

# Symposium Scientometric Maps and Dynamic Models of Science and Scientific Collaboration Networks (SYSM)

jointly organized by  
 Physics of Socio-Economic Systems Division (SOE),  
 Dynamics and Statistical Physics Division (DY),  
 Biological Physics Division (BP) and  
 Working Group young DPG (AGjDPG)

Jens Christian Claussen  
 Jacobs University Bremen  
 Computational Systems Biology Lab  
 Campus Ring 1, Research 2  
 D-28759 Bremen, Germany  
 j.claussen@jacobs-university.de

Andrea Scharnhorst  
 Royal Netherlands Academy of Arts and Sciences  
 Data Archiving and Networked Services  
 Kloveniersburgwal 29  
 NL-1011 JV Amsterdam, Netherlands  
 andrea.scharnhorst@dans.knaw.nl

Scientometrics is collecting quantitative data of scientific publications. Following approaches by Börner et al., linkage and text similarity measures can be utilized to approximately embed publications (and topics, journals, authors) in a metric space, and already two dimensions allow for a visualization of scientific disciplines and their hierarchical structure of subdisciplines. How does this structure emerge dynamically? How can the emergence of the current structure of scientific disciplines be explained? Which mechanisms influence interdisciplinarity and originality? In which parameter regimes are rich-get-richer mechanisms leading to sole continuation of old topics? This symposium provides a dynamical systems perspective on the scientific process, and outlines data-driven and predictive methods by which the influence of research, review and funding policies as well as individual publication strategies can be studied.

## Overview of Invited Talks and Sessions

(Lecture room H1)

### Invited Talks

|          |     |             |    |   |
|----------|-----|-------------|----|---|
| SYSM 1.1 | Thu | 9:30–10:00  | H1 | <b>Science Forecasts: Measuring, Predicting, and Communicating Scientific Developments</b> — •KATY BÖRNER                             |
| SYSM 1.2 | Thu | 10:00–10:30 | H1 | <b>Mapping science with variable-order Markov dynamics reveal overlapping fields and multidisciplinary journals</b> — •MARTIN ROSVALL |
| SYSM 1.3 | Thu | 10:30–11:00 | H1 | <b>Network algorithms for reputation and quality in scholarly data</b> — •MATÚŠ MEDO, MANUEL MARIANI, YI-CHENG ZHANG                  |
| SYSM 1.4 | Thu | 11:15–11:45 | H1 | <b>Modeling scientific networks in social media</b> — •CASSIDY SUGIMOTO   |
| SYSM 1.5 | Thu | 11:45–12:15 | H1 | <b>Modeling scientific collaboration across multiple scales: from individuals to Europe</b> — •ALEXANDER PETERSEN                     |

### Sessions

|              |     |            |    |   |
|--------------|-----|------------|----|---|
| SYSM 1.1–1.5 | Thu | 9:30–12:15 | H1 | <b>Scientometric Maps and Models of Science and Scientific Collaboration Networks</b> |
|--------------|-----|------------|----|---|

**SYSM 1: Scientometric Maps and Models of Science and Scientific Collaboration Networks**

Time: Thursday 9:30–12:15

Location: H1

**Invited Talk** SYSM 1.1 Thu 9:30 H1  
**Science Forecasts: Measuring, Predicting, and Communicating Scientific Developments** — ●KATY BÖRNER — Indiana University

In a knowledge-based economy, science and technology are omnipresent and their importance is undisputed. Equally evident is the need to allocate resources, both monetary and human, in an effective way to foster innovation. In the preceding decades, data mining, metrics, and indicators have been embraced to gain insights into the structure and evolution of science; but there have been no significant efforts into mathematical, statistical, and computational models that can predict future developments in science, technology, and innovation (STI). While it may not be possible to predict the nature, essence, or the precise extent of impact of the next scientific or technological innovation, it is often possible to predict the circumstances leading to it, i.e., where it is most likely to happen and under which conditions. See Scharnhorst, Börner, and Besselaar, eds. 2012. *Models of Science Dynamics: Encounters Between Complexity Theory and Information Science*. Springer Verlag for an overview of major model types.

This talk reviews and demonstrates the power of computational models for simulating and predicting possible STI developments and futures. In addition, it showcases novel means to broadcast moderated STI forecasts to make them accessible and understandable for a general audience.

**Invited Talk** SYSM 1.2 Thu 10:00 H1  
**Mapping science with variable-order Markov dynamics reveal overlapping fields and multidisciplinary journals** — ●MARTIN ROSVALL — Umeå University, Sweden

To better understand the parallel human endeavor of science, we need good maps that both simplify and highlight the flows of ideas and underlying research organization. However, current maps of science cannot effectively identify the multilevel and overlapping fields of science with multidisciplinary journals. For example, whereas maps based on citations between journals in first-order Markov models can only assign each journal to a single field, maps based on multi-step citation chains in higher-order Markov models become computationally infeasible already for moderate-sized systems. To overcome these problems, we introduce a method that uses model selection to find the appropriate variable-order Markov model. We also present interactive maps of science that highlights the assignments of multidisciplinary journals and how ideas flow through those journals.

**Invited Talk** SYSM 1.3 Thu 10:30 H1  
**Network algorithms for reputation and quality in scholarly data** — ●MATÚŠ MEDO, MANUEL MARIANI, and YI-CHENG ZHANG — University of Fribourg, Fribourg, Switzerland

The ever-increasing quantity and complexity of scientific production have made it difficult for researchers to keep track of advances in their own fields. This, together with growing popularity of online scientific communities, calls for the development of effective information filtering tools. Network theory is an important driving aspect for such algorithms. We will first discuss the case of an online scientific community where users and papers form a bipartite network which can be effectively used to evaluate the reputation of users and fitness of papers. We show that when the input data is extended to a multilayer net-

work including users, papers and authors, the resulting performance improves on multiple levels. In particular, top papers have higher citation count and top authors have higher h-index than top papers and top authors chosen by other algorithms. We will then move to stress the role of time in scholarly data. Most research metrics either ignore time (such as the h index) or consider it in an ad-hoc fashion (such as the m quotient). On the example of PageRank which has been used in the past to assess the quality of papers, we show that a demonstrably better ranking of papers can be obtained by considering time in a principled way.

**15 min. break**

**Invited Talk** SYSM 1.4 Thu 11:15 H1  
**Modeling scientific networks in social media** — ●CASSIDY SUGIMOTO — School of Informatics and Computing, Indiana University Bloomington, USA

This talk will examine the role of social media in constructing new or reinforcing old epistemic communities. In particular, we will analyze the interconnectivity of scientists on social media platforms according to their disciplinary affiliation and the degree to which these networks reinforce or contrast with models constructed through collaboration and citation relations. We will analyze the role of gender and other socio-demographic characteristics where possible.

**Invited Talk** SYSM 1.5 Thu 11:45 H1  
**Modeling scientific collaboration across multiple scales: from individuals to Europe** — ●ALEXANDER PETERSEN — IMT Lucca Institute for Advanced Studies, Lucca, Italy

Quantitative measures are becoming increasingly prevalent at all scales of scientific evaluation, largely due to the advent of large comprehensive publication databases that allow for detailed studies of ideas, people, and institutions, and the vast networks connecting them. As such, there is plenty of room to apply methods from complex systems to address policy-oriented issues relevant to the entire science system. In the first half, I will discuss micro-scale patterns of collaboration from a researcher's local 'ego' perspective, showing that scientific collaboration is characterized by a high turnover rate juxtaposed with frequent 'life partners'. I will show that these extremely strong collaborations have a significant positive impact on productivity and citations – the apostle effect – representing the measurable advantage of 'super' social ties. In the second half, I will discuss macro-scale collaboration patterns concerning the evolution the European Research Area (ERA), a cross-border labor, funding, and mobility scheme aimed at fostering innovation and growth within Europe. However, despite decades of integration policies, recent analyses have shown there to be little cross-border integration in the EU above global trends – i.e. Europe remains a collection of national innovation systems. I will show that high-skilled mobility – i.e. brain drain, largely from East to West following the 2004/2007 EU enlargement – can explain why the cross-border integration of R&D within the ERA is lagging.

**Right after the symposium, participants are invited to enjoy a guided tour of the Places and Spaces: Mapping Science exhibition (<http://scimaps.org>) on display in the foyer of the university library.**