Preface

Andrea Scharnhorst, Data Archiving and Networked Services & e-Humanities group, The Royal Netherlands Academy of Arts and Sciences, The Netherlands

Models of science – scattered knowledge

After World War II, scientists were increasingly subject to systematic and large scale measurements efforts. The growth and changing roles of science stimulated the need for governmental and "policy support of science" as well as the need for an empirical basis for "science policy". Since then a wealth of monitoring and evaluative indicators has been created. Sociology of science (Bernal 1939; Kuhn 1962; Merton 1973) as well as Scientometrics (Nalimov and Mulchenko 1969; Price 1963) were established as scientific fields. The Society for Social Studies of Science (4S), the European Association for the Study of Science and Technology (EASST) and the International Society for Scientometrics and Informetrics (ISSI), among others, are active as professional organisations in this field. At their conferences "models of science" occasionally appear, but are not presented in a systematic way on a regular basis. Not only other knowledge domains, such as sociology, philosophy, economics, but also physics apply their models to science (Lucio-Arias and Scharnhorst 2011, Chap. 2), but so far there has been no common reference point such as a conference series, edited books, or monographs devoted to modeling science. The only exception to our knowledge, beyond review sections in journal articles (e.g., Börner et al. 2004), a review in ARIST (Börner et al. 2003), and a recent special issue on Science of Science (Börner and Scharnhorst 2009), is the monograph of Yablonskiĭ published 1986 in Russian with Nauka and not translated into English (Matematicheskie Modeli v Issledovanii Nauki (in Russian) Nauka, Moscow (Yablonskiĭ 1986)). This edited volume aims to fill this gap by presenting an overview about major current trends in modeling of science (Chaps. 3 (Vitanov and Ausloos 2011), 4 (Payette 2011), 5 (Hanauske 2011), 6 (Mali et al. 2011), and 7

(Radicchi et al. 2011)) and a general framework to relate these trends to each other (Börner et al. 2011, Chap. 1).

New possibilities and challenges from information science – mapping science

This book is also an expression of a growing interest in the field of modeling science. One origin of this development can be found in recent achievements in information and computer sciences. They have made it possible to visualize research activities at an unprecedented scale and with a high level of sophistication (Börner et al. 2003). Networks of publications and their citation patterns, word use, collaborating researchers, or topics in e-mail threads have been measured, analysed and visualized over time. With the emergence of network science (Chaps. 6 (Mali et al. 2011) and 7 (Radicchi et al. 2011)) as a new cross-disciplinary approach (Barabási 2002; Barabási et al. 2002) and in particular with the achievements of visualizing knowledge domains in the information sciences (Shiffrin and Börner 2004), old dreams of mapping the sciences (Garfield et al. 1964; Small and Griffith 1974) can now be realized. A prominent example of this approach are the so-called "maps of science" which show all scientific disciplines—as far as their activities are covered by the ISI Thompson Reuters Web of Knowledge, Elsevier's Scopus, or other databases (Boyack et al. 2005). A prominent initiative for mapping science is the NSF funded "Mapping Science" exhibit (http://scimaps.org) informing a wide audience about a new "cartography of science" (Börner 2010). The new maps of science inspire new models as explanatory tools for emergent structures of the science system. Mathematical models of complex systems play a specific role in this discourse.

Beyond mapping – towards explanations

Information gathering about science as a backbone of the knowledge society is only one aspect of these new developments. These instruments are also meant as tools to detect and maybe forecast conditions under which scientific discoveries emerge and areas where these discoveries can be found. At the same time, basic questions about the understanding of science are raised, such as who are the actors driving the development of science: individuals, groups or institutions. Earlier large-scale maps concentrated on scientific communications as manifested in papers and their citation interlinkage (Scharnhorst and Garfield 2010). Partly, this was due to the fact that unique author names are hard to determine because of same names, name variants and misspellings. So, a large part of bibliometrics and scientometrics analyses texts (titles, keywords, words, references). Some automated techniques have partly solved this problem, at least on a higher level of aggregation. In maps of scientific communication, authors as well as institutions can now be made visible with a higher reliability. To explain the networks in which researchers are linked (by publishing or communicating), current research in social-psychology and sociology of science becomes relevant. Resumé analysis, ethnographic observations, and interviews were presented as ways to gain access to local motivations and behavior the collective effect of which is reflected in the large scale global maps of science. We call this the **return of the actors** in scientometrics research. If one thinks in terms of modeling network of scholars, these models entail assumptions about the behavior of the "nodes" (Mali et al. 2011, Chap. 6). This is the moment when qualitative, quantitative, and mathematical models need to come together.

A second observation concerns the increasing need to explain changes in scholarly activities. The design of mostly static maps of science, social science and the humanities is therefore only a starting point. Ultimately, we need to see and understand the dynamics of science (Börner et al. 2004; Leydesdorff and Schank 2008; Börner and Scharnhorst 2009). Visualizations that show the unfolding of scholarly activities in a 'fast forward' mode can help refute or confirm existing theories and trigger questions for novel research into the basic mechanisms of scientific growth. We call this ***the return of time and dynamics***.

Contribution of models – models as heuristic devices. Meeting between information science and physics

Mathematical models represent a very specific *instrumentarium* to analyse elementary processes behind measurable phenomena on a more global scale. As mentioned above, in particular during the 1970s and 1980s, the science system has been conceptualized as a self-organizing system in sociology (Luhmann 1990) as well as modeled using concepts and techniques from physics and cybernetics (Scharnhorst 1988). Nowadays, network models are proposed for studying scientific collaborations or the emergence of topics. These new approaches to the modeling of science look into the growth of scholarly networks (Barabási et al. 2002; Committee on Network Science for Future Army Applications 2005; Börner et al. 2004), the structure of scientific communities (Newman et al. 2006), the epidemics of ideas on collaboration networks (Bettencourt et al. 2006), scholarly information foraging (Sandstrom 1994), the formation of effective teams (Amaral and Uzzi 2007), the competition of groups about paradigms (Chen et al. 2009), the scientific productivity of generations of scientists over time (Fronczak et al. 2007), and modeling the dynamics of actor networks (Snijders et al. 2007). However, as mentioned above, the many existing models of science have been developed in many different scientific fields ranging from physics, sociology to history of science. They exist often unrelated and independently from each other and are seldom linked to other studies of science. Nevertheless, in the last couple of years we have witnessed several encounters between physics and information sciences (Fortunato 2010; Bollen et al. 2009; Barabási 2002). This book aims to contribute to a consolidation of the knowledge about models and their mutual dependencies.

Outline of the book

The book consists of four parts: Part I – Foundations, Part II – Exemplary Model Types, Part III – Exemplary Model Applications and Part IV – Outlook.

Part I contains two chapters. In Chap. 1 "An introduction to modeling science: Basic model types, key definitions, and a general framework for the comparison of process models" (Katy Börner, Kevin W. Boyack, Staša Milojević and Steven Morris) Börner et al. (2011) develop a set of reference or frames along which models can be ordered and compared. Departing from a general definition of the term "model" the authors identify a set of dichotomies, such as descriptive versus process models, which can be used to differentiate between essence, purpose and insights of different models. Even if the reader might want to extend or alter the prosed criteria, he or she has to accept that no comparison of models is possible without a clear articulation of their main elements (units, interactions, targeted phenomena) and their tentative ordering in a common reference framework. With a glossary at the end of this chapter, the authors further deliver jigsaw pieces for a common ground on which models can be related to each other.

One cannot understand the emergence and the essence of certain models without looking into the history of modeling science. The emphasis of certain perspectives of modeling science above others is obviously correlated with the overall *Zeitgeist* in a certain time period. Accordingly, the second chapter ("Mathematical approaches to modeling science from an algorithmic-historiography perspective " by Diana Lucio-Arias and Andrea Scharnhorst) (Lucio-Arias and Scharnhorst 2011) describes the history of science models combining a participant story with a bibliometric reconstruction. Histories are always told on the basis of a set of experiences on the one side and a set of norms and values on the other. Consequently, a variety of histories can be found. Only recently the different perception of members of a scientific community could be made visible by a bibliometric analysis of the citation network of this community (Havemann et al. 2010). Chapter 2 (Lucio-Arias and Scharnhorst 2011) chooses the classical method of algorithmic historiography as introduced by Eugene Garfield. One of the most interesting findings is that current threads in mathematical modeling in scientometrics seem to ignore each other while at the same time relying on the same classical papers.

Part II – Exemplary Model Types contains three chapters which all review models belonging to a certain class of mathematics and partly also introduce new model approaches. We are quite aware that these chapters do not cover all occurring threads in the history and presence of science models. Missing are, for example, system dynamics (Sterman 1985) which has been successfully applied in innovation studies and urban development, or entropy and information measures. The threads reviewed in this part of the book have been selected based on the availability of authors. Ofcourse this sequence could be extended. Although these chapters use an individual language, what binds them together is a more generic perspective of science models. All chapters depart from mathematical techniques available and interrogate to which extent they can be used to obtain a better understanding of the science system. Accordingly, the empirical validation of the models is discussed but not in the foreground. These chapters introduce the reader to the details of the model building process in terms of conceptualization, abstraction, operationalization and extension towards increasingly more complex models. In Chap. 3 (Knowledge epidemics and population dynamics models for describing idea diffusion) Nicolai Vitanov and Marcel Ausloos (Vitanov and Ausloos 2011) present a rich inventory of dynamic models based on the behavior of groups of scientists and suitable to describe the emergence and spreading of new ideas in a competitive process. Groups of scientists can be defined based on their actual acquaintance with a certain idea (epidemic models) or their membership in a certain scientific community. That scientists can change their membership in scientific communities creates an extra challenge for modeling. The authors also discuss the role of fluctuations during the emergence of innovation and when best to turn from deterministic models to more complex stochastic models. This chapter also demonstrates that a further methodological exploration is needed to fill the toolbox of science models. With this challenge in mind the introduction of time-lag elements and the combination of time and space are the most original contributions in this chapter. Nicolas Payette (2011) introduces the reader in Chap. 4 "Agent-Based Models of Science" into the world of agent-based modeling as practiced in computational sociology and computational philosophy. Obviously, the type of rule based modeling as proposed by Epstein and Axtell connects very well to known social theories about the behavior of social beings. Payette digs out the longer history of agent-based modeling, which goes back to John von Neumann. Actually, there are links to spin models (widely applied in sociophysics) waiting for further exploration (Stauffer and Solomon 2007). The chapter provides the reader with excellent and clear insights into the inner logic of different ABM approaches to science. In difference to the dominant mathematical language of the previous chapter, in an interesting contrast, Payette compares models qualitatively by mapping their different conceptual frames. He highlights possible links to other model threads such as network models. Matthias Hanauske returns in Chap. 5 "Evolutionary Game Theory and Complex Networks of Scientific Information" (Hanauske 2011) to the power of mathematics and scientific diagrams. Triggered by a real-world phenomenon – the reorganization of the market of scientific publishing – Hanauske questions the possibilities to model the interaction of different players in this process (authors and scientific journals) with game theory. Game theory is designed to explore the consequence of individual strategic behavior in interactions between many individuals. In particular it allows statements for multi-level networked systems – a suitable description for the complex interaction of producers and disseminators of scientific products where the same individuals often switch roles.

Part III – Exemplary Model Applications describes models for two major aspects of scientific communication: co-authoring and referencing. Not surprisingly a network model approach is applied to both phenomena, relying on the different epistemic traditions of sociology and physics. Co-authoring and referencing are both part of scientific production. Consequently, in Chap. 6 "Dynamic Scientific Co-Authorship Networks", Mali et al. (2011) start with the whole universe of scientific communication before zooming into their specific topic of co-authoring. They also start with an excellent history of Social Network Analysis (SNA). Here the reader is provided with detailed context to obtain a better understanding of the sources of some still existing tensions between different network approaches. Among the dynamic models, blockmodeling applied to evolving networks and stochastic actor-oriented models form the cornerstones of this chapter. Empirical studies are extensively reviewed; ordered alongside of dimensions of cross-disciplinary, cross-sectoral and cross-national collaboration pattern; and linked to SNA model insights. Among their own studies one of the interesting findings points to a tension between strongly local (national) connectivity and the requirements of being interwoven into the international (global) knowledge production. Chapter 7, "Citation Networks" of Radicchi et al. (2011) complementary to Chap. 6 (Mali et al. 2011) looks into (citation) networks from a statistical physics perspective. Again we see a recurrent pattern. Following the epistemic tradition of physics, Radicchi et al. (2011) insist on the search for universality and general organizing principles in their network studies where Mali et al. (2011), in the epistemic tradition of sociology, emphasize how best to incorporate the multi-facet roles of individuals in networks and the different context of their link structures. Nevertheless, there is an overlapping area. Against expectations based on the knowledge of how different the citation behavior is in different disciplines, on a statistical level there are still similarities or universalities. It remains open if these 'general laws' are just mathematical artefacts or if the point to a shared feature in citing across disciplines. Also in SNA the aim is to detect a general pattern in social behavior (as for instance by blockmodeling). In both cases the challenge is to give these patterns a meaningful interpretation. Similar to (Mali et al. 2011, Chap. 6), the authors of Chap. 7 (Radicchi et al. 2011) carefully discuss empirical material. Time is a leading theme through both chapters. Time is the 'hidden constructor' behind specific distributions of networks (such as degree distributions). The authors of Chap. 7 address time more explicitly in dynamic models of the evolution of citation networks and diffusion processes across citation networks. Concerning the latter they take a very elegant and original approach – namely to model papers in terms of their received reward by citations. While citation networks are cumulative in time and the position of a paper in such a network cannot change, its perception can change with each new generation of citing papers; so, reward and recognition of a paper can travel in network topologies and in this way, the diffusion of ideas become visible.

The book concludes with Part IV – Outlook. Chapter 8 "Science policy and the challenges for modeling science" partly also reflects on the process of the making of the book, and the lessons learned from it (van den Besselaar et al. 2011). Despite the character of the book as a collection of chapters, authors and editors have taken specific measures to enhance the consistency of it. This becomes visible in the different appendices of the book. A glossary of relevant terms comes as appendix with Börner et al. (2011) Chap 1. Another group of appendices lists the (historic) knowledge base of the field – adding details to Lucio-Arias and Scharnhorst 2011, Chap. 2. Also all model chapters in Parts II and III contain overviews and short descriptions of the models they address. They also contain text boxes (Key points) highlighting main insights for the general audience and/or science policy makers.

References

Amaral LAN, Uzzi B (2007) Complex systems – a new paradigm for the integrative study of management, physical, and technological systems. Manag Sci 53(7):1033–1035 (DOI: 10.1287/mnsc.1060.0696)

Barabási AL (2002) Linked: The new science of networks. Perseus, Cambridge, MA

Barabási AL, Jeong H, Néda Z, Ravasz E, Schubert A, Vicsek T (2002) Evolution of the social network of scientific collaborations. Physica A 311(3–4):590–614 (DOI: 10.1016/S0378-4371(02)00736-7), also available as arXiv preprint arXiv:cond-mat/0104162

Bernal JD (1939) The social function of science. George Routledge & Sons, London

Bettencourt LMA, Cintrón-Arias A, Kaiser DI, Castillo-Chávez C (2006) The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. Physica A 364:513–536 (DOI: 10.1016/j.physa.2005.08.083), also available as arXiv preprint arXiv:physics/050206

Bollen J, van de Sompel H, Hagberg A, Bettencourt L, Chute R, Rodriguez MA, Balakireva L (2009) Clickstream data yields high-resolution maps of science. PLoS ONE 4(3):e4803, 11 pp (DOI: 10.1371/journal.pone.0004803)

Börner K (2010) Atlas of science: Visualizing what we know. MIT Press, Cambridge, MA

Börner K, Boyack KW, Milojević S, Morris S (2011) Basic model types, key definitions, and a general framework for the comparison of process models. In: Scharnhorst A, Börner K, van den Besselaar P (eds) Models of science dynamics. Springer, Berlin, Ch 1, pp 3–22

Börner K, Chen C, Boyack KW (2003) Visualizing knowledge domains. Ann Rev Inf Sci Technol 37:179–255 (DOI: 10.1002/aris.1440370106)

Börner K, Maru JT, Goldstone RL (2004) The simultaneous evolution of author and paper networks. Proc Natl Acad Sci USA 101(Suppl 1):5266–5273 (DOI: 10.1073/pnas.0307625100), also available as arXiv preprint arXiv:cond-mat/0311459. The article is part of the special issue: (2004) Mapping knowledge domains: Arthur M. Sackler Colloquium of the National Academy of Sciences. Proc Natl Acad Sci USA 101(Suppl 1):5183–5310

Börner K, Scharnhorst A (2009) Visual conceptualizations and models of science. Editorial for the special issue 'Science of science: Conceptualizations and models of science'. J Informetr 3(3):161–172 (DOI: 10.1016/j.joi.2009.03.008), also available as arXiv preprint arXiv:0903.3562

Boyack KW, Klavans R, Börner K (2005) Mapping the backbone of science. Scientometrics 64(3):351–374 (DOI: 10.1007/s11192-005-0255-6)

Chen C, Chen Y, Horowitz M, Hou H, Liu Z, Pellegrino D (2009) Towards an explanatory and computational theory of scientific discovery. J Informetr 3(3):191–209. (DOI: 10.1016/j.joi.2009.03.004), also available as arXiv preprint arXiv:0904.1439

Committee on Network Science for Future Army Applications NRC (2005) Network science. The National Academies Press, Washington, DC, available online at the URL: http://www.nap.edu/catalog.php?record_id=11516

Fortunato S (2010) Community detection in graphs. Phys Reports 486(3–5):75–174 (DOI: 10.1016/j.physrep.2009.11.002), also available as arXiv preprint arXiv:0906.0612

Fronczak P, Fronczak A, Hołyst JA (2007) Publish or perish: Analysis of scientific productivity using maximum entropy principle and fluctuation-dissipation theorem. Phys Rev E 75(2):026103, 9 pp (DOI: 10.1103/PhysRevE.75.026103), also available as arXiv preprint arXiv:physics/0606190

Garfield E, Sher IH, Torpie RJ (1964) The use of citation data in writing the history of science. Technical report, Institute for Scientific Information Inc. Philadelphia, PA. Report of research for the Air Force Office of Scientific Research under contract AF49 (638)-1256, also available online at the URL: http://www.garfield.library.upenn.edu/papers/useofcitdatawritinghistofsci.pdf

Hanauske M (2011) Evolutionary game theory and complex networks of scientific information. In: Scharnhorst A, Börner K, van den Besselaar P (eds) Models of science dynamics. Springer, Berlin, Ch 5, pp 159–191

Havemann F, Heinz M, Struck A, Gläser J (2010) Identification of overlapping communities by locally calculating community-changing resolution levels. Arxiv preprint arXiv:1008.1004

Kuhn TS (1962) The structure of scientific revolutions. University of Chicago Press, Chicago, IL

Leydesdorff L, Schank T (2008) Dynamic animations of journal maps: Indicators of structural changes and interdisciplinary developments. J Am Soc Inf Sci Technol 59(11):1532–2890 (DOI: 10.1002/asi.20891), also available as arXiv preprint arXiv:0911.1437

Lucio-Arias D, Scharnhorst A (2011) Mathematical approaches to modeling science from an algorithmic-historiography perspective. In: Scharnhorst A, Börner K, van den Besselaar P (eds) Models of science dynamics. Springer, Berlin, Ch 2, pp 23–66

Luhmann N (1990) Die Wissenschaft der Gesellschaft. Suhrkamp, Frankfurt am Main

Mali F, Kronegger L, Doreian P, Ferligoj A (2011) Dynamic scientific co-authorship networks. In: Scharnhorst A, Börner K, van den Besselaar P (eds) Models of science dynamics. Springer, Berlin, Ch 6, pp 195–232

Merton RK (1973) The sociology of science: Theoretical and empirical investigations. University of Chicago Press, Chicago, IL

Nalimov VV, Mul'chenko ZM (1969) Naukometriya. Nauka, Moscow

Newman MEJ, Barabási AL, Watts DJ (2006) The structure and dynamics of networks. Princeton Studies in Complexity Princeton University Press, Princeton, NJ

Payette N (2011) Agent-based models of science. In: Scharnhorst A, Börner K, van den Besselaar P (eds) Models of science dynamics. Springer, Berlin, Ch 4, pp 127–158

Price DJ de Solla (1963) Little science, big science. Columbia University Press, New York, NY

Radicchi F, Fortunato S, Vespignani A (2011) Citation networks. In: Scharnhorst A, Börner K, van den Besselaar P (eds) Models of science dynamics. Springer, Berlin, Ch 7, pp 233–257

Sandstrom PE (1994) An optimal foraging approach to information seeking and use. Libr Q 64(4):414–449 (stable JSTOR URL: http://www.jstor.org/stable/4308969)

Scharnhorst A (1988) Philosophische, methodologische und erkenntnistheoretische Probleme der Anwendung von Instrumentarien der Physik der Selbstorganisation auf quantitative Aspekte der Wissenschaftsentwicklung. Dr. phil. (PhD) Thesis. Humboldt-Universität zu Berlin, Berlin

Scharnhorst A, Garfield E (2010) Tracing scientific influence. Dyn Socio-Econ Syst Int J 2(1):1–33, available online at the URL: http://www.dyses.org.ar/ojs2333/index.php/jdyses/article/view/11, also available as arXiv preprint arXiv:1010.3525

Shiffrin RM, Börner K (2004) Mapping knowledge domains. Proc Natl Acad Sci USA 101(Suppl 1):5183–5185, (DOI: 10.1073/pnas.0307852100), also available online at the URL: http://ivl.slis.indiana.edu/km/pub/2004-shiffrin-mapknow.pdf. The article is part of the special issue: (2004) Mapping knowledge domains: Arthur M. Sackler Colloquium of the National Academy of Sciences. Proc Natl Acad Sci U S A 101(Suppl 1):5183–5310

Small H, Griffith BC (1974) The structure of scientific literatures I: Identifying and graphing specialties. Sci Stud 4(1):17–40 (DOI: 10.1177/030631277400400102, stable JSTOR URL: http://www.jstor.org/stable/284536)

Snijders TAB, Steglich CEG, Schweinberger M, Huisman K (2008) Manual for SIENA version 3.2. University of Groningen, ICS, Department of Sociology, Groningen; University of Oxford, Department of Statistics, Oxford, available online at the URL: http://stat.gamma.rug.nl/s_man317.pdf and at the URL: http://www.stats.ox.ac.uk/ snijders /siena/s_man312.pdf

Stauffer D, Solomon S (2007) Ising, Schelling and self-organising segregation. Eur Phys J B 57(4):473–479 (DOI: 10.1140/epjb/e2007-00181-8), also available as arXiv preprint arXiv:physics/0701051

Sterman JD (1985) The growth of knowledge: Testing a theory of scientific revolutions with a formal model. Technol Forecast Soc Change 28(2):93–122 (DOI: 10.1016/0040-1625(85)90009-5)

van den Besselaar P, Börner K, Scharnhorst A (2011) Science policy and the challenges for modeling science. In: Scharnhorst A, Börner K, van den Besselaar P (eds) Models of science dynamics. Springer, Berlin, Ch 8, pp 261–266

Vitanov NK, Ausloos MR (2011) Knowledge epidemics and population dynamics models for describing idea diffusion. In: Scharnhorst A, Börner K, van den Besselaar P (eds) Models of science dynamics. Springer, Berlin, Ch 3, pp 69–125

Yablonskiĭ AI (1986) Matematicheskie modeli v issledovanii nauki. Nauka, Moscow

http://www.springer.com/978-3-642-23067-7

Models of Science Dynamics
Encounters Between Complexity Theory and Information Sciences
Scharnhorst, A.; Börner, K.; van den Besselaar, P. (Eds.)
2012, XXX, 270 p., Hardcover
ISBN: 978-3-642-23067-7