Toward a computational history of science: The dynamics of socio-epistemic networks and the renaissance of general relativity

DOI: 10.12871/978883339402237

Toward a computational history of science: The dynamics of socio-epistemic networks and the renaissance of general relativity

Roberto Lalli – Max Planck Institute for the History of Science – rlalli@mpiwg-berlin.mpg.de
Dirk Wintergrün – Max Planck Institute for the History of Science – dwinter@mpiwg-berlin.mpg.de

Abstract: The exploding amount of available historical data provides intriguing possibilities as well as major challenges to historians of science. In the last years, several quantitative methods have been developed in order to analyze historical data. At the same time, new analytical frameworks need to be developed to bring together quantitative methods with the more traditional historians' toolkit. The present paper has a twofold aim. The first one is to briefly review major quantitative approaches that have been developed in the history of science in two areas: data modeling and network analysis. The second part of the contribution focuses on applications of social network analysis to the evolution of knowledge systems. We propose a methodological and conceptual framework aiming at uncovering the dynamical transformations of intra- and inter-connections within and between different *layers* of the scientific enterprise. We define knowledge networks as being composed of three different layers: the social network, the semiotic network, and the semantic network. The first is defined as the collection of relations involving individuals and institutions. The semiotic network is defined as the collection of the material or formal representations of knowledge. The semantic network is the collection of knowledge elements and their relations. We call *socio-epistemic networks* the interlinked set of these three levels. As an illustration of this methodology results drawn from our own work on social and conceptual changes in the history of general relativity in the 20th century will be presented.

Keywords: Computational history, socio-epistemic networks, general relativity.

1. Introduction

This paper aims at providing a short review of some of the main approaches of an emerging research field called "computational history of science", with a focus on the methodological and conceptual framework of *socio-epistemic networks* elaborated by the authors and others at the Department 1 of the Max Planck Institute for the History of Science (Renn *et al.* 2016). The computational history of science is a branch of research within the history of science that has been growing quite rapidly in the last few years. While still marginal, and in some respects controversial, its relevance within the international landscape has sensibly increased, as shown by a recent issue of the journal *Isis*, which dedicated the Focus session to this topic (Gibson *et al.* 2019; Laubichler *et al.* 2019). In the short space of this article we will not be able to do justice to the variety of approaches and the results obtained with them, nor will we be able to fairly address the various criticisms that have been raised against the excessive use of quantitative approaches. However, we hope that by summarizing some major approaches we shall convey a sense of the richness, potentiality, and challenges of computational perspectives in the history of science.

The paper is structured as follows. In the second section, we will provide a historiographical introduction in which we situate the computational history of science within the disciplinary frameworks of the history of science, digital humanities and quantitative history. In the third section, we will review computational methodologies in the humanities that are particularly relevant to the history of science in two areas: data modeling and network analysis. In the fourth section, we will present the framework of the socio-epistemic networks, which integrates various methods for addressing questions of relevance in the history of science. We will finally present, in section five, some applications of this framework to the case study of the history of general relativity which result in new perspectives on the phenomenon named the renaissance of general relativity, namely, the return of general relativity to the mainstream of physics in the post-World War II period after a thirty-year period of stagnation.

2. Historiographical discussion: between digital humanities and quantitative history

The use of quantitative methods in the history of science, or more generally, in the humanities, is a still controversial topic. Computational approaches, like any other methods, are not neutral and might be highly misleading, if used uncritically. Moreover, they have far-reaching conceptual and epistemological consequences which are yet to be understood. No doubt, the emerging community connected to this field hopes that computational methods will in the future provide sound mathematical models of complex historical processes, in order to understand *longue durée* dynamics taking into account the interrelation of factors of a different nature: social, cultural, economic, scientific, and so on. The search of general principles governing historical processes is, in other words, one ambitious goal of this emerging field in line with the program of *cliodynamics* aiming at transforming history into an analytical science (Turchin 2008, 2010). Most historians of science who are using these methodologies are, however, aware of the possible perils of the misuse of computational techniques. They are rather promoting a mixed approach that combines computational and other techniques of the historians' toolkit in a complementary fashion. This move is supported by the observation that, whatever our personal predisposition toward digital technologies, they are so widespread that they are deeply affecting our scholarly practices anyway. The ideological opposition to computational approach and, *vice versa*, an uncritical embracement of it are both considered unproductive, while a better understanding and a careful application of these methods can do much for the field of the history of science in the near future. This is the view we also embrace in our work, even though it must be recognized that there is still much work to be done in order to create and improve the connections between quantitative and qualitative methods and fully open the black boxes of computational methods.

The current rise of the computational history of science is a consequence of a worldwide radical transformation in the way information is stored, accessed and transferred. An article published in *Science* in 2011 has showed how the greatest amount by far of stored information is now in the digital form (Hilbert and López 2011). One of the authors has even proposed to define the beginning of the digital age in the early 2000s, when digital storage started becoming larger than analog storage at a rapid pace (Hilbert 2012). Like many other human activities, our scholarly profession has now become almost unthinkable without computers, PDFs, search engines, online journals, digital databases, and so on. This is a phenomenon certainly not limited to the history of science or to history, but to the entire range of the humanities. As American historian Lincoln Mullen (2010) wrote in 2010, "we are all digital humanists now!"

But, the turn toward the computational history of science is not simply, nor uniquely, related to the, inevitable, growth of digital methods in historical research and scholarly communication practices. With computational history of science, one does not mean simply the intersection of the history of science with digital humanities. There is a second trend that flows into the computation approaches, and it is what is called the *second wave of quantitative history*. The first wave had its apogee in the 1970s, when quantitative methodologies became popular in historical research, in connection with an increasing interest in social and economic histories, comparative approaches, *long durée* perspectives, all areas that were forcefully stimulated by the *Annales* school of historiography (Trevor-Roper 1972). After an initial enthusiasm, this first wave of quantitative history known as "new social history" encountered substantial criticisms in the 1980s when many saw that quantitative and statistical methods were not providing what they promised to achieve: these approaches were extremely expensive, dragging away huge resources that could be used for many more projects; the authors failed to make the assumptions at the basis of their analysis - as well as part of their results - accessible to the entire scholarship; most of the findings turned out to be either obvious or wrong; many research approaches did not take into due consideration the biases in the historical data themselves, etc. (Stone 1979, Lemercier and Zalc 2019; Gibson and Ermus 2019).

While many of the criticisms raised at the time are still important challenges in the application of computational approach, the widespread accessibility of computational facilities and digitalized historical sources has enormously improved the feasibility of data-driven quantitative approaches. This has given hope that computationally intensive methods can revive the historians' ambition to investigate and understand *longue durée* processes, against the trend of focusing on local practices characterizing much scholarship from the 1990s onward. The authors of *The History Manifesto* made it explicit by underlining how the new technological lens of digital analysis, together with the increasingly availability of big data, make the new *longue durée* perspectives much more dynamic and flexible than previous attempts (Guldi and Armitage 2014).

Strikingly enough, the digital technology-driven second wave of quantitative history is not only going in the direction of sparking a return to grand narratives of the past. It is also increasingly employed to develop new approaches to microhistory. French historians Claire Lemercier and Claire Zalc are promoting the application of quantitative methods to microhistory by arguing that, with the correct use of statistical sampling and careful assessment of data availability, each scholar can successfully apply quantitative methods to small, but indicative, samples that might reveal interesting and under-recognized historical patterns (Lemercier and Zalc 2019).

The emerging field of computational history of science situates itself within these broad movements of digital humanities, and the *second wave* of quantitative history sparked by the digital revolution. The question is then what specifically characterizes the history of science, or, more broadly, of knowledge with respect to other branches of historical research in terms of application of computational methods. In other words, the question is whether the history of science has features that require specific methodological and conceptual approaches within the spectrum of computational techniques.

The response to this question is far from straightforward, and it is difficult to identify a consensus in the emerging community. We believe, however, that the history of knowledge does, in fact, require some specific approaches that look at complex systems comprehending layers of very different types[1]. In this sense, the network theoretical framework emerges as the most promising approach to bring together digital humanities and the history of science, as it allows a joint investigation of the dynamics of such layers and the interrelations between them, as we shall discuss in our presentation of the socio-epistemic networks in Sec. 4. A more general perspective concerns the application of the evolutionary framework drawn from biological sciences to the history of science. Those who propose this view, such as historians of science Jürgen Renn and Manfred Laubichler, don't use the evolutionary framework at the metaphorical level. Rather, they specifically contend that the extended evolutionary theory might explain, rather than describe, what is indeed called the evolution of knowledge (Renn 2020; Renn and Laubichler 2017). Renn and Laubichler are also two of the main proponents of the computational history of knowledge and explicitly relate the use of computational methods with novel interpretations of historical processes within the evolutionary conceptual framework, as computational methods bring a "new ontology and epistemology of knowledge" (Laubichler *et al.* 2019, p. 504). The approach of socio-epistemic networks has certainly its origins within these perspectives, but we contend that the evolutionary conceptual framework is not necessary to make sense of it. The methods themselves, indeed, do not rely on the assumptions connected to the evolutionary view of the history of knowledge, although one might of course give a stronger interpretative account within the evolutionary perspective.

[1] For the relations between history of science and history of knowledge see: (Renn 2015).

3. Review of digital methods in historical research

After this brief introduction, we will give a sketch of approaches in areas that are necessary to conceptualize and build the framework of socio-epistemic networks: data modeling and network analysis.[2]

3.1. Data modeling

Reliable data retrieval is one of the biggest challenges in the computational history of science. There is still no standardized way to keep and store data. Each research project has its own data repository, and the collected data are strongly dependent on the questions the specific project is aimed to address. In general, these data are not openly available, and, in any case, the lack of standards makes them difficult to re-utilize in subsequent analyses. To become more effective, computational approaches need a standardization of the taking, storing and transferring of data (Damerow and Wintergrün 2019).

A relatively recent method to overcome these difficulties consists of building semantic models of data in machine-readable specifications. Semantic data modeling aims at presenting and storing data in a logical way by giving semantic information attached to the data. This modeling allows, in principle, to quickly retrieve a set of interlinked data though semantic queries. The most common of these machine-readable specifications is the RDF (Resource Description Framework) families of abstract specifications, which normally uses the triple subject/predicate/object as the unit of analysis (Wintergrün 2019).

To make these models interpretable outside the specific domain where they were originally build for, these models should be based on reference models, i.e. internationally standardized conceptual frameworks for the description of information that have been designed by a recognized community of experts. Currently, the most promising approach for semantic modeling of historical data is to moderately extend the CIDOC/CRM, namely the Conceptual Reference Model developed by the International Committee for Documentation for the object-based description of cultural heritage documentation. The CIDOC/CRM and its extension FRBRoo (Functional Requirements for Bibliographic Recodes – object oriented) are intended to become the standard reference model for the description of the underlying semantic of bibliographical and museum information (Doerr 2003). One example of complex data modeling based on the CIDOC/CRM framework is the database CorpusTracer created by Matteo Valleriani and Florian Kräutli, for the project *De Sphaera*, which investigates the commentary tradition of the thirteenth century university textbook *Tractatus de Sphaera* by Johannes de Sacrobosco. The data about each book included in the dataset are stored as triples in RDF and modeled using the FRBRoo extension of the CIDOC/CRM reference model. Their

---

[2] In our paper we could have included a review of a number of computational methods for textual analysis that are being developed to extract and visualize data and meta-data from large textual corpora (for an introduction see Moretti, 2013). Within the perspective of the socio-epistemic networks, these methods are extremely relevant to produce what we call semantic networks out of the material representations of knowledge (see Sec. 4). We didn't include such a review because of lack of space.

semantic data model captures bibliographic data of each treatise as well as data on the individual texts they contain (Kräutli and Valleriani 2018).

3.2. Network analysis

Once captured in standardized and machine-readable forms, data can be visualized and analyzed in multiple ways, depending on the historical questions one wants to address. One of the most used computational methods in the history of science is based on notions, methods and tools of network theory (Barabási *et al.* 2006). Network theory is a branch of graph theory, which, in turn, is the branch of mathematics modelling pairwise relations between discrete objects. As a possible formalization of complex dynamic (discrete) systems, network theory applies graph theoretical notions to the study of a wide range of structures in a variety of disciplines including physics, biology, engineering, computer science, economics, sociology etc. In network theory, the entities and the relations between them have been called in very different ways, according to the specific spheres of application. In this paper, we will call *nodes* the entities and *edges* the relations between them. Formal methods of network theory allow answering specific structural questions concerning complex data on relations and about how these relations are connected to the features (*attributes*, in network parlance) of *nodes* (see Fig. 1).

In addition, formal methods require giving some precise assessments of the biases present in the historical sources and on missing elements (Lemercier 2015). Taken together, these features give the possibility of rethinking the history of science from a new perspective and build new narratives, focusing on the relational structures of the scientific enterprise and on their change over time. Only recently, a new generation of scholars has started to use formal network analysis tools to investigate historical data. This trend has given rise to a new research field and scholarly community, which has as a reference point the online platform *Historical Network Research*, and the related online journal established in 2017, which has the purpose of bridging traditional hermeneutics of historical research with the more technical methods of network analysis (Düring 2017).

The application of network theory to sociology, called Social Network Analysis (Wasserman and Faust 1994), constitutes the major conceptual and methodological reference in such historical studies, which aim at investigating from the historical perspective the relevance of individuals through various centrality measures, the structure of communities, the distance between actors, and so forth.[3] While the most studied, social networks are only one of the possible areas of applications of complex network theory to the history of science. Another important application is what physicist Mark Newman (2003) calls the information or knowledge network. This network is not precisely defined in Newman's review, but might be interpreted as the network of relationship between knowledge products. Well-known examples are the various networks that might be created out of the patterns of citations. Another well-known example is the *World Wide Web*,

[3] In network theory, *centrality measures* are a set of formulas aimed at identifying the most important nodes in a network's structure. The *distance* between two nodes is instead defined as the number of edges in the *shortest path* connecting them.

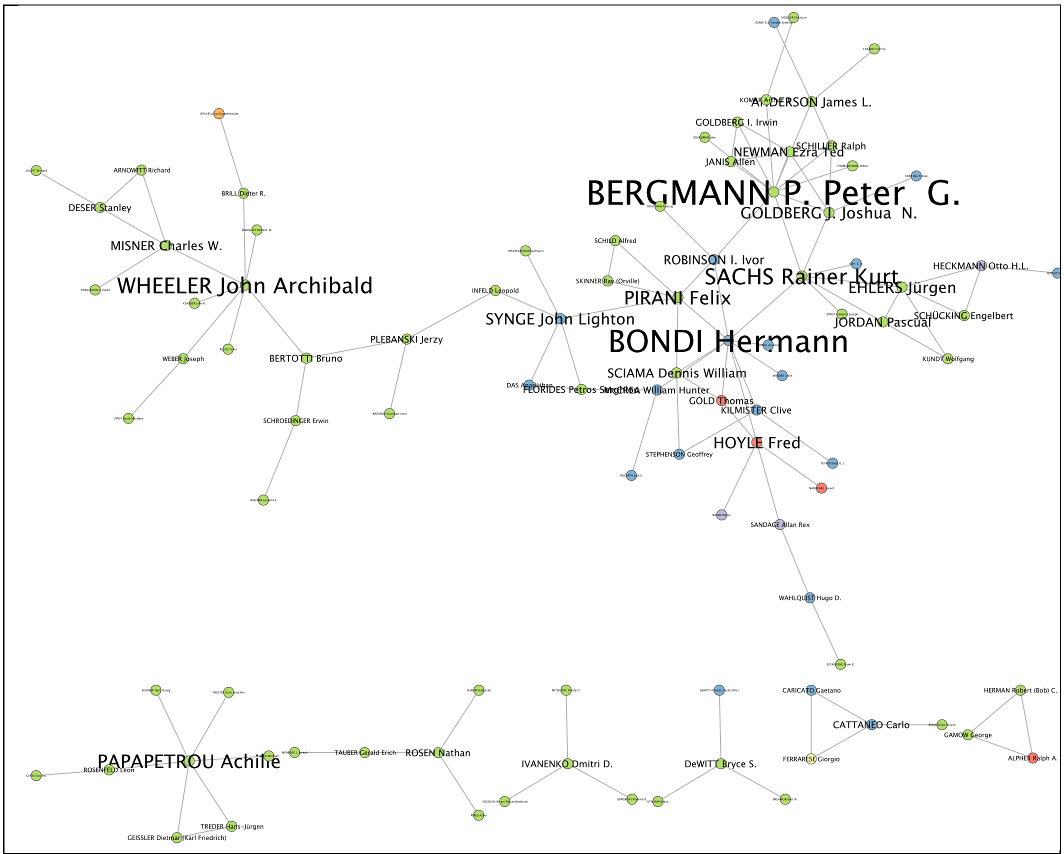

Fig. 1. The image represents a social network in which the nodes are the scientists who were active in research on topics related to general relativity in 1960. The edges represent collaboration links that had taken places within the previous 8 years (namely, between 1952 and 1960). The image shows also one specific attribute of the scientists, namely, the disciplinary domain of the nodes' PhD or, alternatively, of their highest degree: physics (green), mathematics (blue), astronomy (purple), astrophysics (dark orange). The size of the nodes' labels is proportional to the number of edges. This number is called *degree centrality*, which is one way to measure the structural relevance of each node in the network. Historical questions might be, for instance, related to the changing relations between disciplines in the field of general relativity, the importance of specific figures in connecting the network in specific periods, and the identification of different social groups within the network.

which is the network of web pages connected by hyperlinks. A different, and important, kind of network is the one of written words connected through some kinds of relations, such as the co-occurrence in textual units. We can then form a network of words, to which we can apply the various techniques for community detection or *centrality measures*, etc. (see, e.g., Leydesdorff and Welbers 2011).

4. The framework of the socio-epistemic networks

As we mentioned in Sec. 2 the framework of socio-epistemic networks has been elaborated to properly address questions pertaining to the field of the history of knowledge. In the scholarly tradition of the history of science one of the main goals has become to properly address the role of social factors in the production, circulation, and certification of knowledge. In the past decades, various approaches have been developed to connect the social history and the intellectual history of the scientific enterprise. We propose that the approach of socio-epistemic networks has the potential to bridge in a coherent conceptual way the socio-historical analyses of scientists and institutions with the history of ideas and their representations.

Within the perspective of the historical epistemology pursued in our department, we understand knowledge as complex systems of encoded experience represented in material forms by actors in specific social settings. This view is based on the assumptions that knowledge structures have three fundamental layers. The first level is composed of *social actors*, which might of course be individuals, social groups and institutions and the complex relations between them.[4] *Nodes* in the social network might be connected by multiple types of *edges* and have different *attributes*. The social network is itself a very complex structure, whose boundaries strongly depend on the questions one wishes to address. The second level is the *semiotic network*. This rests on the assumption that knowledge is always represented and transmitted by material embodiments that carry meaning. Artifacts, symbol systems, publications are all examples of *semiotic networks*. The *semiotic network* is the set of these material representations of knowledge and the semiotic connections between them. The *semantic network* is composed of cognitive elements that are semantically related to each other. There is no direct access to these cognitive elements in the historical data, so we need proxies in the semiotic network that we might hypothesize to be closely related to the underlying semantic structure. The *nodes* might for instance be words or phrases that represent relevant concepts, and the *edges* might be retrieved by their *co-occurrence* in written texts or parts thereof.[5]

Scholars of various disciplines, including historians, have been interested in all these layers and have developed tools to investigate each layer in a variety of manners. What has been done much more rarely is to connect the analysis of these three layers in order to provide a unified, synthetic framework for the dynamics of knowledge structures (for an exception, see Mutschke and Haase 2001). At the more basic level, this unified approach creates many more *edges* between entities through the process of projecting intra-level and inter-level *edges* into one layer. This approach provides insights on layers of our socio-epistemic network that are notoriously more difficult to analyze on the basis of the available historical sources, such as the *semantic layer*. It is much easier to define and identify social connections *(edges* of the *social layer*), than connections between abstract entities such as concepts, ideas, or research agendas. Let's suppose, for example,

[4] We use institutions and social groups only as actors, if historical evidence does not allow the identification of individuals in actions. A more precise formulation would introduce institutions and groups as graphs, so that interactions between them have to be formulated as hypergraphs.

[5] In textual analysis, *co-occurrence* is the paired presence within a textual unit (e.g., abstracts, paragraphs, sentences, etc.).

that two scientists A and B (*nodes* in the *social layer*) are working on two different research agendas C and D, respectively (*nodes* in the *semantic layer*), which means that there are two inter-layer *edges* connecting the *social* and the *semantic layers* (between A and C, and B and D, respectively). In case A and B start collaborating at a certain point in time (and then create an intra-layer *edge* in the *social layer*), the socio-epistemic network approach leads to the hypothesis that at the same time an intra-layer *edge* is created in the *semantic layer* between research agendas C and D. This procedure is a projection of a connection in the *social layer* we have historical evidence about (e.g., A and B writing a paper together) in the *semantic layer*.

This is just the basic gain that one can obtain if one uses this taxonomy. The more ambitious goal is, however, developing dynamic models of how changes at one level impact the dynamics of other layers, which in turn enables the historians to formulate sound hypotheses on causal relationships on historical processes. The task is formidable, as it requires new mathematical techniques in the generalization of network theory, called multilayer network analysis. This is a relatively new branch of complex system research that studies in a unique formalism the structure and dynamics of various complex networks with different kinds of *nodes* and relations, both between the layers and within the layers (Bianconi 2018).

4.1. The socio-epistemic network of general relativity, 1925-1970

While we are still developing the methodology, we can already present some results from our project on the history of general relativity obtained by using the conceptual apparatus of socio-semantic networks and, partially, its methods. General relativity is now considered one of the pillars of modern physics, but it did not reach this status soon after the famous confirmation by Eddington in 1919, as it is commonly believed. It has been convincingly argued that general relativity remained a marginal theory of little interest to theoretical physicists for many years from the mid-1920s to the mid-1950s, after which it gradually returned to the mainstream of physics (Eisenstaedt 1986). We apply our framework of the socio-epistemic networks to give a quantitative and robust assessment of this change in the status of the theory, between what has been called the *low-water-mark* period and what is coined the *renaissance of general relativity*, and provide a more clear and systematic definition of this transformation (Will 1989).

Through the analysis of the dynamics of the collaboration networks of general relativity, we have provided robust evidence that there was a radical shift in the structure of such networks around 1960, when a giant component started forming at a rapid pace. While the number of authors working on general relativity (the *nodes* of the entire network) started increasing soon after World War II, this did not have immediately a significant impact on the topology of the network and on the connectivity between individuals. A significant shift occurred only fifteen years after the end of World War II, but before the discovery of quasars in 1963. Through this analysis we found that the structural change in the social dimension of general relativity research was, first, independent from serendipitous astrophysical discoveries in the 1960s, and, second, not immediately depending on the rapid increase of the number of physicists following World War II. Our analysis disproves common explanations of the renaissance process. It shows that this phenomenon was not a consequence of astrophysical discoveries in the 1960s, nor was it a simple by-product of socio-economic transformations in the physics landscape after World War II. There was a more complex mechanism that took time to affect the topology of the network, but concerned the internal social dynamics of the scientists working on general relativity-related subjects. With more qualitative analysis and close analysis of the network, one sees that these changes were related to two social changes: the widening of the practice of long postdoctoral education in the 1950s, and the self-organization of the emerging international communities, especially with organization of the first conferences specifically dedicated to gravitation physics from the mid-1950s onward (Lalli *et al.* 2020).

To understand how knowledge changed in the same period, we have studied the *semiotic network* of general relativity using as a *proxy* the co-citation networks of most cited papers related to general relativity from 1947 to 1974 (Small 1973). We used a clustering algorithm to identify research agendas and visualized the outlook in a temporal arrow. For the study of the *semantic layer*, we identified the cluster's topics using algorithms on the titles of the citing articles. This analysis - carried on with the software *Citespace* (Chen 2017) - shows that a major shift occurred also at the *semio-semantic layer* exactly at the same time as the *social layer*, between the mid-1950s and the early 1960s. During the 1940s and 1950s - the last period of the *low-water-mark* phase - the involved scientists were focusing on attempts to substitute the theory of general relativity with a quantum theory of gravity, or a unified field theory of gravitational and electromagnetism or alternative cosmological models. Between the second half of the 1950s and the early 1960s, a unique large cluster of co-cited items emerged out of these different agendas, with papers in which scholars shifted their interest toward problems related to the theory of general relativity and its physical predictions, with a particular attention to gravitational waves. In the 1960s, scientists became involved in particular areas of research concerning the predictions of general relativity that were strongly related to specific astrophysical and astronomical discoveries made in the 1960s.

Combined with the results of our analysis of the social network, we claim that the renaissance process was a two-step process. During the first one, occurring between the mid-1950s and the early 1960s, a highly connected community of scholars formed, that reconfigured the research agendas on general relativity focusing on its role and effectiveness as a proper physical theory rather than a mathematical construct. We call this phase the renaissance of general relativity. The second step, which we call the astrophysical turn, represented a diversification of research agendas in the direction of relativistic astrophysics and physical cosmology of the emerging community strongly shaped by the recent discoveries in astrophysics (quasars in 1963, CMBR in 1964/1965, pulsar in 1968 and even the announcement by Joseph Weber that he had detected gravitational waves) (Blum *et al.* 2018). This analysis provides strong empirical support to previous interpretations proposed by our group (Blum *et al.* 2015, 2016).[6]

[6] While the results of the dynamics of collaboration network has already been published, the part on the *semio-semantic network* will appear in more detail in Lalli R., Howey R., Wintergrün D. (2020), "The Socio-Epistemic

5. Conclusion

In this short piece, we aimed at showing that the computational history of science is an emerging field with strong potential, which comes with a full baggage of approaches, tools, open access software programs, as well as many challenges to be understood and solved. In spite of the potentiality, this approach is at the moment still marginal in the history of science for many reasons and controversies, some of which are connected to the difficulties in fully understanding these methods and not seeing them as black boxes. There is a steep learning curve that is of course quite worrying for practitioners and there is a justifiable resistance against the view that coding has to become indispensable expertise in the historians' toolkit.

To solve these fundamental problems and make full use of the possibilities of these methodologies, research practices and reward systems have to be considerably revised within the humanities. In fact, these sorts of projects might be fruitful only within a scheme of strong and close multi-disciplinary cooperation, including historians of science, mathematicians, statisticians, physicists, data and computer scientists, which is considerably different from the traditional, mostly solitary, works of historians of science. In a certain sense, this field requires a working environment more like the little-science style cooperation in the natural sciences.

If this sort of successful multi-disciplinary collaboration is achieved, these groups have the possibility to address major challenges of these methods: integration of quantitative and qualitative methods based on the greatest openness of assumptions; showing and limiting data biases, making the parameters' choice comprehensible to a non-expert audience, etc. Moreover, the multi-layer network theory is an emerging field and groups working in the computational history of science might contribute to the development of such a field providing questions, data and even possible approaches. Within this context, the approach of socio-semantic network is proposed as the way to address specific questions in the history of science with the aim to bridge tradition in the social and intellectual histories of science. …

References

Barabási A.-L., Watts D.J., Newman M.E.J. (2006). *The structure and dynamics of networks*. Princeton, N.J.: Princeton University Press.

Bianconi G. (2018). *Multilayer Networks: Structure and Function*. Oxford, New York: Oxford University Press.

Blum A., Lalli R., Renn J. (2015). "The Reinvention of General Relativity: A Historiographical Framework for Assessing One Hundred Years of Curved Spacetime". *Isis*, 106, pp. 598–620.

Blum A., Lalli R., Renn J. (2016). "The renaissance of General Relativity: How and why it happened". *Annalen der Physik*, 528, pp. 344–349.

Networks of General Relativity, 1925-1970," in Blum A., Lalli R. and Renn J. (eds.) *The Renaissance of General Relativity in Context* (Boston: Birkhäuser).

Blum A., Lalli R., Renn J. (2018). "Gravitational waves and the long relativity revolution". *Nature Astronomy*, 2, pp. 534–543.

Chen C. (2017). "Science Mapping: A Systematic Review of the Literature". *Journal of Data and Information Science*, 2, pp. 1–40.

Damerow J., Wintergrün D. (2019). "The Hitchhiker's Guide to Data in the History of Science". *Isis*, 110, pp. 513–521.

Doerr M. (2003). "The CIDOC Conceptual Reference Module: An Ontological Approach to Semantic Interoperability of Metadata". *AI Magazine*, 24, pp. 75–75.

Düring M. (2017). *Historical Network Research. Network analysis in the historical disciplines*. [online]. URL: <http://historicalnetworkresearch.org/> [retrieved 14/12/2019].

Eisenstaedt J. (1986). "La relativité générale à l'étiage: 1925 – 1955". *Archive for History of Exact Sciences*, 35, pp. 115–185.

Gibson A., Ermus C. (2019). "The History of Science and the Science of History: Computational Methods, Algorithms, and the Future of the Field". *Isis*, 110, pp. 555–566.

Gibson A., Laubichler M.D., Maienschein J. (2019). "Introduction". *Isis*, 110, pp. 497–501.

Guldi J., Armitage D. (2014). *The history manifesto*. Cambridge, UK: Cambridge University Press.

Hilbert M. (2012). "How much information is there in the 'information society'?". *Significance*, 9, pp. 8–12.

Hilbert M., López P. (2011). "The World's Technological Capacity to Store, Communicate, and Compute Information". *Science*, 332, pp. 60–65.

Kräutli F., Valleriani M. (2018). "CorpusTracer: A CIDOC database for tracing knowledge networks". *Digital Scholarship in the Humanities*, 33, pp. 336–346.

Lalli R., Howey R., Wintergrün D. (2020). "The dynamics of collaboration networks and the history of general relativity, 1925–1970". *Scientometrics*, 122, pp. 1129-1170.

Laubichler M.D., Maienschein J., Renn J. (2019). "Computational History of Knowledge: Challenges and Opportunities". *Isis*, 110, pp. 502–512.

Lemercier C. (2015). Formal network methods in history: why and how?, in Fertig G. (ed.), Social networks, political institutions, and rural societies. Turnhout: Brepols, pp. 281–304.

Lemercier C., Zalc C. (2019). *Quantitative Methods in the Humanities - An Introduction*. Charlottesville: University of Virginia Press.

Leydesdorff L., Welbers K. (2011). "The semantic mapping of words and co-words in contexts". *Journal of Informetrics*, 5, pp. 469–475.

Moretti F. (2013). *Distant Reading.* London: Verso.

Mullen L. (2010). *Digital humanities is a spectrum; or, we're all digital humanists now*. [online]. URL: <https://lincolnmullen.com/blog/digital-humanities-is-a-spectrum-or-we8217re-all-digital-humanists-now/> [retrieved 11/08/2019].

Mutschke P., Haase A.Q. (2001). "Collaboration and Cognitive Structures in Social Science Research Fields. Towards Socio-Cognitive Analysis in Information Systems". *Scientometrics*, 52, pp. 487–502.

Newman M. (2003). “The Structure and Function of Complex Networks”. *SIAM Review*, 45, pp. 167–256.

Renn, J. (2015). “From the History of Science to the History of Knowledge – and Back”. *Centaurus*, 57, pp. 37–53.

Renn J. (2020). *The Evolution of Knowledge*. Princeton: Princeton University Press.

Renn J., Laubichler M. (2017). Extended Evolution and the History of Knowledge. in Stadler F. (ed.), Integrated History and Philosophy of Science: Problems, Perspectives, and Case Studies. Dordrecht: Springer, pp. 109–125.

Renn J., Wintergrün D., Lalli R., Laubichler M., Valleriani M. (2016). *Netzwerke als Wissensspeicher*, in Mittelstraß J., Rüdiger U. (eds.), *Die Zukunft der Wissensspeicher : Forschen, Sammeln und Vermitteln im 21. Jahrhundert*. München: UVK Verlagsgesellschaft Konstanz, pp. 35–79.

Small H. (1973). “Co-citation in the scientific literature: A new measure of the relationship between two documents”. *Journal of the American Society for Information Science*, 24, pp. 265–269.Trevor-Roper, H. R. (1972). “Fernand Braudel, the Annales, and the Mediterranean”. *The Journal of Modern History*, 44, pp. 468–479.

Turchin P. (2008). “Arise ‘cliodynamics’”. *Nature*, 454, pp. 34–35.

Turchin P. (2010). “Launching the Journal”. *Cliodynamics*, 1, pp. 1-2.

Wasserman S., Faust K. (1994). *Social network analysis: methods and applications*. Cambridge, UK: Cambridge University Press.

Will C.M. (1989). *The renaissance of general relativity*, in Davies, P. (ed.), *The New Physics*. Cambridge, UK: Cambridge University Press, pp. 7–33.

Wintergrün D. (2019). Netzwerkanalysen und semantische Datenmodellierung als heuristische Instrumente für die historische Forschung (PhD dissertation). Friedrich-Alexander-Universität Erlangen-Nürnberg.