific Visual Representations in History*, Cham, Springer (forthcoming).

Valleriani Matteo, Kräutli Florian, 2022. "The Necessity of Linked Data alias Thinking Big in Computational History," in Kaspar Gubler, Christian Hesse, Rainer C. Schwinges (Hrsg.), *Person und Wissen: Bilanz und Perspektiven*, Zürich, vdf Hochschulverlag AG, pp. 171-191.

Valleriani Matteo, Kräutli Florian, Zamani Maryam et al., 2019. "The Emergence of Epistemic Communities in the 'Sphaera' Corpus: Mechanisms of Knowledge Evolution," *Journal of Historical Network Research*, 3, pp. 50-91.

Van den Heuvel Charles, 2015. "Mapping Knowledge Exchange in Early Modern Europe Intellectual and Technological Geographies and Network Representations," *International Journal of Humanities and Arts Computing*, 9/1, pp. 95-114.

Van Miert Dirk, 2014. "What Was the Republic of Letters? A Brief Introduction to a Long History (1417-2008)," *Groniek*, 204/5, pp. 269-287.

Van Vugt Ingeborg, 2017. "Using Multi-Layered Networks to Disclose Books in the Republic of Letters," *Journal of Historical Network Research*, 1, pp. 25-51.

Van Vugt Ingeborg, 2018. *The structure and dynamics of scholarly networks between the Dutch Republic and the Grand Duchy of Tuscany in the 17th century*, PhD dissertation, Scuola Normale Superiore di Pisa/University of Amsterdam.

Zamani Maryam, Tejedor Alejandro, Vogl Malte et al., 2020. "Evolution and Transformation of Early Modern Cosmological Knowledge: A Network Study," *Scientific Reports*,10/19822, pp. 1-15.