Steering science through Output Indicators & Data Capitalism

Author: Ulrich Herb, ORCID: 0000-0002-3500-3119

Affiliations:

- Saarland University & State Library
- ./scidecode science consulting & research

Performance tracking in science

Since the mid-1920s, a scientist's performance was tracked by using bibliometric information such as the number of publications or their citations. Today, there are unprecedented possibilities for controlling science by analysing data on production and use of scientific information, so that citations only play a subordinate role in the evaluation of science. To illustrate this change, we take a look at the history of the Dutch publishing house Elsevier and its metamorphosis into a research intelligence service provider, because it ideally reflects the new possibilities of logging and controlling science.

The story of Elsevier

Elsevier was founded in 1880 as a **Publishing House**, today it is one of the three largest scientific publishers in the world. Elsevier currently publishes 2,786 scientific journals (made available through its platform ScienceDirect), Springer Nature just over 3,000, Wiley about 2,500. Since 2004, Elsevier has no longer been just a publishing house, as this year its Scopus citation (or **Impact) Database** was launched. Getting a journal indexed in Scopus is perceived as a quality criterion. For Elsevier, this means that even its competitors are keen to get their content indexed in Scopus and that Elsevier itself can market Scopus not only as an expensive database, but can also evaluate it for its own purposes. In 2009, Elsevier introduced SciVal, a completely new service whose purpose is the **Benchmarking** of scientific institutions. This service will be discussed in more detail later. The development away from the traditional publishing business became more and more obvious when Atira and its **Research Information** system PURE were acquired in 2012. These systems are used *externally* to present the performance of an institution, *internally* to benchmark and record the research performance of an institution's units and members. In 2013, Elsevier's shopping tour continued and the **Reference Management** tool Mendeley was acquired. Mendeley is used for the management of scientific literature and has also a **Social Network** component. Mendeley may provide information about the relevance of published texts that are not (yet) cited or publications

that seem to be of great interest, but which are not or rarely cited (e.g. because they are more relevant for teaching or methodology than research) or whose citations are not captured by citation databases. Since Mendeley also offers online profiles, users may be inclined to post academic status (e.g. student, PhD candidate, post doc, professor), affiliation, and research topics here, which allows a more accurate assessment of research trends. In 2015, Elsevier bought the Newsflo service, which analyses the impact of science on the basis of **Media Coverage**. The service features a high-precision search for author IDs and affiliation through a connection with Scopus. In the same way, Newsflo is linked a) with Mendeley to assess the societal influence of research, b) with SciVal to visualize the resonance of an institution and its researchers, and c) with PURE, for which a media showcase with live feeds on the media coverage of an institution and its researchers exists. In 2015 Elsevier launched the **Research Data** service Mendeley Data. Mendeley Data provides free storage for research data and of course some Elsevier journals or Society Journals published by Elsevier (e.g. CELL Press) use the service to make article-related research data available. In 2016, Elsevier acquired the **Open Access Service** Social Science Research Network (SSRN), a publication platform for the social sciences that mainly publishes preprints and discussion papers. Elsevier now had a technology in its hands that could be rolled out to other disciplines and had access to information from research that had not yet been formally published, e.g. that was not (yet) registered in Scopus or Mendeley. Also in 2016, Hivebench, a software that allows the **Management of Experiments, Protocols, Analyses and Data** was bought by Elsevier. This tool offers a simple and safe way to manage and store protocols and research data. Hivebench recommends the integration of Hivebench and Mendeley Data, e.g. to secure the long-term archiving of research data. In 2017, Elsevier again shopped in the **Open Access** domain and acquired the repository provider Bepress. In comparison to SSRN and its spin-offs BioRN or ChemRN, which are central publication servers and are not offered as hosting solutions, Bepress is a software that research institutions operate themselves in order to offer scientists a platform for Open Access publication. In February 2017, Elsevier acquired the **Altmetrics** provider PLUM Analytics, whose service PLUM X, like other Altmetrics services, aims to capture the impact of scientific objects, which does not (only) manifest itself in citations, but, for example, in the number of tweets, likes, and mentions in Mendeley libraries. PLUM X differs from comparable offerings such as Impactstory or Altmetrics.com in its obviously competitive claim and benchmarking functionalities and brings with it a vast amount of information about the distribution of scientific objects in social media, news, reference management systems and a variety of other communication channels of a scientific as well as non-scientific nature. The list of object types that PLUM X tracks currently covers 67 types, including audio files, blogs, software, datasets, expert opinions, government documents.

From a publishing house to an operating system: dependence through convenience
Elsevier has created an operating system of highly integrated services that a scientist can use at any phase of his/her work. One communicates with his/her contacts in Mendeley, where one also collects and manages his/her literature. Preprints or other documents are deposited in SSRN-like repositories, which (as is partly the case with SSRN) are linked to matching journals, or on a local Bepress server. Reviewed articles are published in journals that are indexed in ScienceDirect and their citations are tracked in Scopus. If a scientific achievement is evaluated, scientometric information from Scopus and PLUM X are used. That is one reason why even editors of journals *not published by Elsevier* make every effort to have them indexed in Scopus or tracked by PLUM X. Research data is managed with Hivebench and published with Mendeley Data, its impact is measured by PLUM X. The resonance of research within science is measured by Scopus and PLUM X, outside

science by Newsflo and PLUM X. Managing, sharing, publishing, indexing, analysing, evaluating and selling literature, research data and research information - all this is offered by Elsevier. However, one does not only not have to leave the system, one cannot bypass it or ignore its apps.

Googleization?
Anyone who comes up with a comparison with Google in view of Elsevier's impressive data collection should be worried. Anyone can escape Google (perhaps by losing comfort) and use alternative search engines like DuckDuckgo that give high priority to privacy issues or do without Android smartphones and GoogleMail. But a scientist cannot prevent a service provider like Elsevier from measuring and tracking him: even those who don't want to publish in Elsevier journals won't be able to prevent their articles from being indexed in Scopus - or even better: they will welcome the indexing in Scopus, even as a possible supporter of one of the many Elsevier boycotts, since this is a sign of reputation. Moreover, this indexing allows citation impact (as a supposed quality indicator) to be calculated for one' s publications. And anyway: Once the publications are out in the world, they will sooner or later leave their own traces in one of the Elsevier services: For example, when readers list them in their Mendeley libraries or mention them in one of the many data sources of PLUM X. And anyone who, as a scientist, once decided not to use Endnote because he/her didn't want to give his data to Elsevier's competitor and Endnote provider Thomson Scientific, but opted for Mendeley, had to admit that all well-intentioned considerations and the use of innovative software from small providers would eventually become obsolete in view of Elsevier's clever acquisition strategy. Such fatalism is also likely to have spread among universities who, for similar reasons, opposed the use of the research information system CONVERIS, also a product of Thomson Scientific, and preferred the PURE system bought later by Elsevier. And the Open Access community, which promoted and actively used the SSRN service and numerous Bepress servers to foster non-commercial access to scientific information, may have been hit even harder - surprised to realize that they had been working eagerly to acquire content for Elsevier for years.

Data is King
Today Elsevier's business is "Information analytics"[1] and thus no longer the distribution and sale of scientific publications. The systematic processing of the data collected in the above-mentioned services is primarily provided by SciVal. Its functionalities are already proving to be seductive or dangerous - depending on one's point of view: numerous parameters can be evaluated and a variety of statements can be made for benchmarking and recommendations on the focus of future research:[2]

- "**Create and select research entities: Test scenarios by modeling any Research Areas or groups** such as newly evolving interdisciplinary Research Areas, groups of researchers to apply for a large-scale grant program, and departmental renovations"
- "**Select and combine any set of metrics** to measure an institution's or a country's productivity, citation impact, collaboration, subject disciplinarity, visibility and more."
- "Benchmark your progress: View the relative performance of your institution, specific departments, research groups or selection of researchers."
- "Develop collaborative partnerships: Identify and analyze existing and potential collaboration opportunities based on publication output and citation impact."

[1] https://www.elsevier.com [accessed 12.06.2019]
[2] https://www.elsevier.com/solutions/scival/features [accessed 12.06.2019]

- "Analyze research trends: Analyze the research trends of any Research Area with citation and usage data, to discover the top performers and rising stars."

The raw material for data analysis, as known from data capitalism in other contexts, is usually provided by the objects (= the users) themselves: Tools as SciVal formulate their instructions on the basis of information produced by the scientists themselves in services that they can use free of charge, such as Mendeley, SSRN or Bepress, as well as those that cost their universities money, such as scientific journals or research information systems.

So what?

A service such as SciVal can be seen as a tool that, in times of scarce resources and the notion of universities as enterprises, helps to efficiently allocate resources and staff. The fact that Elsevier uses these circumstances to develop, buy and market services is not to be criticized, because the RELX Group as Elsevier's parent company must have the goal of making as much money as possible. More problematic might be the fact that publication, curation and evaluation of science are increasingly in the hands of private sector stakeholders. The European Commission's report "Open innovation, open science, open to the world - a vision for Europe" (2016) identifies Elsevier, Springer Nature and its subsidiary Digital Science, Google and Wikimedia as stewards of scientific information - but no public institutions. A tightrope walk, considering what happens if the return on investment falls short of the expectations: Elsevier, for example, quietly dropped its search engine Scirus 2014, which was praised as the "most comprehensive scientific research tool on the web"[3]. Klein & Hüllmann (2018) go one step beyond and note with regard to Elsevier that the intertwining of publication and evaluation services in particular should be regarded as extremely sensitive: "The potential collusion of interests between the roles of publisher and ranking service provider contradicts the rules of good governance, but is part of Elsevier's business model"[4]. A fact that becomes all the more significant when one knows that Elsevier is a contractor for setting up the European Open Science Monitor (Tennant, 2018).

Perspectives

The descriptions and functionalities of SciVal and other services presented are garnished with different metaphors that can give us a foretaste of how science may be operated in the future: The *economic metaphor* is interfused with the jargon of calculation and efficiency, in the form of calculated cooperation with predicted funding revenues and recommended partnerships with generated returns in the form of impact rates and excellence ratings. The *militaristic metaphor* is perhaps best illustrated by a statement from the PLUM X service's website (2012, before it was purchased by Elsevier): "Arm your researchers to compete for funding"[5]. Also in 2018 the advertising has a military connotation: SciVal "will revolutionize the way in which you develop your research strategy"[6]. If such strategic plans are implemented, the scientists will turn from free scientific actors into agents of a plan designed by research intelligence providers. However, this emerging *scientific planned economy* is paradoxically legitimised by radical market competition between universities.

---

[3] http://web.archive.org/web/20130217135017/http://www.scirus.com/ [accessed 12.06.2019]
[4] translated by Ulrich Herb
[5] http://web.archive.org/web/20160317073116/http://plumanalytics.com/products/plumx-grants/ [accessed 12.06.2019]
[6] https://www.elsevier.com/solutions/scival/releases/topic-prominence-in-science [accessed 12.06.2019]

Other statements refer to the fourth metaphor, the *biological one of selection and breeding*: „Test scenarios by modeling (...) groups of researchers to apply for a large-scale grant program"[7].

No matter what metaphor we choose: Certain decisions, e.g. which researcher is seeking third-party funding in cooperation with which colleague, should no longer be left to the researcher alone, but should be made based on data from a research intelligence provider. Similarly, collaborations should no longer be based on the scientists' assessments and expertise, but should be pinpointed according to criteria of effectiveness and profitability.

Those who keep these reflections going inevitably and very quickly come up with very bizarre and unattractive ideas of designer babies, inbreeding, degeneration and sterility, as well as wild marriages between research teams that are not authorized by science partner agencies and that produce project bastards. So hefty, so gloomy: A science whose logging and control are constantly advancing and whose development is increasingly planned through calculated cooperation and selections, in the end will be perfected to death. Finally, evolution teaches us that every improvement and every selection advantage is preceded by a deviation from the genetic plan, a mutation, a violation of rules, in short: a mistake. These are all phenomena that a sophisticated research intelligence promises to eliminate.

References

European Commission. (2016). *Open innovation, open science, open to the world - a vision for Europe*.

https://doi.org/10.2777/061652 [accessed 12.06.2019]

Klein, S., & Hüllmann, J. (2018). Datenkapitalismus akademischer Wissenschaftsverlage.

Wirtschaftsdienst, *97*(7), 477–480.

https://archiv.wirtschaftsdienst.eu/jahr/2018/7/datenkapitalismus-eine-oekonomische-

betrachtung/#res5 [accessed 12.06.2019]

Tennant, J. (2018, Juli 6). Response to President Paul Hofheinz of The Lisbon Council regarding

Elsevier and the Open Science Monitor. http://fossilsandshit.com/response-to-president-

paul-hofheinz-of-the-lisbon-council-regarding-elsevier-and-the-open-science-monitor/

[accessed 12.06.2019]

---

[7] https://www.elsevier.com/solutions/scival/features [accessed 12.06.2019]