

Schwerpunkt Big Data

Ulrich Herb*

Zwangsehen und Bastarde

Wohin steuert Big Data die Wissenschaft?

<https://doi.org/10.1515/iwp-2018-0021>

Zusammenfassung: Taucht das Wort „Wissenschaft“ im Zusammenhang mit Big Data auf, denkt man zumeist an Wissenschaftler, die durch fundige Analysen treffsichere Prognosen über zukünftige Entwicklungen, Ereignisse und Handlungen fällen – oder aber mit chirurgischer Präzision Verhalten beeinflussen, wie etwa vom erfolgreichen Wahlkampf Donald Trumps kolportiert. Allerdings sind Wissenschaftler längst nicht mehr nur Agenten in Big-Data-Szenarien, sondern auch Objekte. Die Vermessung der Wissenschaft erfolgt heute nicht allein anhand simpler Zitationszählungen oder erwirtschafteter Projektgelder: Kommerzielle Akteure schicken sich an, ein ganzes Betriebssystem inklusive integrierter Werkzeuge zu schaffen, das unablässig Verhalten von Wissenschaftlern und Performanz von Wissenschaft protokolliert. Der Beitrag illustriert am Wandel des früheren Verlagshauses Elsevier zum Research Intelligence Provider exemplarisch, wie Wissenschaft selbst in den Fokus von Big Data gerät und wie eine Big-Data-gesteuerte Wissenschaft aussehen könnte.

Deskriptoren: Wissenschaft, Wissenschaftliches Arbeiten, Benchmarking, Bibliometrie, Forschungsplanung, Big Data

Forced marriages and bastards. Where does Big Data drive science?

Abstract: When the word “science” comes up in the context of big data, one usually thinks of scientists who make accurate predictions about future developments, events and actions by means of sophisticated analyses – or who influence behaviour with surgical precision, such as it is conveyed from Donald Trump's successful election campaign. However, scientists are no longer just agents in big data scenarios, but also objects. The measurement of science today is not only based on simple citation counts or obtained project funds: Commercial actors are creating an

entire operating system with integrated tools that continuously record the behaviour of scientists and the performance of science. By describing the transformation of the former publishing house Elsevier to a research intelligence provider this article illustrates how science itself gets into focus of Big Data and what a Big Data driven science could look like.

Descriptors: Science, Scientific work, Benchmarking, Bibliometrics, Research planning, Big Data

Des mariages forcés et des bâtards. Dans quelle direction le Big Data oriente-t-il la science?

Résumé: Si on évoque le mot « science » dans le contexte du Big Data, on pense le plus souvent à des scientifiques qui font des prévisions précises sur des évolutions, des événements et des actions futurs grâce à des analyses inventives – ou bien qui influencent avec une précision chirurgicale notre comportement, comme colporté de la campagne réussie de Donald Trump. Cependant, les scientifiques ne sont depuis longtemps plus seulement des agents dans les scénarios du Big Data, mais aussi des objets. Le mesurage de performance de la science ne repose pas seulement sur de simples recensements de citations ou sur les montants des fonds récoltés dans des projets : des acteurs commerciaux commencent à créer un système d'exploitation complet incluant des outils intégrés qui enregistrent en permanence le comportement des scientifiques et la performance de la science. En se basant sur la transformation de l'ancienne maison d'édition Elsevier en un fournisseur de renseignements de la recherche, l'article illustre comment la science elle-même est observée à l'aide de Big Data et à quoi pourrait ressembler une science dirigée par le Big Data.

Describeurs: Science, Travail scientifique, Benchmarking, Bibliométrie, Planification de la recherche, Grandes données

*Kontaktperson: Dr. Ulrich Herb, Saarländische Universitäts- und Landesbibliothek, Postfach 15 11 41, 66041 Saarbrücken, E-Mail: u.herb@sulb.uni-saarland.de

Wissenschaftsbewertung in alter Manier

Die Bewertung von Wissenschaft erfolgt anhand von Event Logging, worunter der Autor im Folgenden zwei Arten von Informationssammlung versteht:

- a) Die Protokollierung der (digitalen) Spuren, die Personen oder Werke/ Informationen, die von ihnen geschaffen wurden, hinterlassen.
- b) Die Protokollierung der Referenzen/ Reaktionen auf diese Personen bzw. deren Werke/ Informationen.

Dass Wissenschaftler durch Event Logging bewertet werden, ist nichts Neues: Schon immer zog man bibliometrische Informationen heran, wenn es um die Einschätzung der Leistung eines Wissenschaftlers ging. Sei es, dass diese durch eigenes Verhalten (im Sinne von a) produziert wurden (z. B. wenn die Zahl der Publikationen erhoben wurde), sei es, dass diese durch reaktives Verhalten anderer (b) produziert wurden (z. B. durch Zitationen auf Publikationen). Daher sind die im Folgenden dargestellten Neigungen zur Steuerung von Wissenschaft nichts völlig Unbekanntes, neu hingegen sind die Möglichkeiten, Informationen zu produzierten und genutzten Informationen über eine Vielzahl von Datentypen und -quellen zu sammeln und zu kombinieren, so dass produzierte Texte eine nur noch untergeordnete Rolle spielen.

Vom Verlag zum Research Intelligence Provider

Zur Illustration dieses Wandels bietet sich die Beschreibung der Metamorphose des Verlags Elsevier zum Research Intelligence Dienstleister (Elsevier, 2018e) an, spiegelt dessen strategische Umgestaltung die neuen Möglichkeiten der Protokollierung und Steuerung von Wissenschaft doch nahezu idealtypisch wider.

Verlagsgeschäft (1880)

Der Informationsanbieter Elsevier wurde im Jahr 1880 als Verlag gegründet, das Geschäftsmodell bestand im Verkauf und Vertrieb wissenschaftlicher Publikationen. Im Februar 2018 legt man 2.971 wissenschaftliche Zeitschriften auf. Laut Outsell Report war Elsevier 2013/2014 mit 2.525 Journalen der Verlag, der nach Springer (inkl. Bio-Med Central, BMC) die zweitmeisten Zeitschriften publi-

zierte (Outsell, 2015, S. 16), Springer und BMC kamen zusammen auf 2.984 Journale.

Geschäftsfeld Bibliometrie (2004)

Spätestens seit 2004 ist Elsevier nicht mehr nur ein Verlag, denn in diesem Jahr brachte man die Datenbank *Scopus* auf den Markt. Scopus ist eine Recherche- und Zitationsdatenbank und folglich ist für Herausgeber und Verlage eine Indexierung ihrer Journale in Scopus attraktiv: Allein schon die Aufnahme einer Zeitschrift in eine Recherche-Datenbank wird als Gütekriterium wahrgenommen. Dies gilt noch viel mehr, wenn diese Datenbank Auskunft über den Impact, die meist – und umstrittener Weise – als Qualität interpretierte quantitative und durch Zitationszählungen ermittelte Resonanz einer Publikation gibt. Aus der Perspektive des Scopus-Anbieters bedeutet dies, dass fremde Verlage ihm Content geradezu zur Auswertung aufdrängen und Elsevier selbst Daten über das Geflecht wissenschaftlicher Informationen nicht nur als Zitationsdatenbank teuer verkaufen, sondern auch für eigene Zwecke auswerten kann. Mit Scopus erschuf Elsevier, neben der eigenen Verlagsdatenbank *ScienceDirect*, einen zweiten Datenpool, gefüllt mit Informationen fremder Anbieter, gespeickt mit Daten über die Kommunikation in wissenschaftlichen Journalen und mit Meta-Informationen dazu wer (als Person oder Organisation) wieviel, mit wem, worüber publiziert und wer wie häufig von wem zitiert wird.

Geschäftsfeld Benchmarking (2009)

2009 schuf Elsevier mit *SciVal* ein weiteres, ganz neuartiges Angebot, das weder der Recherche nach Publikationen noch nach Zitationszahlen dient, sondern vielmehr dem Benchmarking von Wissenschaftseinrichtungen. Wir werden später auf diesen Dienst zurückkommen.

Geschäftsfeld Forschungsinformation (2012)

Die Entwicklung weg vom traditionellen Verlagsgeschäft wurde immer offensichtlicher als man 2012 den Anbieter Atira und dessen Forschungsinformationssystem *PURE* erwarb. Forschungsinformationssysteme dienen nach außen der Darstellung der Leistungsfähigkeit einer Einrichtung, intern zu Benchmarking und Protokollierung der Forschungsleistung dieser. Unter anderem werden solche Systeme üblicher Weise mit Projektdaten, Finanzdaten, Publikationsinformation und Impactziffern gespeist.

Geschäftsfeld Reference Management (2013)

2013 hielt die Shopping Tour Elseviers weiter an und man erwarb die Literaturverwaltung *Mendeley*. Mendeley dient, vergleichbar den Tools Endnote, Zotero, Citavi, Refworks, der Verwaltung wissenschaftlicher Literatur und hat darüber hinaus eine Social-Network-Komponente zur Kontaktpflege mit Fachkollegen. Literaturverwaltungswerkzeuge erleichtern die Textproduktion und erlauben das bequeme Einfügen von Referenzen in Publikationen. Ob Mendeley dem wirtschaftlich gut aufgestellten Anbieter Elsevier wirklich nennenswerte Einnahmen beschert, mag fraglich erscheinen: Die kostenfreie Endnutzer-Version bietet kaum beachtenswerte Unterschiede zur kostenpflichtigen Variante und auch die Campus-Lizenzen dürften finanziell nicht sonderlich ins Gewicht fallen. Interessanter könnte Mendeley als Datenquelle sein, denn daraus lassen sich Informationen über die Relevanz publizierter Texte gewinnen, die noch zu jung sind, als dass sie hätten zitiert werden können. Weiterhin lässt sich die Relevanz von Texten erfassen, die zwar von großem Interesse zu sein scheinen, da viele Wissenschaftler sie sich in Mendeley merken, die jedoch nicht oder selten zitiert werden (z. B. weil sie eher für die Lehre oder Methodik als die Forschung relevant sind) oder deren Zitationen von klassischen Zitationsdatenbanken nicht erfasst werden, weil die zitierenden Quellen nicht von diesen indexiert sind. Die Möglichkeit, Dokumente in Mendeley als Favorit zu markieren, ermöglicht sogar eine noch exaktere Bewertung deren Inhalts. Besser noch: Es ließen sich gar Informationen über *zukünftige* Forschung gewinnen, man müsste dazu nur auslesen, welche Themen in den Bibliotheken der Literaturverwaltungsnutzer gerade Konjunktur haben. Da Mendeley Nutzern auch Online-Profile bietet, sind diese geneigt hier akademischen Status (z. B. Student, Promovierender, Post Doc, Professor), Affiliation und fachliche Zuordnung zu hinterlegen, was noch treffsichere Einschätzungen zu thematischen Moden ermöglicht. Noch feinere Daten fallen an, wenn Mendeley-Nutzer ihren Account mit ihrer Auto-ID in Scopus verknüpfen.

Geschäftsfeld Medienmonitoring (2015)

2015 kaufte Elsevier den Service *Newsflo*, der die Wirkung von Wissenschaft, Publikationen und Forschungsergebnissen in Medienberichten analysiert. Der Dienst wertet laut Website (Elsevier, 2018d) 45.000 News Outlets in über 20 Ländern aus (darunter die USA, Indien, China, Brasilien und mehrere europäische Länder) und wirbt mit einer hochpräzisen Suche nach Autoren-IDs und Affiliation

durch eine Verbindung mit Scopus. Genauso betont man die Verknüpfung mit:

- Mendeley zur Bestimmung des gesellschaftlichen Einflusses von Forschung,
- SciVal zur Visualisierung der Resonanz einer Einrichtung und ihrer Forscher sowie
- PURE, für das ein Media Showcase mit Live-Feeds existiert, um aggregierte Daten der Presseberichterstattung über Forscher einer Institution zu erhalten und anzuzeigen.

Geschäftsfeld Forschungsdaten-Management (2015)

2015 führte Elsevier schließlich den Forschungsdaten-Dienst *Mendeley Data* ein. Mendeley Data stellt kostenlosen Speicherplatz zum Ablegen von Forschungsdaten bereit. Naheliegender Weise nutzen manche Elsevier-Journale oder von Elsevier aufgelegte Society Journals (z. B. CELL Press) Mendeley Data, um Artikeln zugrundeliegende Forschungsdaten verfügbar zu machen. Solche mit Publikationen verbundene Datensätze werden auch in ScienceDirect mit dem entsprechenden Artikel nachgewiesen und verlinkt.

Geschäftsfeld Disziplinäre Open Access Repositories (2016)

2016 erstand Elsevier den Open-Access-Dienst *Social Science Research Network (SSRN)*, eine Publikationsplattform für die Sozial- und Geisteswissenschaften, auf der vor allem Preprints und Discussion Papers erscheinen. Der Clou: Man hielt eine Technik in den Händen, die sich auf andere Fachbereiche ausrollen lassen sollte (und die sich perspektivisch als Plattform vermarkten lassen konnte) und hatte Zugriff auf Informationen aus noch nicht formal publizierter Forschung, die also z. B. noch nicht in Scopus oder Mendeley hinterlegt war. Noch weniger als bei Mendeley dürften direkt an SSRN gebundene finanzielle Erwartungen für dessen Kauf gesprochen haben, denn das Publizieren auf SSRN ist nach wie vor entgeltfrei möglich. Das Ausrollen der SSRN-Technik ließ auch nicht lange auf sich warten, noch 2016 startete Elsevier zwei Publikationsdienste für Working Papers aus den Naturwissenschaften, ChemRN und BioRN.

Geschäftsfeld Elektronische Labordaten (2016)

Ebenfalls 2016 erwarb man einen Dienst, der in erster Linie Natur- und Lebenswissenschaftlern nützlich ist: *Hivebench*. Die Software erlaubt die Verwaltung von Experimenten, Protokollen, Analysen und Daten. Man wirbt mit den Möglichkeiten der umfassenden, konsistenten und strukturierten Datenerfassung mittels *Hivebench*, das eine einfache und sichere Möglichkeit zur Verwaltung und Aufbewahrung von Protokollen und Forschungsdaten bietet (Elsevier, 2018b).

Als Anreiz zur Nutzung von *Hivebench* in Kombination mit anderen Elsevier-Diensten sprechen Komfort und einfache Datenhaltung, exemplarisch werden auf der Website die Integration von *Hivebench* und *Mendeley Data* empfohlen, u. a. zur Sicherung der Langzeitarchivierung von Forschungsdaten. Ebenfalls regt man 2018 die integrierte Nutzung von *Hivebench* und einem institutionellen Open Access Repository an (Elsevier, 2018b). Eine 2016 im Elsevier-Portfolio noch klaffende Lücke, die zu schließen man sich bereits anschickte.

Geschäftsfeld Institutionelle Open Access Repositories (2017)

2017 bediente sich Elsevier nochmals im Open-Access-Segment und übernahm den Repository-Anbieter *Bepress*. Anders als SSRN, BioRN oder ChemRN, die zentrale fachliche Publikationsserver sind und nicht als Hosting-Lösungen angeboten werden, ist *Bepress* eine Software, die Forschungseinrichtungen selbst betreiben, um Wissenschaftlern eine Plattform zur Open-Access-Publikation zu bieten. Elsevier (2017a) bewirbt *Bepress* mit den bekannten Lockmitteln *Showcase* und *Impact*: „Campuses can promote faculty and their expertise with scholarly profiles and expert galleries. And administrators can use industry-leading readership analytics to track impact and share it with stakeholders.“ Die Pressemitteilung Elseviers (2017b) anlässlich des Erwerbs von *Bepress* zitiert *Bepress* CEO Jean-Gabriel Bankier mit den Worten „Now with Elsevier we'll be stronger and better by applying more technologies and data and analytics capabilities to help more institutions achieve their research goals.“ Ein Kommentar, der vermitteln soll, wie unerlässlich Datenanalysen für eine Wissenschaftseinrichtung sind, die nicht scheitern, sondern ihre Ziele erreichen will. Und auch das Datenvolumen, die Ressource für weitere, noch bessere Analysen, auf die Elsevier nun zugreifen konnte, wird mit „over 2 million articles and 100 million annual downloads“ beziffert (Elsevier, 2017b).

Geschäftsfeld Altmetrics (2017)

Im Februar 2017 akquirierte Elsevier schließlich den Altmetrics-Anbieter *PLUM Analytics*, dessen Service *PLUM X*, wie andere Altmetrics-Dienste auch, Impact wissenschaftlicher Objekte zu erfassen sucht, der sich *nicht (nur)* in Zitationen manifestiert, sondern z. B. in der Anzahl der Tweets, der Likes, des Vorkommens in *Mendeley-Bibliotheken*. *PLUM X* unterscheidet sich von vergleichbaren Angeboten wie *Impactstory* oder *Altmetrics.com* durch den offenkundig kompetitiven Anspruch und seine Benchmarking-Funktionalitäten (Herb, 2016, S. 398) und bringt eine Fülle an Informationen über die Verbreitung wissenschaftlicher Objekte in Social Media, Nachrichten, Literaturverwaltungssystemen und einer Vielzahl anderer Kommunikationskanäle wissenschaftlicher wie nicht-wissenschaftlicher Natur mit sich. Die Liste der Objektarten (oder in der *PLUM*-Terminologie: Artefakte), die *PLUM X* trackt¹, umfasst derzeit 67 Typen, unter anderem auch Audio-Dateien, Blogs, Broschüren, Software Code, Datensätze, Designs, Expertenmeinungen, Regierungsdokumente² (*PLUM Analytics*, 2018).

Vom Verlag zum Betriebssystem

Allein diese, nicht abschließende Liste der Angebote Elseviers³ zeigt: Elsevier hat ein Betriebssystem mit hoch-integrierten Diensten erschaffen, das ein Wissenschaftler in jeder Arbeitsphase nutzen kann. Man tauscht sich mit seinen Kontakten in *Mendeley* aus, dort sammelt und verwaltet man zugleich seine Literatur. Preprints oder andere Dokumente stellt man in SSRN-ähnliche Repositories ein, die (wie bei SSRN teils der Fall) mit einschlägigen Journalen gekoppelt sind⁴, oder auf einen lokalen *Bepress*-Server. Artikel erscheinen in Elsevier-Journalen, die in *ScienceDi-*

1 Tracken kann hier bedeuten, dass a) für diese Objekte (z. B. Code) Impact-Werte ermittelt werden, dass b) für diese und aus diesen Objekten (z. B. Artikel) Impact-Werte ermittelt werden oder dass c) aus diesen Objekten (z. B. Nachrichtenmeldungen) Impact-Werte ermittelt werden.

2 Hier zeigt sich ein spezieller Wert für Thinktanks.

3 Eine umfassende, wenn auch nicht vollständige Liste findet sich unter: <https://www.elsevier.com/solutions> [27.2.2018]. Unter anderem fehlt an dieser Stelle *PLUM X*.

4 Hier, wie auch anderen Stellen böten sich förmlich strategische Auswertungen an, von denen man nicht in jedem Fall sicher sagen kann, dass sie angewandt werden: Erreichen die Preprints hohe Downloadzahlen oder Altmetrics-Werte, könnten diese für ein einschlägiges Journal vorgeschlagen werden. Schließlich deuten hohe Download- und Altmetrics-Werte oft auf hohe Zitationszahlen hin (Herb, 2015, S. 207–223).

rect nachgewiesen werden und deren Zitationen in Scopus erfasst sind. Wird eine wissenschaftliche Leistung dann bewertet, gehen darin Informationen aus Scopus und PLUM X ein. Sogar Herausgeber eines nicht bei Elsevier erscheinenden Journals bemühen sich darum, dass dieses in Scopus indexiert oder via PLUM X ausgewertet wird. Forschungsdaten verwaltet man mit Hivebench und publiziert sie mit Mendeley Data, ihr Impact wird mittels PLUM X erfasst. Literatur, Forschungsdaten und Forschungsinformationen verwalten, teilen, publizieren, indexieren, analysieren, bewerten und jede dieser Tätigkeiten verkaufen – das alles bietet Elsevier. Die Resonanz und Bewertung der Forschung innerhalb der Wissenschaft messen Scopus und PLUM X, außerhalb der Wissenschaft erledigt dies Newsflo. Allerdings, man ahnt es, *muss* man das System nicht nur nicht verlassen, man *kann* es nicht umgehen und seine Apps nicht ignorieren.

Googleisierung?

Wer nun angesichts Elseviers beeindruckender Datensammlung auf den Vergleich mit Google verfällt, sei beunruhigt. Jedermann kann Googles Sammeleifer (durch Verlust auf Komfort) einigermaßen entgehen und alternative, starken Wert auf Privatsphäre setzende Suchmaschinen wie DuckDuckgo nutzen oder auf ein Android-Smartphone und GoogleMail verzichten. Ein Wissenschaftler kann allerdings nicht verhindern, dass ein Dienstanbieter wie Elsevier ihn ausmisst: Selbst wer nicht in Elsevier-Journalen publizieren will, wird nicht verhindern können, dass seine Artikel in Scopus indiziert werden – besser noch: Er wird, sogar als etwaiger Anhänger eines der vielen Elsevier-Boykotte, die Indizierung in Scopus begrüßen, ist diese doch allein schon ein Zeichen von Reputation. Überdies erlaubt diese Indexierung, dass der Zitationsimpact (als vermeintlicher Qualitätsindikator) für seine Publikationen ermittelt wird und die Datenbank einen Hirsch-Index für ihn ermitteln kann. Und sowieso: Sind die Publikationen erst einmal in der Welt, werden sie in einem der Elsevier-Dienste früher oder später wie von selbst Spuren hinterlassen: Etwa, wenn Leser sie in ihre Mendeley-Bibliotheken speichern oder sie in einer der vielen Datenquellen des Altmetrics-Dienstes PLUM X erwähnen. Und wer sich als Wissenschaftler ehemals gegen die Nutzung der Literaturverwaltung Endnote entschloss, weil er seine Daten partout nicht dem Elsevier-Konkurrenten und Endnote-Anbieter Thomson Scientific anvertrauen wollte⁵, sondern sich

für Mendeley entschied, musste sich (wie der Autor dieses Artikels) eingestehen, dass alle gutgemeinten Überlegungen und die Nutzung innovativer Software kleiner Anbieter angesichts Elseviers geschickter Akquise-Strategie irgendwann obsolet sind. Derartiger Fatalismus dürfte sich genauso unter Hochschulen verbreitet haben, die sich aus ähnlichem Dafürhalten gegen die Nutzung des Forschungsinformationssystems CONVERIS, ebenfalls eines Produktes von Thomson Scientific⁶, aussprachen und lieber auf das später von Elsevier gekaufte System PURE setzten. Schlimmer noch dürfte es die Open-Access-Community getroffen haben, die den Dienst SSRN und zahlreiche Bepress-Server bewarb und aktiv nutzte, um den nicht-kommerziellen Zugang zu wissenschaftlichen Informationen zu fördern – nur um eines Tages verduzt festzustellen, dass man jahrelang sehr eifrig Content-Akquise für Elsevier betrieben hatte.⁷

Man darf jedoch keinesfalls vergessen, was geschieht, wenn die Kapitalrendite aus den bereitgestellten Diensten hinter den Erwartungen der kommerziellen Anbieter zurückbleibt: Elsevier etwa stellte seine als „most comprehensive scientific research tool on the web“ (Elsevier, 2013) angepriesene Suchmaschine Scirus 2014 klammheimlich ein.⁸ Dies verweist auf ein leider wenig diskutiertes Problem, das an dieser Stelle nur kurz angesprochen sei: Die Ökonomisierung der Wissenschaft hat längst die Kuratierung wissenschaftlicher Informationen erreicht (Haider, 2018, S. 25 f.). So nennt der Report „Open innovation, open science, open to the world – a vision for Europe“ der European Commission (2016, S. 50) Elsevier, Springer Nature und dessen Tochter Digital Science, Google und Wikimedia als Verwalter wissenschaftlicher Informationen. Eine Vorstellung, die angesichts der Scirus-Malaise als riskant betrachtet werden könnte.

Der König ist tot, es lebe der König: Data is King

Elsevier hat offensichtlich die Zeichen der Zeit früher erkannt und besser gedeutet als andere Anbieter, wie z. B. Wiley oder Springer Nature, die dem Motto Content is King

⁶ CONVERIS wird heute von Clarivate Analytics angeboten.

⁷ Auch hier war der Autor dieses Textes betroffen, der zwar Unterzeichner des Elsevier-Boykotts *The Cost of Knowledge* ist, von dem allerdings dennoch ein Preprint auf SSRN zu finden ist, wenn auch publiziert vor dem Kauf durch Elsevier.

⁸ Mein ausdrücklicher Dank gebührt Jürgen Plieninger für den Hinweis auf diese Volte.

⁵ Endnote wird heute von Clarivate Analytics angeboten.

verhaftet scheinen, und versteht sich lange nicht mehr als Verlag, sondern als „Global Information Analytics Company“ (Elsevier, 2018a). Diese Neuausrichtung der Geschäftsstrategie erklärt sicher teilweise Elseviers bislang hartnäckige Weigerung, ein landesweites Deutschland-Konsortium mit dem Projekt DEAL einzugehen. Erklärte Dienstleistung Elseviers sind „Information analytics“ und damit längst nicht mehr der Vertrieb und Verkauf wissenschaftlicher Publikationen. Die Mission lautet nun: „Uniting content and technology to serve the science, technology and health communities“ (Elsevier, 2018c). Die systematische Aufbereitung der in den oben erwähnten Services gesammelten Daten leistet in erster Linie das Angebot SciVal. Schon jetzt erweisen sich dessen Funktionalitäten als – je nach Sichtweise – verführerisch oder gefährlich: Für das Benchmarking sowie für Empfehlungen der Ausrichtung zukünftiger Forschung können zahlreiche Parameter ausgewertet und vielfältige Aussagen getroffen werden. Elsevier macht Einrichtungen SciVal unter anderem durch folgende Funktionen schmackhaft (Elsevier, 2018f):

- „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 metrics: 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.“
- „Access comprehensive research performance summaries of any desired research entities, identify their unique research strengths and multidisciplinary Research Areas.“
- „Compare the performance of any institutions, countries, and pre-defined groups, or create your own research area and monitor progress over time.“
- „Develop collaborative partnerships: Identify and analyze existing and potential collaboration opportunities based on publication output and citation impact.“
- „Analyze research trends: Analyze the research trends of any Research Area with citation and usage data, to discover the top performers and rising stars.“

Auch die Weiterentwicklung SciVals wird skizziert (Elsevier, 2018g): „We are expanding SciVal from being a purely evaluative and analytical tool to being an integral part of your research planning process“. SciVal soll, das belegt dieses Statement (genau wie die obige Liste seiner Features,) eine Planungsinstanz mit Eigenschaften der Predict-

ve Analysis sein: „You are now able to run a complete portfolio analysis to see which Topics your institution is currently active in, and which Topics have high momentum, those therefore more likely to be well-funded. It will provide insight into which researchers are active in those Topics, which Topics your peers and competitors are active in and the related Topics of which you should be aware.“ (Elsevier, 2018g).

Na und?

Ein Angebot wie SciVal kann als Werkzeug verstanden werden, das, in Zeiten knapper Mittel und der Vorstellung von Hochschulen als Unternehmen, Mittel und Personal ökonomisch einzusetzen hilft. Diese Ökonomisierung ist hinreichender Kritik, die hier nicht rekapituliert werden soll, ausgesetzt. Dass Elsevier diese Randbedingungen nutzt, um Angebote zu entwickeln, zu kaufen und zu vermarkten, ist nicht zu kritisieren, denn die RELX-Gruppe als Mutterkonzern Elseviers *muss* als Wirtschaftsunternehmen das erklärte Ziel haben, möglichst viel Geld zu verdienen.⁹ Auch die Nutzung der bereitstehenden Daten zum Monitoring und zur Steuerung von Wissenschaft ist kein Alleinstellungsmerkmal Elseviers, denn etwas später verfiel beispielsweise auch Springer Nature auf die Idee, ähnliche Dienstleistungen anzubieten. Springer Nature kennt mit dem SciGraph ein Angebot, das SciVal ähnelt. Man beschreibt SciGraph als: „...our Linked Open Data offering which aggregates data sources from Springer Nature and key partners from the scholarly domain. The Linked Open Data platform collates information from across the research landscape, for example funders, research projects, conferences, affiliations and publications“ (Springer Nature, 2018). Auch SciGraph wird mit dirigistischen Funktionalitäten beworben: „Funders, librarians, conference organizers find optimal data for analysis and recommendation tools“ (Springer Nature, 2018). Zwar stehen die SciGraph-Daten unter der Creative-Commons-Lizenz CC-BY und als Linked Open Data bereit, die faktische Verwendbarkeit dieser RDF-Daten jedoch erfordert Einiges an Know-how, so dass die Hoheit über die Auswertung weitgehend bei Springer Nature liegen dürfte. Zugleich könnte unter dem Vorzeichen einer derart liberalen Lizenzierung die Open Science Community sogar gegebenenfalls als Beiträger für den Daten-Corpus gewonnen werden. Die

⁹ Elsevier erwirtschaftet seit Jahren regelmäßig Gewinne von zwischen 30 und 40 Prozent, im Jahr 2016 beliefen sich die Einnahmen auf 2,478 Milliarden £, der bereinigte operative Gewinn 0,913 Milliarden £ oder 36,8 Prozent (RELX Group, 2018, S. 17).

Entwicklung von kostenpflichtigen Mehrwertdiensten, wie bei SciVal, wird – so viel kann man vermuten – erst durch die Anreicherung mit nicht-offenen Daten attraktiv. Gleichviel: Es bleibt festzuhalten, dass weder Elsevier noch Springer Nature dafür verurteilt werden können, Geld mit Diensten zu verdienen, nach denen der „Mode 2“ der Hochschulentwicklung giert: Man bedient einen Markt.

Wissenschaft als Zuchtbetrieb

Was jedoch einer Reflektion bedarf, sind die Folgen dieser Art der datengetriebenen Organisation von Wissenschaft. Die aus Big-Data-Analysen gewonnenen Schlüsse und Empfehlungen haben, genauer betrachtet, beachtenswerte Implikationen und zwar gleichermaßen innerhalb wie außerhalb der Wissenschaft (Bendel, 2018): So ergeben sich, in einem System wie dem geschilderten, umfassende Fragen nach Datenschutz und Persönlichkeitsrechten. Die heute mögliche, in unvorhersehbarer Weise exakte Identifikation¹⁰ eines einzelnen Wissenschaftlers in all seinen beruflichen Schattierungen und Aktivitäten erlauben zudem eine umfassende Kontrolle, ein unaufhörliches Benchmarking und ein anhaltendes Bewerten der Forschung z. B. hinsichtlich der Berücksichtigung ihrer Anschlussfähigkeit an die (internationalen) Trends. Dies wiederum dürfte zahlreiche Reaktionen auslösen, so eine aus Befürchtung vor negativer Sanktionierung erfolgende Wahl von Forschungsthemen, Kooperations- und Projektpartnern oder Ko-Autoren und damit zur „digitale[r] Bevormundung“ sowie zur Aufgabe „informationelle[r] Autonomie“ (Bendel, 2018) führen.

Wohin aber könnte uns eine derart auf daten-basierten Empfehlungen beruhende Wissenschaftssteuerung führen? Es drängen sich zwei Metaphern auf, um solche Szenarien zu beschreiben. Eine *betriebswirtschaftliche* Metapher, durchsetzt mit dem Jargon der *Kalkulation und Effizienz*, in Gestalt berechneter Kooperationen mit prognostizierten Drittmittel-Einnahmen und empfohlenen Partnerschaften mit erwirtschafteten Renditen in Form von Impact-Raten sowie Exzellenz-Attributen. Ein Szenario, das neben den Gewinnern selbstredend Verlierer braucht, weswegen das Moment des Wettbewerbs nicht fehlen darf: „Showcasing research is critical as competition increases among institutions to secure funding and attract faculty and students“

¹⁰ RELX erwarb übrigens im Januar 2018 den Dienst Threatmetrix, dessen Spezialität die Analyse von Verbindungen zwischen Geräten, von Standorten und von anonymisierten Identitätsinformationen ist (Lunden, 2018).

verlautbarte Elsevier anlässlich der Bepress-Akquise (Elsevier, 2017b). Der Nutzen des Dienstes PLUM X wurde (vor dem Kauf durch Elsevier) folgerichtig mit Worten, die weniger an Wissenschaft als öffentliches Gut gemahnen, sondern an Wissenschaft als darwinistischen Überlebenskampf, geschildert: „Arm your researchers to compete for funding“ hieß es 2016 (nach Herb, 2016, S. 398). Auch 2018 wohnt der Verkaufsargumentation eine militärische Konnotation inne, denn Festlegungen müssen strategisch sein: „...we will show you how to get the most out of Topic Prominence in Science and utilize it in the development and execution of your research strategy“ (Elsevier, 2018g).

Andere Ausführungen auf der früheren PLUM X Website, „Determine who should apply for grants“ und „Find collaboration partners“ (Plum Analytics, 2016), verwiesen auf die zweite Metapher, die *biologische der Auswahl und Zucht*. Diese findet sich, wie erwähnt, auch in den werbenden Beschreibungen zu SciVal (Elsevier, 2018f), wenn es heißt: „Identify and analyze (...) potential collaboration opportunities based on publication output and citation impact“ bzw. „Test scenarios by modeling (...) groups of researchers to apply for a large-scale grant program“.¹¹ Bestimmte Entscheidungen, z. B. welcher Forscher sich in Kooperation mit andern Forschern um Drittmittel einer Förderorganisation bemüht, sollen nicht mehr so ohne Weiteres alleine dem Forscher selbst überlassen sein, sondern mit Hilfe der Daten eines Research Intelligence Anbieters getroffen werden. Ähnlich sollen Kooperationen nicht mehr basierend auf Einschätzungen und Expertisen der Wissenschaftler zustande kommen, sondern gezielt nach Effektivitäts- und Rentabilitätskriterien ausgerichtet werden. Eine Vorstellung, die (bislang?) eher an Zuchtprogramme („Determine who should apply for grants“ / „modeling (...) groups (...) to apply for a large-scale grant program“) oder Zwangsehen („Find collaboration partners“ / „Identify (...) potential collaboration opportunities“), denn an freie Wissenschaft denken lässt.

Wer diese Assoziationen weitertreibt, gelangt unweigerlich und sehr schnell zu sehr skurrilen bis unschönen Vorstellungen von Designer-Babys, Inzucht, Degeneration und Sterilität bis hin zu nicht von Wissenschaftspartnernermittlungen¹² bewilligten wilden Ehen zwischen Forscherteams, die man mag es kaum schreiben, Projektbastarde hervorbringen. So deftig, so düster: Denn eine Wissenschaft deren Protokollierung und Steuerung sich

¹¹ Wobei diese dirigistische Komponente der betriebswirtschaftlichen implizit zuwiderläuft.

¹² Die Partnervermittlung ElitePartner trägt bereits einen passenden Namen und könnte womöglich Dienstleistungen der Beziehungs-/Kooperationsanbahnung für Hochschulen anbieten.

immer weiter vorantreiben und deren Entwicklung durch herbeikalkulierte Kooperationen und Selektionen immer geplanter wird, perfektioniert sich am Ende womöglich zu Tode. Schließlich lehrt uns die Evolution, dass jeder Verbesserung und jedem Selektionsvorteil eine Abweichung vom genetischen *Plan*, eine Mutation, ein Regelbruch, kurzum: ein *Fehler* vorhergehen muss. Allesamt Phänomene, die eine ausgefeilte Research Intelligence zu eliminieren verspricht.

Literatur

- Bendel, O. (2018). Big Data. In Gablers Wirtschaftslexikon. Springer Gabler. <http://wirtschaftslexikon.gabler.de/Archiv/-2046774198/big-data-v6.html> [27.2.2018].
- Elsevier. (2013). Scirus Homepage. <http://web.archive.org/web/20130217135017/http://www.scirus.com/> [27.2.2018].
- Elsevier. (2017a). Bepress and Elsevier – an update. <https://www.elsevier.com/connect/bepress-and-elsevier-an-update> [27.2.2018].
- Elsevier. (2017b). Elsevier acquires bepress, a leading service provider used by academic institutions to showcase their research. <https://www.elsevier.com/about/press-releases/corporate/elsevier-acquires-bepress,-a-leading-service-provider-used-by-academic-institutions-to-showcase-their-research> [27.2.2018].
- Elsevier. (2018a). Elsevier | An Information Analytics Business | Empowering Knowledge. <https://www.elsevier.com/> [27.2.2018].
- Elsevier. (2018b). Hivebench | Electronic Laboratory Notebook (ELN) designed for biology research workflows. <https://www.elsevier.com/solutions/hivebench> [27.2.2018].
- Elsevier. (2018c). Information analytics: Uniting content and technology to serve the science, technology and health communities. <https://www.elsevier.com/connect/information-analytics-uniting-content-and-technology-to-serve-the-science-technology-and-health-communities> [27.2.2018].
- Elsevier. (2018d). Newsflo | Measures an academic's societal impact. <https://www.elsevier.com/solutions/newsflo> [27.2.2018].
- Elsevier. (2018e). Research Intelligence | Reliable Data and Information | Elsevier. <https://www.elsevier.com/research-intelligence> [27.2.2018].
- Elsevier. (2018f). SciVal Features. <https://www.elsevier.com/solutions/scival/features> [27.2.2018].
- Elsevier. (2018g). Topic Prominence in Science. <https://www.elsevier.com/solutions/scival/releases/topic-prominence-in-science> [27.2.2018].
- European Commission. (2016). Open innovation, open science, open to the world – a vision for Europe. <http://doi.org/10.2777/061652> [27.2.2018].
- Haider, J. (2018). Openness as Tool for Acceleration and Measurement: Reflections on Problem Representations Underpinning Open Access and Open Science. In U. Herb & J. Schöpfel (Hrsg.), *Open Divide? Critical Studies on Open Access* (S. 17–30). Sacramento, USA: Litwin Books. <https://lup.lub.lu.se/search/publication/070c067e-5675-455e-a4b2-81f82b6c75a7> [27.2.2018].
- Herb, U. (2015). *Open Science in der Soziologie: Eine interdisziplinäre Bestandsaufnahme zur offenen Wissenschaft und eine Untersuchung ihrer Verbreitung in der Soziologie*. Glückstadt, Germany: Verlag Werner Hülsbusch. <http://doi.org/10.5281/zenodo.31234> [27.2.2018].
- Herb, U. (2016). Altmetrics zwischen Revolution und Dienstleistung: Eine methodische und konzeptionelle Kritik. In H. Staubmann (Hrsg.), *Soziologie in Österreich – Internationale Verflechtungen*. Kongresspublikation der Österreichischen Gesellschaft für Soziologie (S. 387–410). Österreichische Gesellschaft für Soziologie ÖGS. <http://doi.org/10.15203/3122-56-7> [27.2.2018].
- Lunden, I. (2018). Relx acquires ThreatMetrix for \$817M to ramp up in risk-based authentication. TechCrunch (29.01.2018) <https://techcrunch.com/2018/01/29/relx-threatmetrix-risk-authentication-lexisnexis/> [27.2.2018].
- Outsell. (2015). Open Access 2015: Market Size, Share, Forecast, and Trends. <http://esac-initiative.org/outsell-media-business-report-on-open-access-observes-lack-of-robust-infrastructure-for-oa/> [27.2.2018].
- Plum Analytics. (2016). PlumX +Grants. <http://plumanalytics.com/products/plumx-grants/> [03.2.2016].
- PLUM Analytics. (2018). About Artifacts – Plum Analytics. <https://plumanalytics.com/learn/about-artifacts/> [27.2.2018].
- RELX Group. (2018). Annual Reports and Financial Statements 2017. <https://www.relx.com/~media/Files/R/RELX-Group/documents/reports/annual-reports/relx2017-annual-report.pdf> [27.2.2018].
- Springer Nature. (2018). SciGraph | For Researchers. <https://www.springernature.com/de/researchers/scigraph> [27.2.2018].

Anmerkung: Dieser Artikel beruht auf einer früheren Fassung, vortragen auf dem Symposium „BIG DATA – Perspektiven kritischer Sozial- und Kulturwissenschaften“, veranstaltet im April 2017 an der Johannes Kepler Universität Linz (JKU) Linz/Österreich. Veranstalter: Institut für Philosophie & Wissenschaftstheorie und Kulturinstitut ebendorf.



Dr. Ulrich Herb

Universität des Saarlandes
Saarländische Universitäts- und
Landesbibliothek
Postfach 15 11 41
66041 Saarbrücken
u.herb@sulb.uni-saarland.de
<http://www.sulb.uni-saarland.de/de/service/publikationsangebote>

Dr. Ulrich Herb ist studierter Soziologe und promovierter Informationswissenschaftler. Er ist seit 2001 an der Saarländischen Universitäts- und Landesbibliothek tätig und für die Betreuung von Drittmittelprojekten und elektronischen Publikationsangeboten zuständig. Zudem ist er freiberuflicher Wissenschaftsberater und Journalist.