Altmetrics and Grey Literature: Perspectives and Challenges

Joachim Schöpfel, Hélène Prost

Altmetrics and Grey Literature: Perspectives and Challenges

Joachim Schöpfel (corresponding author), GERiiCO Laboratory, University of Lille, France
joachim.schopfel@univ-lille3.fr

Hélène Prost, CNRS, associated member of GERiiCO Laboratory, France
helene.prost@inist .fr

Abstract

Traditional metrics largely overlook grey literature. The new altmetrics introduced in 2010 as "new, online scholarly tools (that allow) to make new filters" (Altmetrics Manifesto), can include all kinds of scholarly output which makes them interesting for grey literature. The topic of our paper is the connection between altmetrics and grey literature. Do altmetrics offer new opportunities for the development and impact of grey literature? In particular, the paper explores how altmetrics could add value to grey literature, in particular how reference managers, repositories, academic search engines and social networks can produce altmetrics of dissertations, reports, conference papers etc. We explore, too, how new altmetric tools incorporate grey literature as source for impact assessment, and if they do. The discussion analyses the potential but also the limits of the actual application of altmetrics to grey literatures and highlights the importance of unique identifiers, above all the DOI. For the moment, grey literature missed the opportunity to get on board of the new movement. However, getting grey literature into the heart of the coming mainstream adoption of altmetrics is not only essential for the future of grey literature in open science but also for academic and institutional control of research output and societal impact. This can be a special mission for academic librarians.

Introduction

Traditional metrics largely overlook grey literature. Worse, they basically disregard grey literature as irrelevant for the evaluation of research. Established metrics for individuals and organisations are journal-centric. Measuring the performance and popularity of scientists or research structures means counting the number of articles citing other articles, resulting in journal impact factors, normalized citation rates and the h-index. Even those rare studies including conference papers are limited to published proceedings[1]. Grey literature remains out of scope. The most important reason is the way these metrics are produced – they rely on bibliographic tools like the Web of Sciences (WoS) and Scopus which from the beginning on were (nearly) exclusively journal and monograph A&I services, dismissing other vectors of scientific communication outside of the academic publishing market[2].

The emergence of webometrics, i.e. the "study of the quantitative aspects of the construction and use of information resources, structures and technologies on the web drawing on bibliometric and informetric approaches" (Björneborn & Ingwersen 2004, p. 1217), change the situation. As many scholarly activities today are web-based, the field of webometrics is partially covered by scientometrics (figure 1). These new or alternative metrics are not limited to journals but apply to academic content (scholarly work) at large, insofar and as long as this content is available on the web, in particular on the social web (Galligan & Dyas-Correia 2013). They are sometimes called scholarly metrics or social media metrics, and most often defined as altmetrics.

[1] See for instance Ingwersen et al. 2014, also for similar, older studies
[2] The methodological problems to identify theses in bibliographic databases in Larivière et al. (2008) confirm the situation

The fact that these new metrics can include all kinds of scholarly output makes them interesting for grey literature. In a draft on altmetrics definitions and use cases, the National Information Standards Organization describes scholarly output as "a product created or executed by scholars and investigators in the course of their academic and/or research efforts. Scholarly output may include but is not limited to journal articles, conference proceedings, books and book chapters, reports, theses and dissertations, edited volumes, working papers, scholarly editions, oral presentations, performances, artifacts, exhibitions, online events, software and multimedia, composition, designs, online publications, and other forms of intellectual property" (NISO 2016, p.9). One part of this output clearly belongs to grey literature, especially when citable and accessible[3].

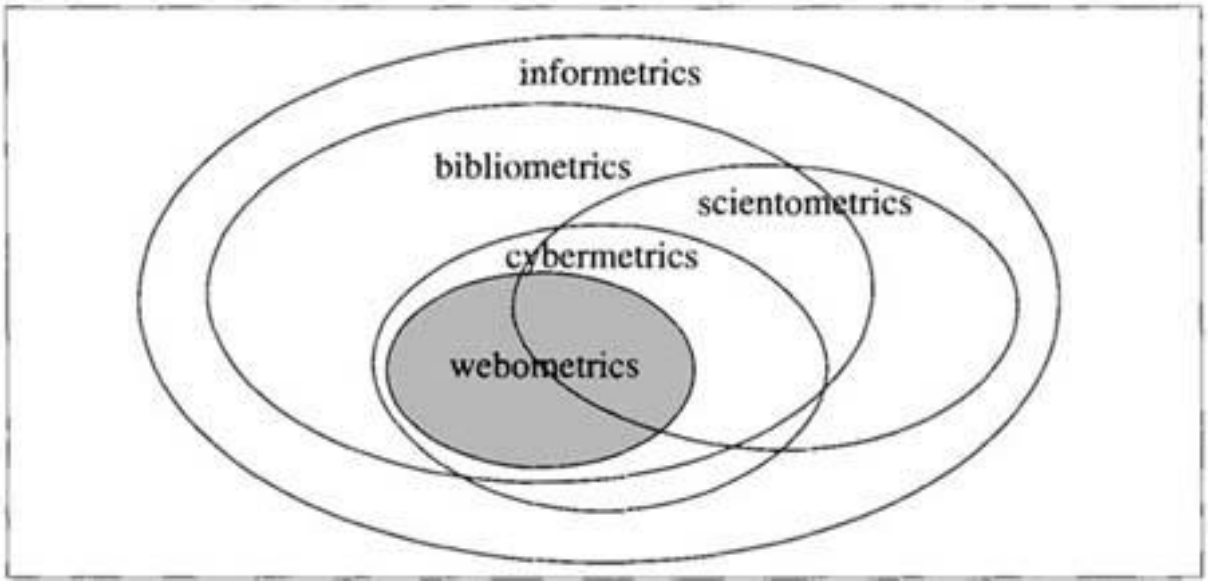

Figure 1: Webometrics in the field of library and information sciences (source: Björneborn & Ingwersen 2004)

The topic of our paper is the connection between altmetrics and conference proceedings, reports, theses and dissertations, and working papers. Do altmetrics offer new opportunities for the development and impact of grey literature? Are there already examples of good practice? Are there any barriers? However, before we outline the potential of altmetrics for grey literature, we will provide some elements for a better understanding of this concept.

A short history of altmetrics

Altmetrics have a short history[4]. The term was introduced by Jason Priem from Chapel Hill in 2010, in a tweet published on the 29th September 2010: "I like the term #articlelevelmetrics, but it fails to imply *diversity* of measures. Lately, I'm liking #altmetrics"[5]. This came after the global success of the web 2.0 tools and media, such as Facebook, Twitter etc., and it became popular as a kind of marketing umbrella for a broad range of new metrics of scholarly impact on the social web (Priem & Hemminger 2010).

[3] See the definition of "acceptable products" by the National Science Foundation, *Grant Proposal Guide II-12 NSF 14-1*, November 2013 http://www.nsf.gov/pubs/policydocs/pappguide/nsf14001/gpgprint.pdf
[4] See comprehensive reviews by Erdt et al. (2016) and Sugimoto et al. (2016)
[5] https://twitter.com/jasonpriem/status/25844968813 by @jasonpriem

The Altmetrics Manifesto[6] from 26 October 2010 merges article-level metrics and distributed scientific evaluation with social media into research on altmetrics and defines them as fast and open filters to relevant and significant scholarly sources, not in continuity but in disruption with webometrics or citations; "given the crisis facing existing filters and the rapid evolution of scholarly communication, the speed, richness, and breadth of altmetrics make them worth investing in" (Priem et al. 2010).

From that moment on, the interest for altmetrics increased steadily to join and finally exceed scientometrics, according to Google Trends (figure 2). Two years after the Manifesto, the San Francisco Declaration on Research Assessment (DORA), initiated by the American Society for Cell Biology (ASCB), recognizes the need to improve the ways in which the outputs of scientific research are evaluated and suggests the "use of a range of article metrics and indicators on personal/supporting statements, as evidence of the impact of individual published articles and other research" (DORA 2012). Signed by nearly 12,500 individuals and 800+ organizations[7], DORA fostered the awareness for altmetrics and became a reference for the debate, research and development in the field.

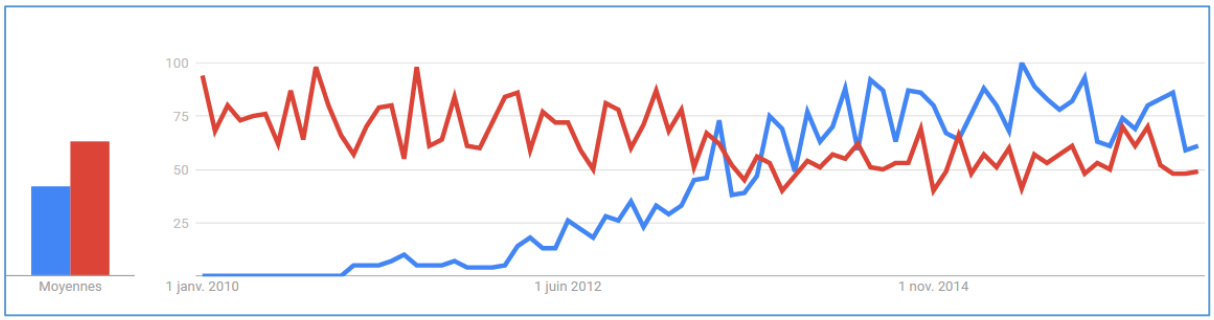

Figure 2: Altmetrics (blue) on Google Trends, compared to scientometrics (red) (2010-2016)[8]

The increasing number of scholarly work dedicated to altmetrics reveals the same trend (figure 3). No study on altmetrics before 2010, and then a steadily growth from 8 references in 2010 to 122 in 2015.

[6] http://altmetrics.org/manifesto/
[7] http://www.ascb.org/dora/ accessed 7 September 2016
[8] Data source: Google Trends www.google.com/trends accessed Sept 3, 2016

The Google Scholar statistics confirm the Google Trend figures – the interest for scientometrics remains relatively stable, with 50-70 publications per year, but is exceeded by works on altmetrics from 2013 on.

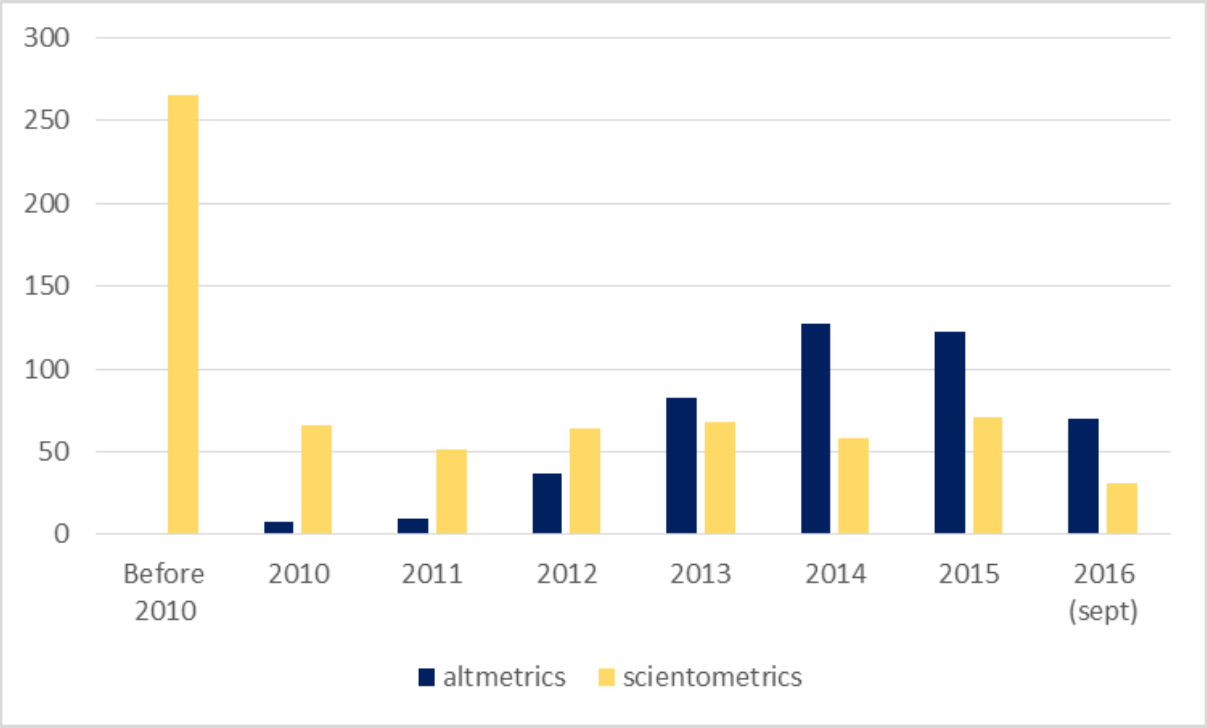

Figure 3: Publications on altmetrics and scientometrics[9]

Basically, altmetrics are "social web metrics for academic publications" (Sud & Thelwall 2014, p.1131) and particularly interesting for measuring societal impact, beyond the academic community (Piwowar 2013), through the count of views, downloads, clicks, likes, tags, posts (blogging) and tweets (micro-blogging), shares, discussions etc. The term "usually describes metrics that are alternative to the established citation counts and usage stats—and/or metrics about alternative research outputs, as opposed to journal articles" (NISO 2014, p.4).

Variety is one main feature of altmetrics, a class of indicators measuring attention, dissemination and influence[10], even if the distinction between attention, dissemination and influence is not self-evident. The main areas of altmetrics are shown in figure 4. Impact on the (social) web can be assessed through the count of PDF or HTML downloads (viewed), the creation of references in online reference managers like CiteULike, Zotero or Mendeley (saved), the number of posts in blogs and micro-blogs, on Facebook or Wikipedia (discussed), the number of mentions in editorials or tools like F1000 (recommended) or as usual, simply via the number of citations in the WoS, Scopus, PubMed Central or CrossRef (cited).

[9] Data source: Google Scholar https://scholar.google.fr allintitle: altmetrics (or scientometrics), accessed Sept 5, 2016

[10] See https://www.altmetric.com/about-altmetrics/what-are-altmetrics/

The NISO Alternative Assessment Metrics Initiative (2016) defines altmetrics as a broad concept that includes "multiple forms of assessment that are derived from activity and engagement among diverse stakeholders and scholarly outputs in the research ecosystem".

Today, a clear, common, widely accepted definition is not in sight. Altmetrics comprise many different types of metrics in a constantly changing landscape and "refer to a heterogeneous subset of scholarly metrics and are a proper subset of informetrics, scientometrics and webometrics" (Haustein 2016, p.416). Perhaps a pragmatic approach like Altmetric's recent definition will fit best, for the moment: "Altmetrics are attention data from the social web that can help librarians understand which articles, journals, books, datasets, or other scholarly outputs are being discussed, shared, recommended, saved, or otherwise used online. They can be reported at the item-, journal-, or author-level"[11].

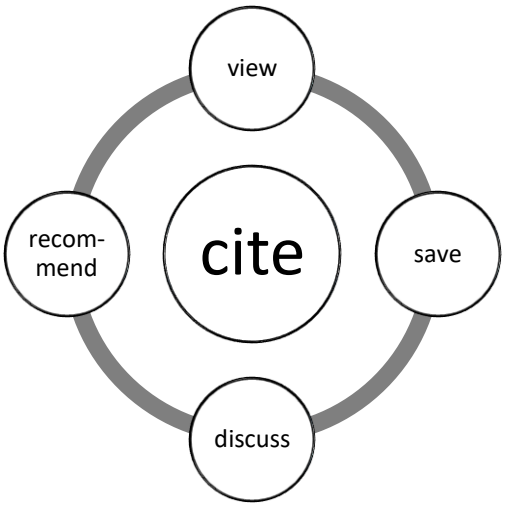

Figure 4: Altmetrics areas of assessment

Six years after the Manifesto, however, it is not quite clear if altmetrics are "an alternative or enhancement to the use of journal impact factors and click-through rate analysis to measure the impact and value of scholarly work" (Galligan & Dyas-Correia 2013, p.56). But they are already relevant for research evaluation. The European Commission DG Research and Innovation has established an Expert Group on Altmetrics which describes the emergence of altmetrics as part of the "transition to a more accountable and transparent research system"[12], more efficient, open to society, and expects "robust, responsible, transparent and interoperable uses of metrics and altmetrics in open science". Altmetrics are levers in support of open science. Up to now, including altmetrics in decisions on grants, hiring and tenure still requires careful consideration but they may soon become a normal part of a CV (Kwok 2013).

What does this mean for grey literature? What is the potential of altmetrics for grey literature? The next section tries to provide a global answer.

[11] https://www.altmetric.com/blog/altmetrics-collection-development/
[12] Next-generation altmetrics: responsible metrics and evaluation for open science, available at https://ec.europa.eu/research/openscience/index.cfm?pg=altmetrics_eg

The potential

Bornmann (2014) mentions four benefits of altmetrics compared to traditional metrics: they measure impact beyond science, they can include scholarly products other than papers (articles), they allow impact to be measured shortly after the output, and as a rule, it is easy to obtain altmetric data (figure 5).

Compared to traditional, citation-based metrics, altmetrics endorse two different developments: "Widening the definition of research outputs to include more than just books and journal articles, and looking beyond citations for a quantitative way of assessing or discovering them" (Adie 2016, p.67). Thus, at least in theory, altmetrics are not limited to a coverage similar to the WoS or Scopus. As stated by Andy Tattersall, "altmetrics focuses on research artefact level metrics that are not exclusive to traditional papers but also extend to book chapters, posters and data sets among other items" (2016, p.1). "Among other items" – this could or should bring in non-traditional, non-commercial items, like working papers, dissertations, conference papers, reports etc.

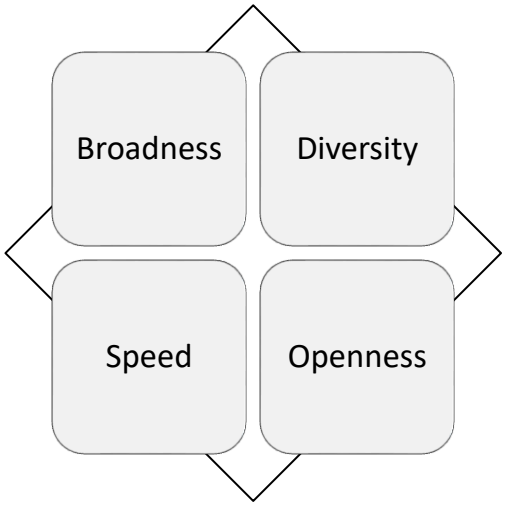

Figure 5: Benefits of altmetrics (source: Bornmann 2014)

And even if altmetrics still focus primarily on practices relating to research articles as "central research output that informs research assessment (they) can and should be extended by recognizing additional products, such as datasets (…)" (DORA 2012). Therefore, their potential for grey literature is twofold (figure 6):

- Diversity: Impact assessment on article level such as download counts also applies to grey literature. "Altmetrics (…) allow for evaluation of a greater diversity of products, i.e., not just publications (…). These products might be datasets, software, copyrights, algorithms, grey literature, and slides (...). Altmetrics now offer the opportunity to determine the impact of these products both in science (…) and beyond science" (Bornmann 2014, p.898). Diversity, as said above, is considered as one crucial advantage of altmetrics, and this includes grey literature.

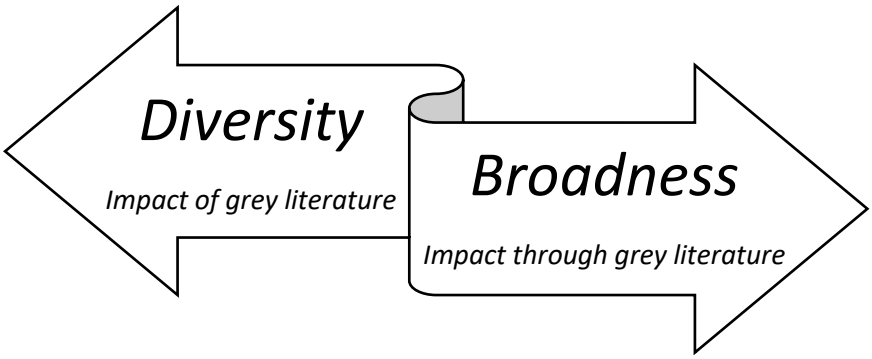

Figure 6: Double potential of altmetrics for grey literature

- Broadness: In contrast to traditional metrics which usually exclude "documents such as technical reports or professional papers which some label as 'grey literature' (…) due to lack of indexing" (Moed & Halevi 2014), altmetrics are not limited to scientometric databases; also their impact assessment based on citations, links and referrals can take account of a broader range of scientific information, including citations in dissertations, reports, white papers etc. Thus, grey literature can serve as material to measure impact of scientific output.

Beyond these two major perspectives for grey literature, recent studies on altmetrics mention two other potential benefits:

1. Dissemination of reports: Altmetrics may be a way to foster more efficient information practices by research organizations and funding bodies (foundations); one obstacle to disseminate reports etc. via open and shared systems is that often "organizations aren't sure people are even reading this stuff (…) Altmetrics (…) hopefully can better inform our expectations and measures of readership" (Brooks & Fitz 2015, p.43; see also Dinsmore et al. 2014).
2. Scientific information in developing countries: Neylon et al. (2014) insist on the application of altmetrics, especially social media usage metrics, for grey literature in "a developing country context, such as in sub-Saharan Africa (where) the importance of 'grey literature' – policy briefs, working papers, media articles and other scholarship aimed at lay audiences – is massive, satisfying both the need for social engagement as well as scholars' professional expectations" (p.2).

In the following, we will address the first two issues, diversity ("altmetrics for grey literature") and broadness ("altmetrics through grey literature"), with some examples and a focus on special conditions and prerequisites.

Altmetrics for grey literature

Our first issue is about impact assessment of grey literature. As said above, traditional metrics have largely overlooked grey literature. Altmetrics can offer new and unique opportunities for the web-

based impact measurement of reports, conferences, dissertations etc. But do they really? And if so, how? To what extent?

In 2012, a study funded by the Dutch SURF-Foundation assessed fifteen "novel impact monitors", such as reference managers, academic search engines and new altmetrics tools (Wouters & Costas 2012). At least nine out of the fifteen "monitors" can produce impact data for grey literature[13]. These seem rather favourable and promising conditions. Let's get some empirical insight for a better understanding.

Repositories

In the GreyNet community, repositories, especially institutional repositories, are generally considered as "natural home" for grey literature, as institution-based platforms for the dissemination and preservation of the institutional scientific output (Banks & de Blaaij 2006). Most of the open repositories contain one or more categories of grey literature, often theses and dissertations (particularly in university repositories), but also conference and workshop papers, unpublished reports, working papers or other "special items"[14].

All repository servers produce log files of views and downloads which can be transformed into statistics and metrics, useful as well for institutions and hosting organisations as for authors and readers. However, a couple of years ago only few repositories made these metrics freely available on their website, along with the metadata and deposited files, and even less did so in a standard, interoperable way (Schöpfel & Prost 2009, Prost et al. 2010).

The debate on new metrics accelerated the movement, and following the Altmetrics Manifesto and the DORA Declaration, repository hosts and managers started to improve the availability of web analytics and to implement new altmetrics tools. As a result, today "institutional repositories are (…) embracing altmetrics as a means of both tracking and encouraging engagement with the resources, and the ability to track and measure engagement with grey literature can be a good source of evidence of the role these outputs play in the research and publication life-cycle" (Priego 2014). Four very different examples may show the potential but also the limits of this development.

HAL[15]

The French national repository HAL contains 400,000+ documents and 1,1m records. Nearly half of the full text deposits (46%) are grey literature, with nearly 60,000 dissertations and more than 80,000 conference papers. For all these documents, HAL produces usage statistics on the item level, of full text downloads and retrievals (views) of the records (metadata). Since the launch of HAL in 2001, authors as well as collection managers have access to detailed and customizable usage statistics for each item or, cumulated, for a collection, an institution, an author etc.

On the public interface, HAL displays for each record two metrics, cumulated metadata views and full text downloads. In our example (figure 7), HAL shows that the Lille 3 White Paper on research data in PhD theses received so far 3,591 record views and 2,150 successful download requests[16]. But HAL does not offer comparative metrics (average statistics per document type and/or domain etc.).

[13] GS, MAS, ArnetMiner, Mendeley, CiteULike, Zotero, ReaderMeter, ImpactStory, SURE2
[14] See the Directory of Open Access Repositories OpenDOAR, available at http://www.opendoar.org/
[15] HAL = Hyper articles en ligne https://hal.archives-ouvertes.fr/
[16] Accessed 9 September 2016

Since 2015, HAL displays an Altmetric badge with metrics based on the unique identifiers DOI, arXiv-id and PubMed ID. So far, HAL does not allocate DOIs to deposits without identifiers and does not use its own identifier HAL Id or other identifiers like the French national dissertation number (NNT) for the assessment of altmetrics. Thus, the only conference papers with Altmetric badges we could find in HAL are those published by Springer, IEEE or other commercial publishers specialised in proceedings and members of CrossRef. Probably, this means that while all grey literature in HAL is displayed with usage statistics, no grey item has received an Altmetric badge up to now.

Figure 7: Display of usage statistics in a HAL record

figshare[17]

The online digital repository figshare where researchers can preserve and share their research outputs contains above all figures, datasets and filesets but also some papers, dissertations, posters and presentations. Figshare has "three basic functions: it acts (1) as a personal repository for yet unpublished materials, (2) as a platform for newly published research materials, and (3) as an archive for PLOS" (Kraker et al. 2015). In fact, almost 90% of the input comes from PLOS – mostly figures, while text files represent less than 2% of all entries, and the part of dissertations (all kinds of short or long unpublished written texts), posters and presentations is extremely low (0.3%).

[17] https://figshare.com/

Figshare exhibits view and download counts for all deposits. In April 2016 figshare implemented Altmetric badges to showcase attention surrounding research output (figure 8).

Figure 8: Usage statistics and altmetrics in figshare

The reader can click through to detailed impact information on the Altmetric server. All figshare entries have a DOI; 89% of the DOIs are provided by PLOS, the other 11% are allocated by figshare (with DataCite) as soon as a user makes an uploaded material publicly available. This means that dissertations may get a DataCite DOI[18]. The systematic allocation of DataCite DOIs facilitates the generation of attentions scores with the Altmetric tool. However, obviously not all figshare entries have an Altmetric badge.

IRUS-UK[19]

IRUS-UK is a Jisc-funded national aggregation service which collects raw usage data from 113 institutional repositories and transforms them into COUNTER-compliant statistics. Insofar as these repositories contain unpublished grey items, IRUS-UK aggregates usage statistics from more than 200,000 conference papers, reports, dissertations and working papers which represent 34% of the repositories' content (figure 9). All these deposits received nearly 30m successful download requests, or 43% of all aggregated IRUS-UK downloads. These statistics are interesting for three reasons.

- Often grey literature usage statistics are not standardized. Here, as the IRUS-UK statistics are COUNTER-compliant, the data are comparable, authoritative, and standard-based.

[18] See the following example, a two-page dissertation: Harper, Danny (2016): Plagiarism in college essays and assignments.docx. figshare. https://dx.doi.org/10.6084/m9.figshare.3528812.v1 Retrieved: 10 47, Sep 09, 2016 (GMT)

[19] Institutional Repository Usage Statistics UK http://www.irus.mimas.ac.uk/

- For this reason, they can be compared to download figures from other document types, in particular with article statistics. This direct comparison reveals for instance that in early September 2016, the average downloads of journal articles are similar to conference items, proceedings or reports, but two times lower than the usage statistics of dissertations and working papers. Taken together, the average usage of grey literature is one third higher than for articles or books.
- The aggregated usage statistics should allow for further standardization, e.g. for document- and/or domain-specific average download figures that could be used as a kind of reference set for individual items, like the PLOS metrics.

Document type	Total number	Total downloads
Conference or Workshop Item - Other	50 892	4 536 836
Conference Papers /Posters	7 823	389 402
Conference Proceedings	5 256	252 382
Report	14 265	1 704 870
Thesis or dissertation	123 014	21 409 166
Working Paper	5 915	795 653

Figure 9: Grey literature in IRUS-UK[20]

CORE[21]

For five years now, the CORE project aggregated and enriched content from nearly 1 000 repositories from all over the world, in order to increase the discoverability and reusability of open access papers (Pontika et al. 2016). As CORE harvests not only the repository metadata but also the full-text and caches this PDF version in its own database, it can provide IRUS usage statistics on full-text downloads. Among the 37m harvested items (called “articles” or “manuscripts”), CORE also contains reports, conference papers and other unpublished documents. However, the CORE portal does not allow for document-type specific browsing or search, a fact which reduces its interest for our purpose. -

To sum up, these four examples confirm the potential of repositories for the production of altmetrics, on a continuum from usage statistics (views, downloads) to impact measures based on social media and the possibility to display standard-based data and reference sets. The limits or pre-requisites are the need for rich metadata, including the document type, and the allocation of an established unique identifier.

Social networks

[20] All IRUS-UK figures and statistics accessed 7 September 2016
[21] COnnecting REpositories https://core.ac.uk/

Much has been said about academic social networks like ResearchGate and Academia, about their functionalities, their uptake by the research communities[22] and their impact on scientific communication[23], their competitive strategy challenging above all institutional repositories, and their business model. Because of the increasing number of users, records, documents and other material shared via these networks, they are part of these "novel impact monitors" mentioned above. At least three different aspects are relevant for altmetrics: the creation of metadata, the deposit of the document and the mention of a document in a debate or in an answer to a question.

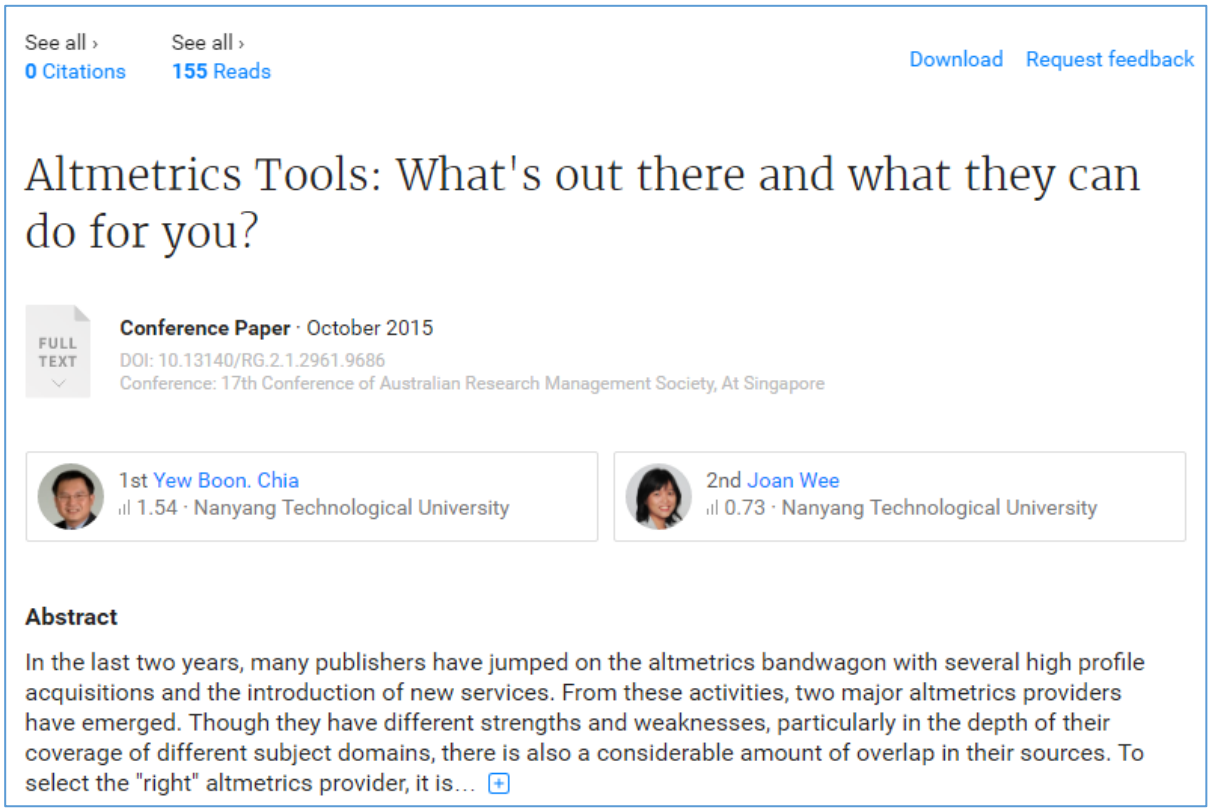

Figure 10: A shared conference paper with DOI (ResearchGate)

Basically, academic social networks invite researchers to share their results, without imposing limits or specific items, i.e. all major types of grey literature can be deposited in social networks. ResearchGate for instance suggests 18 types of "publications", including conference papers, posters, presentations, technical reports, theses and working papers; but also unpublished articles (preprints) and working copies, datasets, negative results and raw data. Also, if necessary, a new format (category) can be created for a specific deposit. Clearly they have become large reservoirs for all kinds of unpublished, grey literature, with the potential to make them available for impact measurement. However, this potential is conditioned by the quality of the metadata, in particular of an identifier. The social networks do not allocate unique identifiers but invite to add or import existing DOIs to the deposit (Figure *10).*

[22] 41m accounts in Academia, 10m accounts in ResearchGate
[23] 14.7m papers in Academia, 100m papers and other items in ResearchGate (24% papers with full-text), mostly STM

Recently, a study was conducted on the effectiveness of six ResearchGate metrics on the author level (ResearchGate score, impact points, number of downloads, number of publication views, number of citations, and number of profile views), concluding that in a small sample and a specific field "the ResearchGate score can be an effective indicator for measuring an individual researcher's performance" (Yu et al. 2016, p.1005).

Beyond academic networks, scientists share and discuss results also on Facebook and Google+; yet, on the one hand these networks are not designed for documentary metadata; on the other hand, the coverage of scientific documents still seems low, producing unreliable metrics (Haustein 2015).

Reference managers

Reference managers like CiteULike, Zotero and Mendeley can provide relevant information for altmetrics, in particular about the number of copies of a given reference. Mendeley for instance is a large database of "white papers, conference proceedings, book and journal references, and other kinds of grey literature that is searchable by other Mendeley users (…)" (Tattersall 2016, p.114). Mendeley provides how many users have a copy for each item.

Figure 11: Number of bookmarks in CiteULike

CiteULike is said to contain 8.3m references and proposes 17 item types, including conference papers, technical reports, Master's and PhD theses, unpublished work and "miscellaneous". There is no available reliable data on the actual number of references for each of these categories. Like Mendeley, CiteULike inform about the number of copies (bookmarks) for each reference (figure 11). CiteULike allocates its own identifier and supports DOI and Pubmed ID, for importing, creating (generating) and searching references.

Bookmarks can be used as a complement to citation metrics. Traditional citation-based indicators, in particular the journal impact factor and author mean citation per paper, are correlated with bookmark-based indicators (altmetrics), such as journal mean bookmarks per paper, the percentage of bookmarked articles in the journal and author mean bookmarks per paper; an analysis of data from the WoS and CiteULike reports the correlations slightly higher for journals than for authors (Sotudeh et al. 2015). Zoller et al. (2016) conclude that a bookmarking system's most inherent feature – tagging – is suitable for identifying topic subsets of publications where usage and future citations exhibit higher correlations. Yet, apparently no data has been published on bookmark-based metrics of grey literature.

Academic search engines

The SURF-study on altmetrics tools mentions the academic search engines Google Scholar and Microsoft Academic Search because they include "cited by" data for some items, whenever they can identify citations (figure 12). This data can be analysed and interpreted as an indicator for impact on the web.

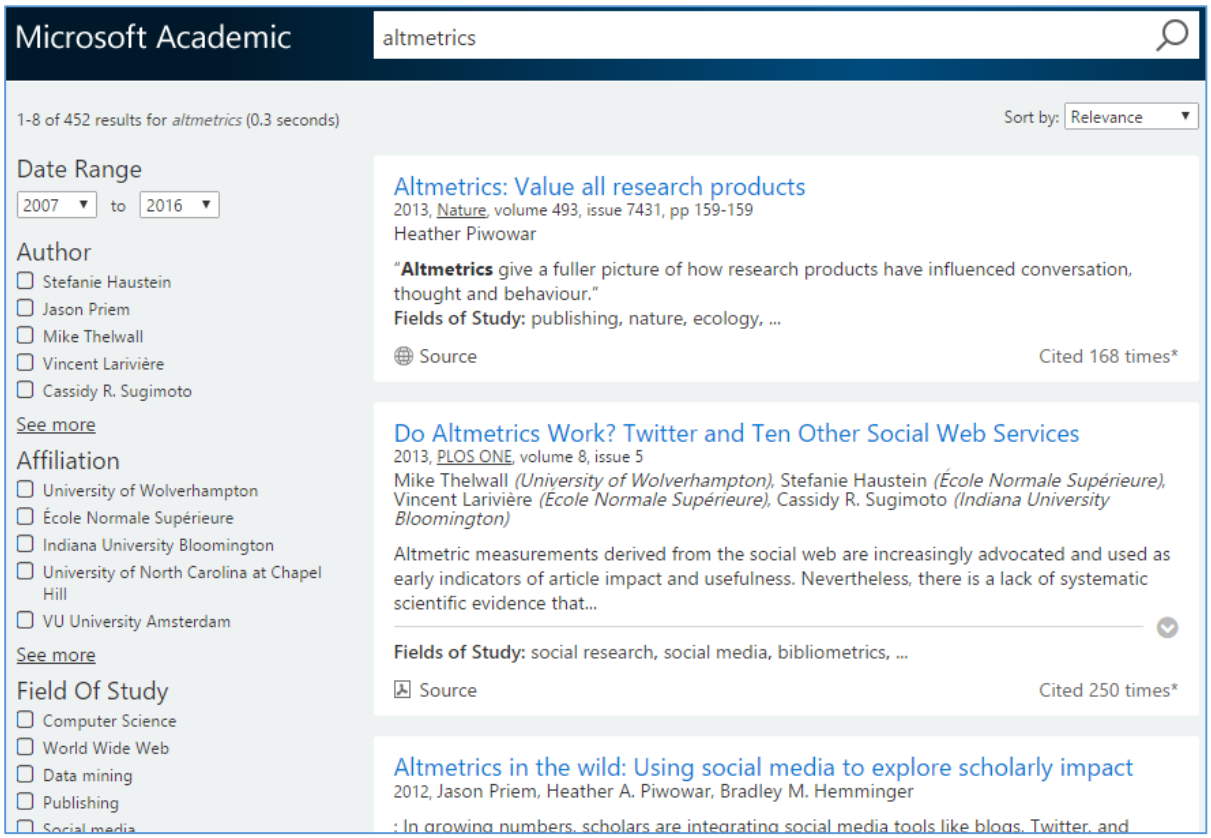

Figure 12: "Cited by" data in Microsoft Academic Search

As these search engines cover a large part of the academic web, in particular institutional repositories and other non-commercial platforms, their crawling and indexing include preprints, dissertations, reports, conference papers etc. For example, figure 13 shows citation data for a workshop paper available on figshare and not published elsewhere, without an allocated DOI.

Obviously, the academic search engine are able to produce impact data for all kinds of scientific papers, as long as they are made available on referenced and indexed platforms, in particular institutional and other repositories. Unique identifiers like the DOI are not indispensable but may improve the reliability of the search results.

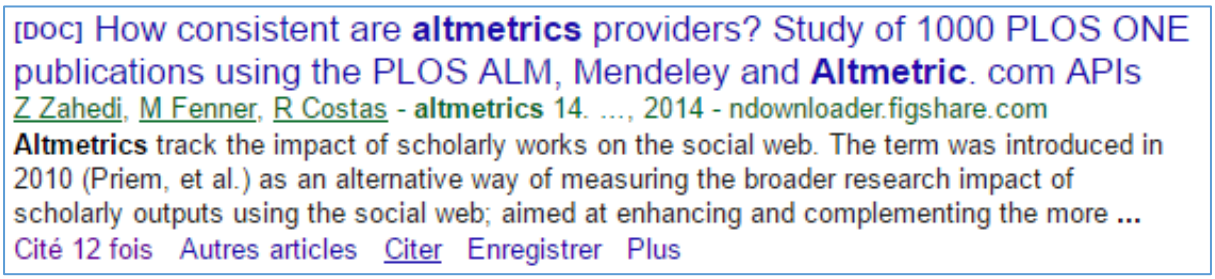

Figure 13: "Cited by" data for a workshop paper on figshare, in Google Scholar

New altmetrics tools

Following Kraker et al. (2015), the most important data provider for altmetrics are not reference managers, academic social networks or search engines but Twitter: "In the altmetrics analysis, we found that Twitter was the social media service where research data gained most attention". A growing number of new tools and platforms aggregate these online events (tweets, likes, comments, downloads etc.) as well as derived metrics from repositories, reference managers etc. (NISO 2016).

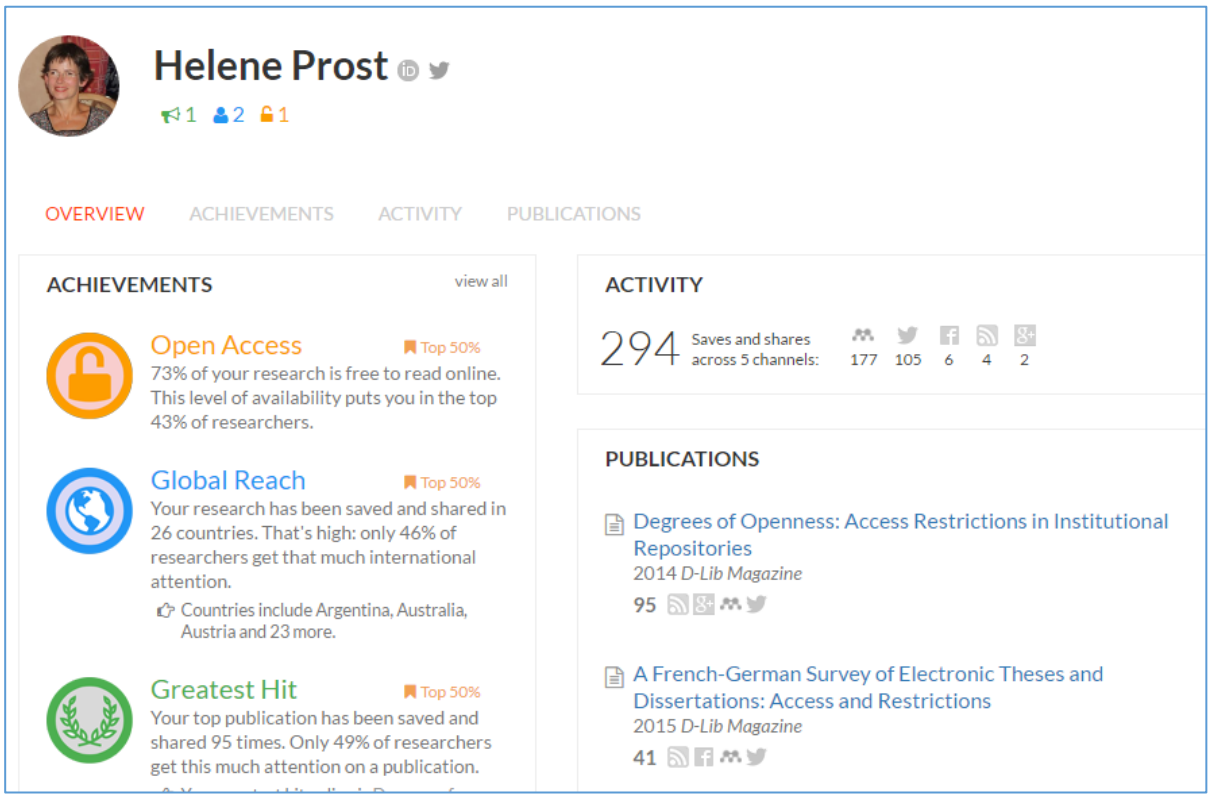

Figure 14: Author's page on ImpactStory

Among these data aggregators one can find Altmetric.com, PlumX, ImpactStory etc. (figure 14). They do not measure the same aspects, and they do not generate the same metrics. While PlumX from PlumAnalytics fits more with libraries' and institutions' needs, especially for repositories (Lindsay 2016), ImpactStory is aimed at individual researchers, and Altmetric offers services for individual researchers, institutions and funders but/and above all for commercial publishers (Konkiel 2012). PlumX detects considerably more items in social media and also finds higher altmetric scores than ImpactStory; but comparison of altmetrics tools is difficult due to differences in assignments to categories, which result in different counts (Kraker et al. 2015).

Basically, these tools can produce social impact metrics for working and conference papers, dissertations and other grey items. However, only few studies have been published on aggregated altmetrics data incorporating grey literature. Altmetric does not track non-traditional outputs. Wilkinson et al. (2014) made use of the Web Impact Report (WIRe) as a novel solution to assess the impact of organisational reports, especially when in open access. WIRe consists of a "range of web-derived statistics about the frequency and geographic location of online mentions of an organisation's reports (…)" (p. 797), such as online citations, site domain and genre of the citing site (blog, governmental sources etc.). Nevertheless, this case study with a small corpus of 20 reports reveals two major issues, i.e. a relatively high percentage of incorrect matches and a time-consuming human workload (content analysis).

ArnetMiner[24] aims to provide comprehensive search and mining services for academic social networks, with a special focus on 6,000+ conferences, mostly in computer sciences, and with a ranking based on the H5-Index, top-cited authors and papers, and data on the social network and semantics for each conference (figure 15).

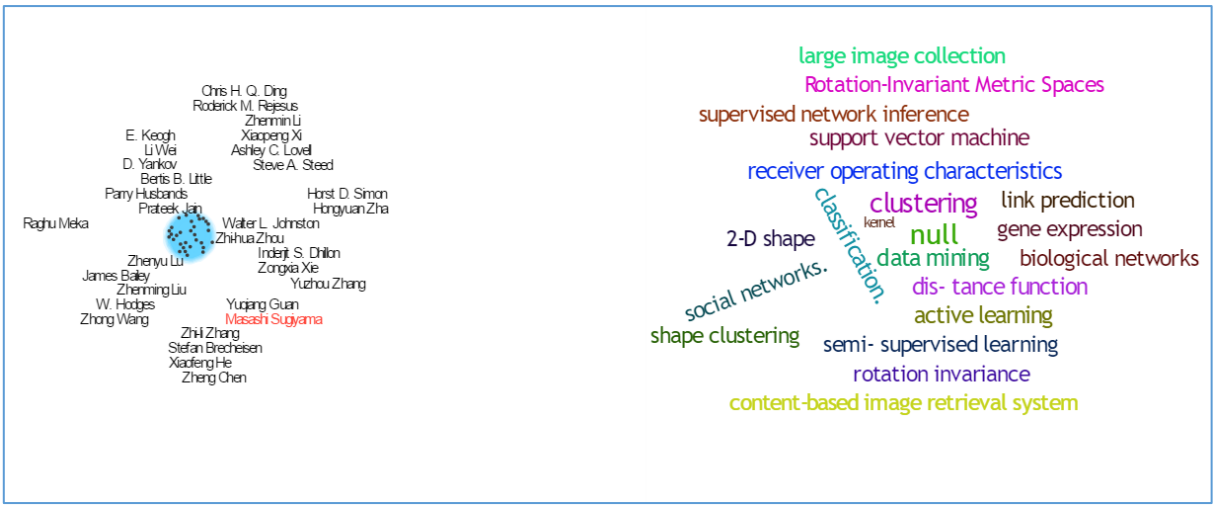

Figure 15: Conference analysis of SIAM International Conference on Data Mining on ArnetMiner

[24] http://www.aminer.org

Literature about altmetrics mentions other tools like PaperCritic, PeerEvaluation or ReaderMeter. Some are operational, others are not; some are part of PLOS, Mendeley etc. Basically they are all open for grey and other items, primary data etc. But we could not find any reliable information about their real value and interest for grey literature, in terms of specificity and impact.

Altmetrics through grey literature

"Broadness", the possibility to take account of a broader range of scientific information is one of the four major benefits of altmetrics. They could be "more diverse in kinds of data and accordingly numbers of data sources (whereas for traditional citations only the cited references in journals serve as data source)" (Bornmann 2014, p.898), thus revealing more diverse and nuanced forms of impact than traditional indicators. So, how can grey literature contribute to altmetrics? Do altmetrics tools make use of grey literature? For Euan Adie, CEO of Altmetric, grey literature[25] "presents great opportunities for alternative metrics, providing data and indicators that cannot be found anywhere else" but also drew attention to the specific characteristics and challenges, e.g. missing identifiers, no "canonical metadata", lack of long term preservation and availability (Adie 2014). Five examples may illustrate potential benefits and limits of this "broadness".

1. Web-based grey literature can serve as source to increase impact of other grey literature. Wilkinson et al. (2014) conducted their study on WIRe with a small sample of 20 research reports. Their results showed that most of them (17) had been cited by other reports, conference papers, white papers, MA and PhD theses and speeches and/or dissertations available on the web. But without standard or automated procedures, including grey literature involves a lot of human work.
2. In some fields, grey literature may be more relevant than in others. Working on subfields of sustainable energy research, Ingwersen et al. (2014) insist that such analyses "should include proceedings papers – because this document type does have significant (…) influence on the overall citation impact of a research field, in particular in proceedings-dominant fields" (p.1290). The same observation would probably apply to economics (working papers), physics (preprints) and computer science (conference papers).
3. New content mining tools improve the efficiency and broadness of data aggregators. Thus, Altmetric has developed a text-mining solution (*Altmetric Policy Miner*) to discover mentions of publications in policy documents on selected websites[26]. Due to this APM-software, Bornmann et al. (2016) were able to assess the societal impact of climate change publications mainly through grey literature from governmental agencies, international organizations and NGOs.
4. Human knowledge, manual searching and browsing are the price of "broadness" and inclusion of grey literature in alternative impact assessment. Urquhart & Dunn (2013) evaluated the impact of the National Minimum Dataset for Social Care (NMDS-SC). References to the dataset (citations) were identified in 175 separate publications, with 50% policy and practice reports, 35% media communications and only 15% academic journal articles. Google Scholar fits more with this kind of analysis than the Web of Science, because of a greater range of included material. Other relevant sources are field-specific databases,

[25] Described as "theses, posters, preprints, patents and policy documents and similar"

[26] Such as European Food Safety Authority (EFSA), GOV.UK–Policy papers, Research & Analysis, Intergovernmental Panel on Climate Change (IPCC), International Committee of the Red Cross (ICRC), World Health Organization (WHO), International Monetary Fund (IMF), Médicins sans Frontières (MSF), NICE Evidence, Oxfam Policy & Practice, UNESCO and World Bank

aggregators' and publishers' platforms[27]. The procedure of such an extensive targeted grey literature search is rather complex, far from automated and quick processing of large amounts of bibliographic information: "Organisations considered likely to be publishing materials drawing on NMDS-SC (…) were identified from existing knowledge of the sector, by the client and ourselves, and the initial findings from the bibliometric survey. A total of 24 organisational websites were manually searched and browsed, including UK government departments, sectoral bodies, knowledge intermediary organisations such as independent research organisations (…), campaigning organisations, think tanks, trade/employer organisations and the professional and mainstream press. We also conducted a limited search of social media, using social media aggregator sites" (p.297).

5. Altmetrics with grey literature produce more content but are time-consuming. Sibbald et al. (2015) conducted a case study on the inclusion of grey literature in citation analysis, based on one published article in the field of violence against women. Google Scholar and the Web of Science produced eighty journal articles citing the paper. The grey literature searches[28] found 29 other sources (27% of all results). But "this method requires additional resources. The much broader range of potential search venues demands more time and expertise. Delving into gray literature is a challenging task and requires planning and coordination, including consideration-specific inclusion/exclusion searching. Unlike database searching, common nomenclatures rarely exist for searching diverse gray literature sources; therefore, the concept of consistency in search terms across sources is difficult to achieve."

These examples confirm the potential of grey literature as source for altmetric impact assessment, with significant and complementary results based on citations, links and referrals from a broader range of scientific information, including dissertations, reports, white papers etc. But they also show that this approach is more complicated and time-consuming than the usual WoS or Scopus-based work. In contrast to traditional metrics which usually exclude grey literature, altmetrics are not limited to scientometric databases. But when it comes to larger empirical studies, this exploitation of grey literature remains an exception and is sometimes limited to science blogs, while proceedings, dissertations etc. are dismissed (see for instance Thelwall et al. 2013 or Costas et al. 2015).

Discussion

Our objective was to clarify the connection between altmetrics and grey literature. Traditional metrics have largely overlooked grey literature. Do altmetrics offer new opportunities for the development and impact of grey literature? In fact, we explored two different issues:

1. Impact assessment of grey literature – do altmetrics offer new and unique opportunities for the web-based impact measurement of reports, conferences, dissertations etc.? Do they contribute to improved visibility and impact? And if so, how?
2. Impact assessment through grey literature - how can grey literature contribute to altmetrics? Do altmetrics tools make use of grey literature?

These two issues have been described in terms of diversity and broadness, as specific benefits of altmetrics compared to traditional indicators.

[27] In this study NHS Evidence, LG Search, Hein Law Online, EBSCO Business Source Complete, Nexis and Emerald Journals

[28] Searches were conducted with Google and in Scopus, MedlinePlus, MDConsult, UpToDate, Factiva, Lexis Nexis, Google News, and Proquest Canadian Major Dailies. Major health care associations and professional organisations likely to include related content were identified, and their websites were individually searched.

Our review of recent publications, together with some altmetrics tools presents a contradictory situation:

- The potential of altmetrics for grey literature is real. Altmetric data providers like Twitter, Mendeley, Facebook or figshare but also reference managers and institutional repositories are tailored for grey literature, and they already contain significant amounts of unpublished documents.
- Assessment studies on the grey literature's web-based impact show partly higher impact than journal articles or books. Apparently, altmetrics offer a unique opportunity to exhibit the real impact of unpublished research results in conference papers, dissertations, working papers etc. and to contribute to improved visibility of these documents.
- But this work remains more or less exceptional. Most studies on F1000, Mendeley etc. include only journal articles (see for instance Mohammadi & Thelwall 2013, 2014). The main reason is that altmetrics tools need unique identifiers, standard metadata and good availability. "One of the critical issues is that these aggregators concentrate on documents that have a unique object identifier, which inevitably neglects certain document types (…) For example, Altmetric.com (…) focuses its data collection on DOIs, which has led to a de facto reduction of altmetrics studies to journal articles, excluding many types of documents and journals" (Sugimoto et al 2016).
- Impact assessment with grey literature is difficult, time-consuming and manual work, and requests expert knowledge of the scientific information landscape, especially when the grey resources are not available on open repositories but somewhere in the dark web, e.g. on less-referenced, personal or other websites.
- And then there may be other reasons to dismiss grey literature. In Hammarfelt's (2014) study research impact in the humanities, all grey items - 1,006 conference papers, dissertations and reports (20%) - were skipped from the initial corpus of 5,091 scholarly works[29] because of the "scarcity" of altmetrics data in particular for the Swedish language documents.

No identifier, lack of bibliographic control and no standard metadata, unsatisfying availability – all this is not new in the field of grey literature, and Adie's (2014) suggestions to improve the situation is only too familiar for the grey community: minimum standards for metadata (PRISM[30], DC), persistent identifiers (handle, DOI), discoverability (index, repository). His suggestion: "An open, central index of scholarly grey literature that enforced a minimum level of metadata for each item (…) An alternative would be to maintain a central index of grey literature repositories (…) and to allow harvesting from each (…)".

A central index of grey literature – this sounds like utopia. Probably the main issue is that altmetrics need DOI (Adie 2016); and the DOI appears to be the only realistic option for the assignment of permanent and citable identifiers to grey literature when it comes to prepare academic output in repositories for alternative metrics (Gerritsma 2015[31], see also Brooks & Fitz 2015). But given the history of failed initiatives for standard identifiers and metadata, we must admit that this may be just another missed opportunity.

[29] Extraction from the SwePub database of academic publications at Swedish universities http://swepub.kb.se/
[30] Publisher Requirements for Industry Standard Metadata, see http://www.idealliance.org/specifications/prism-metadata-initiative
[31] Theses, working papers, reports, conference contribution – in Gerritsma's example (VU Amsterdam) grey items represents 14% of the whole output

Perspectives

Are altmetrics the future of scientometrics? For the moment, they are still "in infancy" (Erdt et al. 2016), and for many researchers, impact factor and large citation databases are still preferred for determining impact, with 'pure' altmetrics tools scoring much lower, especially in physical sciences, engineering and technology[32]. Likewise, because of not-yet achieved critical mass, lack of theory, lack of quality control mechanisms, inconsistencies and multiplicity of social web sources, data, tools and methods, Sotudeh et al. (2015) speak of "immaturity of the field" and call for cautious application and interpretation, even as a complement to traditional metrics. The risk of misuse and rankings based on such arbitrary information is real.

Are download counts really a metric of scholarship or only of computer activity? Is popularity an indicator of quality? How does one deal with multiple versions of the same item? For these and other reasons, Booth (2016) condemns the limited validity of the new generation of altmetrics and suspects that they follow a logic of easiness to get the data; "(they are) neither a more accurate representation of academic 'quality' nor immune to critics" (p.41). In particular, the composite, "all-in-one" Altmetric Score has been critically appraised, because of lacking of transparency, reproducibility and stability, questionable validity and significance, and problems with data sources, consistency and completeness (see Gumpenberger et al. 2016).

The "pressure of various stakeholders" and the dependency on aggregators and social media as data providers may explain one part of the criticisms (Haustein 2016). Lack of transparency and conceptual deficit are at the opposite of the purpose of the Leiden Manifesto for Research Metrics (Hicks et al. 2015) but may be related to the increasing commercial take-over of these new tools and services by those who already dominate the scientific information market.

According to Gartner's famous Hype Cycle model[33], new technology go through a typical five-phase life cycle (figure 16). After a potential technology breakthrough kicks things off ("technology trigger") and a growing number of success stories ("peak of inflated expectations") comes the "trough of disillusionment", with growing criticisms, failures and dissatisfaction.

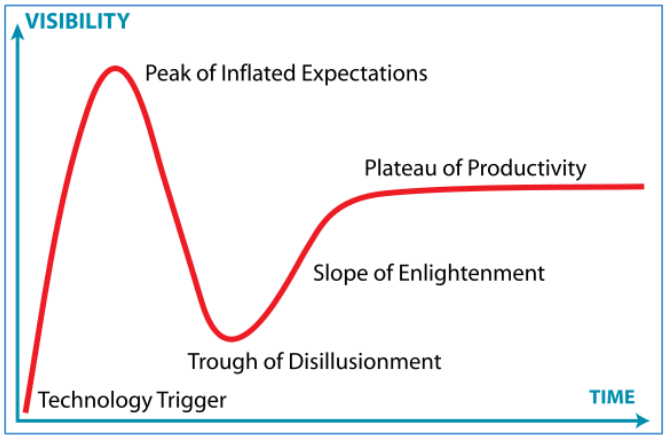

Figure 16: Gartner's Hype Cycle (source: Wikipedia[34])

[32] Innovations in Scholarly Communication Survey, http://altmetricsconference.com/who-is-using-altmetrics-tools/
[33] http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp
[34] https://en.wikipedia.org/wiki/File:Gartner_Hype_Cycle.svg

On their own technology life cycle, altmetrics probably have passed by the peak of inflated expectations and are moving forward to this "trough of disillusionment", which is a necessary and salutary transition to a more realistic and satisfying situation where this new generation of metrics is no longer considered as the one and only alternative to traditional performance assessment but as new and interesting methods to assess impact of research output, complementary to traditional metrics.

Metrics shape the science, said Paul Wouters from the Centre for Science and Technology Studies at Leiden University, and we can reasonably expect that altmetrics will be part of the game. Altmetrics are already a major topic of the European Open Science Agenda and will contribute to a new rewarding and funding system.

To come back to our initial question – what is the role of grey literature in this emerging world of new assessment tools? When second- and third-generation products will appear from technology providers and later, when mainstream adoption will take off, will grey literature be part of the game or remain out of scope, just as before? For the moment, grey literature missed the opportunity to get on board. Since the Altmetrics Manifesto 2010, no real effort has been made to adapt the new assessment tools to grey literature or to make this literature suitable for altmetrics. Publications in the field of scientometrics show that journal (and sometimes book) publishing is still at the heart of research and development, not only for traditional metrics but also for alternative metrics. For instance, most of the contributions to the last Altmetrics Conference in Bucharest[35] are about journal publishing, and the rare exceptions deal with datasets and software, not with grey literature. Today, the future development of this new technology bears the risk of dismissing large parts of scientific literature – those parts not controlled by commercial publishers. Just as before, it is business as usual. Sometimes you don't get a second chance. But you have to be at the station when the train arrives. To get to the station means to:

- Contribute to research on altmetrics for or with grey literature, for instance in the fields of economics (working papers) or computer science (conference papers).
- Cooperate with altmetrics companies and teams for the development of appropriate tools that fit with grey literature.
- Accelerate the allocation of unique identifiers for grey literature and their authors and why not their institutions, above all this means partnership with DOI, ORCID and CASRAI[36], in particular for electronic theses and dissertations and for scientific reports.
- Contribute to further standardization of grey literature metadata.
- Contribute to increasing availability of grey literature in institutional repositories.

Getting grey literature into the heart of the coming mainstream adoption of altmetrics is essential not only for the future of grey literature in open science but also for academic and institutional control of research output and societal impact. This can be a special mission for academic librarians. Grey literature has always been a library-driven concept (Schöpfel 2010); today, as a recent survey shows, academic librarians demonstrate a higher awareness for altmetrics tools than researchers[37]. Perhaps this convergence or happy coincidence may be helpful.

[35] 3:AM Conference, Bucharest 28-29th September 2016, http://altmetricsconference.com/schedule/

[36] See the Jisc CASRAI-UK pilot on organisational identifiers https://jisccasraipilot.jiscinvolve.org/wp/2015/03/06/organisational-identifiers-working-group-outputs-and-update/

[37] Innovations in Scholarly Communication Survey, http://altmetricsconference.com/who-is-using-altmetrics-tools/ See also Malone & Burke (2016).

References

Adie, E., 2014. The grey literature from an altmetrics perspective – opportunity and challenges. *Research Trends* (37).

Adie, E., 2016. The rise of altmetrics. In: Tattersall, A. (Ed.), pp. 67-82.

Banks, M. A., de Blaaij, C., 2006. Implications of copyright evolution for the future of scholarly communication and grey literature. In: GL8 *Eighth International Conference on Grey Literature*, New Orleans, 4-5 December 2006.

Björneborn, L., Ingwersen, P., 2004. Towards a basic framework of webometrics. *Journal of the American Society for Information Science and Technology* 55 (14), 1216-1227.

Booth, A., 2016. 'Metrics of the trade': where have we come from? In: Tattersall, A. (Ed.), pp. 21-47.

Bornmann, L., 2014. Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics* 8 (4), 895-903.

Bornmann, L., Haunschild, R., Marx, W., 2016. Policy documents as sources for measuring societal impact: how often is climate change research mentioned in policy-related documents? *Scientometrics,* 1-19.

Brooks, L., Fitz, G., 2015. Grey matter(s): Embracing the publisher within. *The Foundation Review* 7 (2), 38-50.

Costas, R., Zahedi, Z., Wouters, P., 2015. Do "altmetrics" correlate with citations? Extensive comparison of Altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology* 66 (10), 2003-2019.

Dinsmore, A., Allen, L., Dolby, K., 2014. Alternative perspectives on impact: The potential of ALMs and altmetrics to inform funders about research impact. *PLOS Biol* 12 (11), e1002003+.

DORA (2012). `*San Francisco Declaration on Research Assessment*'. American Society for Cell Biology (ASCB), San Francisco.

Erdt, M., Nagarajan, A., Sin, S.-C. J., Theng, Y.-L., 2016. Altmetrics: an analysis of the state-of-the-art in measuring research impact on social media. *Scientometrics* (first online 10 August 2016).

Galligan, F., Dyas-Correia, S., Mar. 2013. Altmetrics: Rethinking the way we measure. *Serials Review* 39 (1), 56-61.

Gerritsma, W., 2015. Altmetric opportunities for libraries. In: *2:AM Conference (Altmetrics Conference),* University Amsterdam, 8th November 2015.

Gumpenberger, C., Glänzel, W., Gorraiz, J., 2016. The ecstasy and the agony of the Altmetric score. *Scientometrics* 108 (2), 977-982.

Hammarfelt, B., 2014. Using altmetrics for assessing research impact in the humanities. *Scientometrics* 101 (2), 1419-1430.

Haustein, S., Costas, R., Larivière, V., 2015. Characterizing social media metrics of scholarly papers: The effect of document properties and collaboration patterns. *PLoS ONE* 10 (3), e0120495+.

Haustein, S., 2016. Grand challenges in altmetrics: heterogeneity, data quality and dependencies. *Scientometrics* 108 (1), 413-423.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., Rafols, I., 2015. Bibliometrics: The Leiden Manifesto for research metrics. *Nature* 520 (7548), 429-431.

Ingwersen, P., Larsen, B., Carlos Garcia-Zorita, J., Serrano-López, A., Sanz-Casado, E., 2014. Influence of proceedings papers on citation impact in seven sub-fields of sustainable energy research 2005–2011. *Scientometrics* 101 (2), 1273-1292.

Konkiel, S., 2012. Altmetrics: An app review. In: *OCLC Innovation in Libraries, post-conference event, LITA Forum 2012*, Columbus, OH, USA, October 8, 2012.

Kraker, P., Lex, E., Gorraiz, J., Gumpenberger, C., Peters, I., 2015. Research data explored II: the anatomy and reception of figshare. In: *STI2015 20th International Conference on Science and Technology Indicators*, Lugano, 4 September 2015.

Kwok, R., 2013. Research impact: Altmetrics make their mark. *Nature* 500 (7463), 491-493.

Larivière, V., Zuccala, A., Archambault, E., 2008. The declining scientific impact of theses: Implications for electronic thesis and dissertation repositories and graduate studies. *Scientometrics* 74 (1), 109-121.

Lindsay, J. M., 2016. PlumX from plum analytics: Not just altmetrics. *Journal of Electronic Resources in Medical Libraries* 13 (1), 8-17.

Malone, T., Burke, S., 2016. Academic librarians' knowledge of bibliometrics and altmetrics. *Evidence Based Library and Information Practice* 11 (3), 34+.

Moed, H. F., Halevi, G., 2014. The multidimensional assessment of scholarly research impact. *Journal of the Association for Information Science and Technology* 66 (10), 1988-2002.

Mohammadi, E., Thelwall, M., 2013. Assessing non-standard article impact using f1000 labels. *Scientometrics* 97 (2), 383-395.

Mohammadi, E., Thelwall, M., 2014. Mendeley readership altmetrics for the social sciences and humanities: Research evaluation and knowledge flows. *Journal of the Association for Information Science and Technology* 65 (8), 1627-1638.

Neylon, C., Willmers, M., King, T., 2014. *Rethinking impact: Applying altmetrics to Southern African research.* Report, University of Cape Town. Scholarly Communication in Africa Programme, Paper 1.

NISO, 2014. *Alternative metrics initiative. Phase 1 White Paper*. National Information Standards Organization, Baltimore MD.

NISO, 2016. *Altmetrics definitions and use cases. A recommended practice of the National Information Standards Organization*. Draft. NISO RP-25-201x-1, National Information Standards Organization, Baltimore MD.

Piwowar, H., 2013. Altmetrics: Value all research products. *Nature* 493 (7431), 159.

Pontika, N., Knoth, P., Cancellieri, M., Pearce, S., 2016. Developing infrastructure to support closer collaboration of aggregators with open repositories. *LIBER Quaterly* 25 (4).

Priego, E., 2014. *On metrics and research assessment*.

Priem, J., Taraborelli, D., Groth, P., Neylon, C., 2010. *Altmetrics: A Manifesto*.

Priem, J., Hemminger, B. H., 2010. Scientometrics 2.0: New metrics of scholarly impact on the social web. *First Monday* 15 (7).

Prost, H., Le Bescond, I., Schöpfel, J., 2010. Usage assessment of an institutional repository: a case study. In: *GL12 Twelfth International Conference on Grey Literature*, Prague, 6-7 December 2010.

Schöpfel, J., 2010. Towards a Prague definition of grey literature. In: *GL12 Twelfth International Conference on Grey Literature*. Prague, 6-7 December 2010

Schöpfel, J., Prost, H., 2009. Usage of grey literature in open archives: state of the art and empirical results. In: *GL11 Eleventh International Conference on Grey Literature*. Washington D.C., 14-15 December 2009.

Sibbald, S. L., MacGregor, J. C. D., Surmacz, M., Wathen, C. N., 2015. Into the gray: a modified approach to citation analysis to better understand research impact. *Journal of the Medical Library Association: JMLA* 103 (1), 49-54.

Sotudeh, H., Mazarei, Z., Mirzabeigi, M., 2015. CiteULike bookmarks are correlated to citations at journal and author levels in library and information science. *Scientometrics* 105 (3), 2237-2248.

Sud, P., Thelwall, M., 2014. Evaluating altmetrics. *Scientometrics* 98 (2), 1131-1143.

Sugimoto, C. R., Work, S., Larivière, V., Haustein, S., submitted. Scholarly use of social media and altmetrics: a review of the literature. *Journal of the American Society for Information Science and Technology*.

Tattersall, A. (Ed.), 2016. *Altmetrics: a practical guide for librarians, researchers and academics*. Facet Publishing, London.

Thelwall, M., Haustein, S., Larivière, V., Sugimoto, C. R., 2013. Do altmetrics work? Twitter and ten other social web services. *PLOS ONE* 8 (5), e64841+.

Urquhart, C., Dunn, S., 2013. A bibliometric approach demonstrates the impact of a social care data set on research and policy. *Health Information & Libraries Journal* 30 (4), 294-302.

Wilkinson, D., Sud, P., Thelwall, M., 2014. Substance without citation: evaluating the online impact of grey literature. *Scientometrics* 98 (2), 797-806.

Wouters, P., Costas, R., 2012. Users, narcissism and Control—Tracking the impact of scholarly publications in the 21st century. In: *STI2012 17th International Conference on Science and Technology Indicators (STI)*, 5-8 September, 2012, Montreal, Quebec, Canada.

Yu, M.-C., Wu, Y.-C. J., Alhalabi, W., Kao, H.-Y., Wu, W.-H., 2016. ResearchGate: An effective altmetric indicator for active researchers? *Computers in Human Behavior* 55, 1001-1006.

Zoller, D., Doerfel, S., Jäschke, R., Stumme, G., Hotho, A., 2016. Posted, visited, exported: Altmetrics in the social tagging system BibSonomy. *Journal of Informetrics* 10 (3), 732-749.

The complete bibliography with links and further readings is available at http://www.citeulike.org/user/Schopfel/tag/gl18

All websites visited in August and September 2016.

Acknowledgment to Paul Needham, JISC, for helpful advice with the IRUS-UK project.