

# Infrastructures as analytical framework for mapping research evaluation landscapes and practices

This is a work in process, please do not re-distribute or cite without the authors permission. A 'cite-able' version (minus the empirical example) is:

Åström, F. (2016). Outlining an analytical framework for mapping research evaluation landscapes. Paper presented at the STI/Enid Conference 2016, Valencia, September 14-16, 2016".

Fredrik Åström

fredrik.astrom@ub.lu.se

Lund University Library, P.O. Box 3, SE 221 00 Lund (Sweden)

## Abstract

This paper suggests an infrastructure perspective, as suggested by Star and Bowker (2006), as an analytical framework for studying the research evaluation landscape. An infrastructure is suggested to be understood, not as a concrete technology, but as a system of contextual factors including 'Actors/Stakeholders', 'Technical systems', and 'Evaluation practices'. How the framework can be operationalized is exemplified by examples from previous and ongoing research, as well as by identify gaps in current research.

## Introduction

Research evaluation and resource allocation systems permeates academic research, and while evaluation practices per se are well established, there is also a growing literature on research evaluation systems and the effects they are having on 'the science system' (de Rijcke et.al, In press).

The aim of this paper is to briefly outline a framework for understanding the complex landscape of research evaluation; and in particular evaluation systems based on the use of bibliometric indicators, to identify from what different perspectives these systems can be analyzed and understood as an infrastructure (Star & Bowker, 2006). The basis for developing the framework is examples from previous and current research, as well by identifying gaps in research so far.

## Background

Over the last three or so decades, we have seen substantial changes in the governance of science (e.g. Whitley and Gläser, 2007); a change that from a policy perspective has been described as change from a linear model to an innovation systems model (e.g. Elzinga, 1995). These changes are often seen as related to the notion of 'new public management' (NPM) and the concepts of the audit and/or evaluation society (Dahler-Larsen, 2012).

There have been different suggestions on how we can gain a theoretical understanding of the development of research evaluation systems, both as a general development in research policy and governance, and suggestions of theories contributing to our understanding of particular aspects of the research evaluation systems. There is a long standing discussion in bibliometrics and STS research on the meaning of citations, e.g. drawing on semiotics (Cronin, 2000), or more along the lines of this paper, Wouters' (2014) suggestion to view the citation as an infrastructure. Recently, Åström and colleagues (2016) suggested 'boundary objects' as a way to theoretically conceptualize scholarly and scientific publications in relation to bibliometrics based research evaluation systems. To understand some of the stakeholders involved in research evaluation processes, Petersohn (In press) has utilized theories on how professions develop. In relation to bibliometrics based research evaluation systems, the conceptualization of research fields and disciplines is also an important aspect, both in terms of how we understand what constitutes fields and disciplines as entities per se (Sugimoto & Weingart, 2015); and how fields are defined in bibliometric analyses and research evaluation systems (Åström et.al, 2016).

Research on the evaluation landscape has been described as having four main research foci: how academic institutions are affected by decreased governmental funding at the same time as NPM related forms of academic governance are introduced, what assessment mechanisms are utilized in national and regional evaluation systems, identifying the dynamics in science and innovation systems, and the effects of indicator use on knowledge production. This last focus address issues of for instance strategic behavior of scholars/scientists in response to evaluation indicators; and when discussing indicator use in research practices, research on different stakeholders is also brought to attention (de Rijcke et.al, In press).

### **Infrastructures**

Star and Bowker (2006) describes infrastructures as representing "one of a number of possible distributions of tasks and properties between hardware, software and people" (Star & Bowker, 2006, p. 232). Drawing on this perspective, we suggest that the evaluation landscape can be understood through the concept of infrastructures, supplying us with an analytical framework for studying evaluation practices. Furthermore, we suggest a categorization of the elements in the evaluation infrastructure in correspondence with Star and Bowker, where "people" take into account the various actors or stakeholders involved in evaluation processes, where "hardware" is understood from the perspective of technical and auxiliary systems, and where "software" represents the evaluation practices per se.

The aspects defined in the categorization are by no means supposed to be considered mutually exclusive, in the same way that categories within these aspects are also often overlapping in many ways. The framework presented here is an attempt at conceptualizing the different aspects of the research evaluation landscape for structured analyses.

#### *"People": Actors/Stakeholders*

The research evaluation landscape is populated by a great variety of actors, such as individual scholars, scientists and research groups; research institutes studying research evaluation; local research administration and services; research funding agencies; national government agencies; research evaluation organizations; and 'content providers' (de Rijcke et.al, In press).

There is a variation of types of organizations, from commercial enterprises, over independent research institutes, to public universities and government organizations, all of which taking part in evaluation practices, in academic research on evaluation practices and the formation of research evaluation policies. The roles of these different actors are often intersecting and overlapping; and there is a substantial diffusion of roles and interests both in-between and within groups of actors. The role of university libraries, as part of local research administration and governance, as well as a service institutions for scholars and scientists has been analyzed by Åström and Hansson (2013) and Sabrina Petersohn (2016); and Petersohn (Forthcoming) is also studying organizations bordering between being academic research institutes and research evaluation consultants; and how such expert organizations provide professional expertise for the implementation of national research policy measures.

#### *“Hardware”: Technical & auxiliary systems*

The aspect traditionally most associated with infrastructures is technical systems, in the case of bibliometrics based research evaluation, primarily bibliographical databases, citation indices and publication repositories. These exist on many different levels: local, national and international, in terms of coverage, and in terms of where and by whom the databases are developed, from locally developed institutional repositories to international databases produced by large commercial entities. To this can also be added a development where traditional databases are appended by a number of new systems of various kinds: there is a growing market for ‘Current Research Information Systems’ (CRIS), as well as for instance research funding application systems; and to this should also be added systems for bibliometric analyses, where there is a great variation from software developed by individuals to commercial research evaluation tools.

This technical infrastructure has primarily been analyzed from perspectives of technical evaluations of the functionality of the systems per se; and the practical applicability of systems in relation to certain evaluation systems and/or practices. Research on the technical infrastructure in a larger context of the research evaluation landscape, however, is rare. This is not for the lack of interesting research questions to address. One issue is of course the implications of – and the different dynamics created by – the use of for instance international citation indices as opposed to locally developed systems. Another complex of questions is related to the increasing communication between systems, where data is being communicated between local publication archives, national research funding application systems, and international citation indices. An example of an attempt at addressing questions related to the technical infrastructures and bibliometrics based research evaluation is recently initiated research on classification issues in relation to bibliometric indicators, where classification systems is seen as a part of a technical infrastructure understood from the point of view of ‘boundary object’ theory (Åström et.al, 2016).

#### *“Software”: Evaluation practices*

The part of the research evaluation infrastructure that arguably have received the most attention from scholars and scientists, is the evaluation practices per se. For instance, the relation between national and local resource allocation systems have been investigated in the Swedish context (Hammarfelt et.al, In press), while Hicks (2012) have analyzed performance-based university research funding systems from a broader perspective.

An important aspect of the evaluation practices is how they relate to wider research policy issues. The most immediate example is of course resource allocation systems building on publication and/or citation indicators, but equally important is other funding and reward programs, mandates on issues related to for instance research data management and open access issues.

### Example

To exemplify the intended use of infrastructures as an analytical framework, the research group CWTS will be described and briefly analyzed. The Centre for Science and Technology Studies (CWTS) at Leiden University (<https://www.cwts.nl/>) in the Netherlands is one of the leading research groups in science studies, and in bibliometrics research in particular. However, CWTS is not only a research institution; at the same time as the members of the center is doing research, CWTS is also a consultancy business providing research performance and evaluation studies to customers such as universities, funding bodies and governments. In addition to the research and the contract evaluations, CWTS are also developing both research performance indicators as well as software for bibliometric analyses. Given the multiple roles of CWTS in terms of research evaluation practices, they become a good example for concretizing infrastructures as an analytical framework (Figure 1).

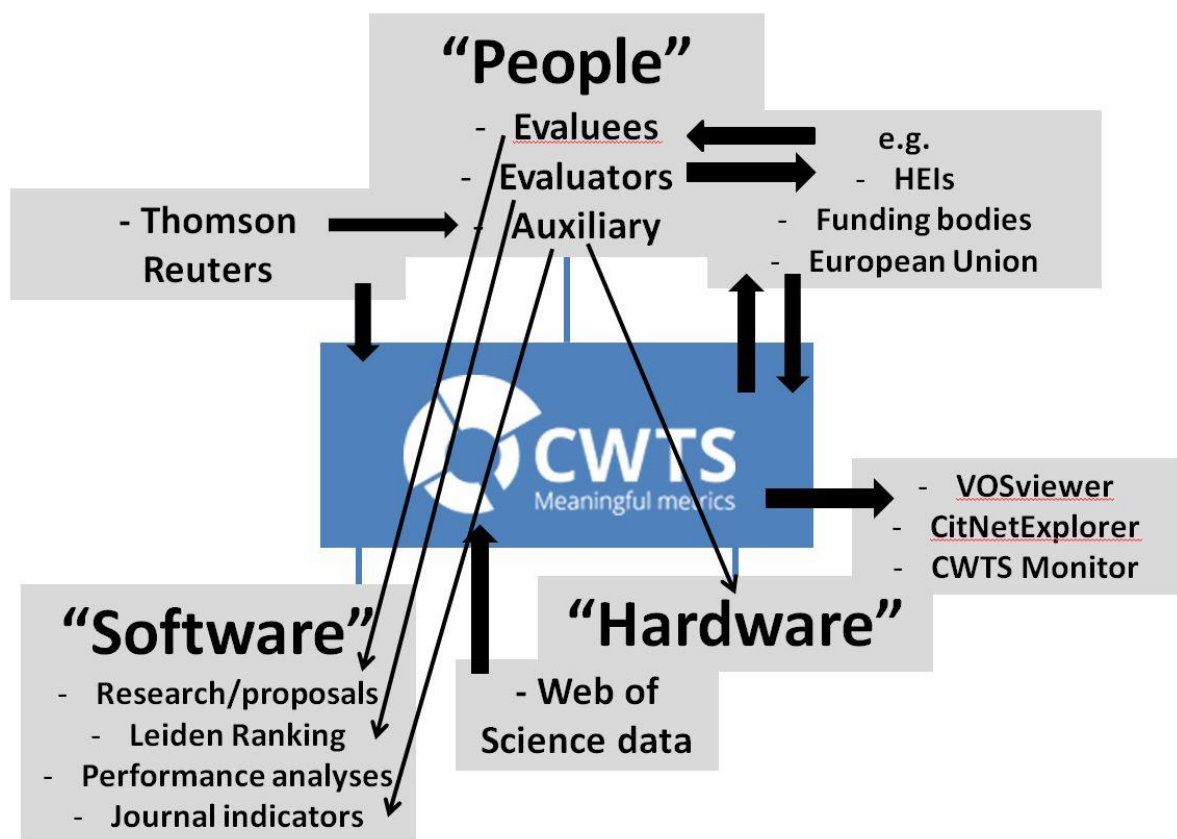


Figure 1. The Centre for Science and Technology Studies (CWTS) at Leiden University from an infrastructure perspective.

In terms of stakeholders involved in research evaluation practices, a distinction has been made between those being evaluated, those evaluating and the auxiliary stakeholders providing for instance data and analytical software. As a research organization CWTS are being evaluated by other

scholars in the field when their members submit articles for publication; and by funding bodies when submitting research proposals. And as part of Leiden University, CWTS are also being evaluated in the Dutch Standard Evaluation Protocol; an assessment scheme performed every six years and published under the authority of the Association of Universities in the Netherlands (VSNU), the Netherlands Organisation for Scientific Research (NWO), and the Royal Netherlands Academy of Arts and Sciences (KNAW) to be presented to the Dutch Minister of Education, Culture and Science.

CWTS also has a role as an evaluator, not only for individual scholars at the center being peer reviewers of manuscripts and research proposals, but also as contracted consultants to universities, funding bodies and governments, performing analyses of higher education institutions (HEIs) using bibliometric indicators. These activities are formally organized separately from the research activities of the center as a company, CWTS B.W. (<http://www.cwtsbv.nl/>), but the ties between the research center and the company are strong, both in terms of people involved and through the ownership of the company, where the center and the university are the main owners.

The output – or the ‘software’ as it has been described in the analytical framework – of CWTS as evaluators is quite diverse. One form of output is the analyses of different HEIs and the evaluation reports formulated around these. Another is the yearly published Leiden ranking (<http://www.leidenranking.com/>), where some 850 universities are ranked through a set of different bibliometrics based indicators. Both these are outputs that are based on the results of the analyses they perform, but another form of output that can be seen as software within the framework of the analytical framework is the development of indicators, such as indicators for analyzing the impact of journals (<http://www.journalindicators.com/>).

When analyzing CWTS from the perspective of auxiliary stakeholders and ‘hardware’, we find on the one hand that CWTS are buying the raw data used in their analyses from Thomson Reuters (although Thomson Reuters have now sold their Intellectual Property & Science Business to Onex and Baring Asia). On the other hand, CWTS can also be seen as a provider of hardware through their development of tools for bibliometric analyses, like VOSviewer, CitNetExplorer, and previously also the CWTS Monitor; an analytical tool for making performance analyses of for instance HEIs. Another aspect of CWTS as a stakeholder with an auxiliary function is the organization of courses for professionals working with issues related to bibliometrics and research evaluation.

Taking one stakeholder as point of departure for mapping serves to illustrate the complexities found in the research evaluation landscape today, where we find a variety of different interactions between stakeholders, as well as overlapping roles even among individual stakeholders. It should be pointed out that CWTS are by no means unique in this sense. Being both evaluators and ‘evaluatees’, as well as developers of both software and hardware, is a common trait in the research evaluation landscape.

## **Discussion**

The purpose of this paper has been to suggest an analytical framework for understanding the effects of research evaluation systems on academia and academic research. Aside from studying the effects per se, as in how for instance scholars and scientists adapt to evaluation criteria in their work, a focus on a broad understanding of the infrastructure is presented, taking into account stakeholders, technical systems and practices. This allows for a structured mapping the evaluation landscape, not the least from a perspective of understanding the ‘materialities’ of research evaluation; and how different aspects of the infrastructure interact.

The complexities found in the evaluation landscape, not the least in terms of how different roles and practices interact, are brought up as an important aspects to consider when analyzing regimes of accountability together with the citation infrastructure (Wouters, 2014); strengthening our claim that the infrastructure perspective can be a valuable framework for understanding research evaluation practices as an activity on the borders between science, science policy and research evaluation as a commercial enterprise.

## References

- Åström, F., Hammarfelt, B. & Hansson, J. (2016). Scientific publications as boundary objects: theorizing the intersection of classification and research evaluation. Paper presented the 9th Conference in the Conceptions of Library and Information Science (CoLIS 9), Uppsala, Sweden, June 27, 2016.
- Åström, F. & Hansson, J. (2013). How implementation of bibliometric practice affects the role of academic libraries. *Journal of Librarianship and Information Science*, 45(4), 316-322.
- Cronin, B. (2000) Semiotics and evaluative bibliometrics. *Journal of Documentation*, 56(4), 440 – 453.
- Dahler-Larsen, P. (2012). *The Evaluation Society*. Stanford: Stanford University.
- de Rijcke, S., Wouters, P.F., Rushforth, A.D., Franssen, T.P. & Hammarfelt, B. (In press). Evaluation practices and effects of indicator use: A literature review. *Research Evaluation*.
- Elzinga, A. (1995). Reflections on Research Evaluation. *Science Studies*, 8(1), 5-23.
- Hammarfelt, B, Nelhans, G., Eklund, P. & Åström, F. (In press). The heterogeneous landscape of bibliometric indicators: Evaluating models for allocating resources at Swedish universities. *Research Evaluation*.
- Hicks D. (2012). Performance-based university research funding systems. *Research Policy*, 41(2), 251–61.
- Petersohn, S. (Forthcoming). Professional evaluation of research excellence: case study of the Netherlands. Paper to be presented at the 4S/EASST Conference, Barcelona 2016.
- Petersohn, S. (2016). Professional competencies and jurisdictional claims in evaluative bibliometrics: The educational mandate of academic librarians. *Education for Information*, 32(2), 165-193.
- Star, S.L. & Bowker, G.C. (2006). How to infrastructure. In: L.A. Lievrouw & S. Livingstone (Eds), *Handbook of New Media: Social Shaping and Social Consequences of ICTs* (pp. 230-245. London: Sage.
- Sugimoto, C.R & Weingart, S. (2015). The kaleidoscope of disciplinarity. *Journal of Documentation*, 71(4), 775-794
- Whitley, R. & Gläser, J. (Eds., 2007). *The Changing governance of the sciences: The advent of research evaluation systems*. Dordrecht: Springer.
- Wouters, P. (2014). The citation: From culture to infrastructure. In: B. Cronin & C.R. Sugimoto (Eds), *Beyond Bibliometrics: Harnessing Multidimensional Indicators of Scholarly Impact* (pp. 47-66). Cambridge, Mass.: MIT Press