

Coombs, Sarah K.; Peters, Isabella

Article

The Leiden Manifesto Under Review: What Libraries Can Learn From It

Digital Library Perspectives

Suggested Citation: Coombs, Sarah K.; Peters, Isabella (2017) : The Leiden Manifesto Under Review: What Libraries Can Learn From It, Digital Library Perspectives, ISSN 2059-5816, Vol. 33, Iss. 4, pp. 324-338,
<http://dx.doi.org/10.1108/DLP-01-2017-0004>

This version is available at:
<http://hdl.handle.net/11108/323>

Kontakt/Contact

ZBW – Leibniz-Informationszentrum Wirtschaft/Leibniz Information Centre for Economics
Düsternbrooker Weg 120
24105 Kiel (Germany)
E-Mail: info@zbw.eu
<http://zbw.eu/de/ueber-uns/profil/veroeffentlichungen-zbw/>

Standard-Nutzungsbedingungen:

Dieses Dokument darf zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden. Sie dürfen dieses Dokument nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen. Sofern für das Dokument eine Open-Content-Lizenz verwendet wurde, so gelten abweichend von diesen Nutzungsbedingungen die in der Lizenz gewährten Nutzungsrechte.

Terms of use:

This document may be saved and copied for your personal and scholarly purposes. You are not to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public. If the document is made available under a Creative Commons Licence you may exercise further usage rights as specified in the licence.



<https://creativecommons.org/licenses/by/4.0/>



The Leiden Manifesto Under Review: What Libraries Can Learn From It

Please cite as: Coombs, S. K., & Peters, I. (2017). The Leiden Manifesto Under Review: What Libraries Can Learn From It. *Digital Library Perspectives*. [accepted for publishing].

Purpose

This paper is a critical discussion of the Leiden Manifesto for libraries already engaged in bibliometric practices. It offers practical recommendations based on the work of the Association of European Research Libraries (LIBER) Working Group on Metrics. This work is in the beginning phase and summarizes literature on the topic as well as the experiences of the members of the Working Group. The discussion reflects today's growing popularity of (quantitative) research assessment which is seen in enthusiasts introducing new metrics (i.e. altmetrics) and by critics demanding responsible metrics that increase objectivity and equity in evaluations.

Design/methodology/approach

This paper is the result of the Working Group on Metrics of the Association of European Research Libraries (LIBER) that critically discussed the practicality of the Leiden Manifesto for libraries.

Findings

Full compliance with the Manifesto is time-consuming, expensive and requires a significant increase in bibliometric expertise with respect to both staffing and skill level. Despite these apparent disadvantages, it is recommended that all libraries embrace the Manifesto's principles. To increase practicality it is advised that libraries collaborate with researchers, management and other libraries at home and around the world to jointly design and provide services that can be reused within the library community.

Originality/value

Libraries have increasingly been confronted with questions about research assessment, responsible metrics, and the role of digital products in evaluations and funding decisions. Although a wide range of recommendations and initiatives are available (e.g., DORA San Francisco Declaration on Research Assessment) many recommendations are not straightforward enough to be implemented from a library perspective. This paper provides assistance for libraries to implement these principles by acknowledging the heterogeneous backgrounds the libraries may stem from.

Keywords

Research evaluation, metrics, bibliometrics, altmetrics, Leiden Manifesto, libraries, responsible metrics

Type of article

viewpoint





Recently, there has been increased interest in the assessment of research products and researchers. Some of the interest stems from the enthusiasm of having new countable items (such as data sets) and data sources (such as Twitter) available, the other is a reaction to criticism of bibliometrics. Enthusiasts have embraced the field of altmetrics (Priem *et al.*, 2010) that assesses diverse research products and forms of engagement with them (e.g., pageviews, savings to a Mendeley profile, tweets; Piwowar, 2012). The engagement primarily occurs on social media platforms, which offer the option of almost immediate (Shuai *et al.*, 2012) “impact measurement from areas outside science” (Bornmann and Haunschild, 2016) because of the broad user base of social media. Of late, the European Commission under the Commissioner for Research, Science and Innovation, Carlos Moedas, has also emphasized the role of alternative metrics and ways of research assessments as incentives in open scientific practice, referred to as Open Science (EC, 2016). The NISO Alternative Assessment Metrics Project [1] aimed at formalizing altmetrics by providing definitions and use cases, has recommended the use of persistent identifiers for alternative outputs in scholarly communication, data metrics, and a code of conduct for altmetrics data suppliers (NISO, 2016).

These recommendations have produced two benefits: NISO responded to the need for standardization in light of the fact that more and more publishers have included alternative metrics in their portfolios (e.g., Springer [2]); and, NISO provided the environment to upgrade altmetrics to the level of serious metrics that are able to complement citation analyses and derived procedures.

The work of NISO is reflected in several manifestos and initiatives whose aim it is to bring awareness to the problems of quantitative research assessment. This is particularly true for quantitative approaches that when collapsed into one single indicator are reductionist, and fundamentally flawed (Pickton, 2015). Additionally, journal indicators (such as the impact factor) as proxies for quality of research are insufficient (see DORA San Francisco Declaration of Research Assessment [3]). Consequently, research assessment has to be carried out responsibly [4] and should embrace concepts such as robustness, humility, transparency, diversity, and reflexivity during bibliometric exercises (Wilsdon *et al.*, 2015).

Probably the most popular example [5] of recommendations in this vein is the Leiden Manifesto (Hicks *et al.*, 2015) that has been published by renowned researchers in bibliometrics. It is named after the place where it was first developed, the STI Conference 2014, in Leiden, The Netherlands. The manifesto recommends 10 principles that are derived from the best practices of (quantitative) bibliometric exercises. In 2016, the European Association for the Study of Science and Technology awarded the Leiden Manifesto with the Ziman Award for a “collaborative promotion of public interaction with science and technology [...] [that] ”is designed to influence evaluation practice rather than simply to critique it. This is an impressive effort to take specialized scientometric knowledge into a wider policy arena” [6]. By now the Leiden recommendations have found their way into practice: The Loughborough University in the United Kingdom is among the first universities to adopt the Leiden Manifesto as part of their ‘building excellence’ strategy [7].

Regardless of which report or recommendations are read, the same issues with respect to research assessment based on bibliometric methods appear. Each suggest

- a need for transparency of underlying data and methods;
- a default combination of qualitative and quantitative analyses;
- alignment of purpose of assessment, choice of indicators, and discipline;
- acknowledgement of the limitations of data and methods used; and
- regular scrutinization of data, indicators, and methods (Pickton, 2015).

These reports and recommendations assist interested parties to practice quantitative assessment of research in a responsible way. Because these issues have been discussed continuously in the scholarly community, not only among bibliometricians and librarians (Cronin and Sugimoto, 2015), stakeholders familiar with bibliometrics may have already internalized most of the principles discussed in the Leiden Manifesto and other documents – although the state of knowledge among them surely is very heterogeneous. Consequently, these principles may be more relevant to newcomers in the field (Gornitzki, 2015), but they also support development of standardized practices in bibliometric assessments (e.g., regarding their reproducibility or the provision of open data).

Role of Libraries in Research Assessment

Metrics and indicators which track the attention around scholarly products can be considered reflections of their impact. As signals of popularity they increase visibility of scholarly work and can be used for optimizing publication strategies. They are not only a helpful evaluation instrument to complement the peer review system but are also a compass for researchers in the ‘publish or perish’ dilemma. In addition, metrics and indicators aid discovery of scholarly content and provide context on how users engage with it.

Academic libraries have a long experience in dealing with bibliographic data and outcomes of scholarly publications. As the link between the library, publications (open access or traditional), and repositories strengthens and the demands for evaluation statistics by grant awarders and financers increase, this traditional role will most likely become of greater importance. Libraries that do not provide services around metrics may find themselves in a precarious position unless they incorporate this into their daily activities. Libraries already offering bibliometric services may need to expand these services. In addition to specific service organisations such as CWTS [8], research libraries are essential executives of the recommendations of Leiden and other similar documents. Through their expertise and services as well as their direct access to researchers, research libraries are essential for the implementation and success of these recommendations.

Although the recommendations of the reports are predominantly viewed positively, and the need for more reflection and discussion about how to conduct research assessments and the value of science is generally acknowledged, questions about the practicality of the manifestos have been raised. In particular, libraries have asked how to efficiently transfer the principles into the ‘real world’ (FIN, 2016). These issues have gained urgency in light of the increased popularity of alternative metrics (amongst other altmetrics) and the European Commission’s move towards Open Science as a fundamental principle for conducting science in the near future (EC, 2016).

Motivation and Approach

This article presents a research library-focused discussion on the relevance of the recommendations of the Leiden Manifesto for the development and provision of metrics services in libraries. It presents work in progress by the Working Group on Metrics [9] established by the Association of European Research Libraries (LIBER) to support its mission to re-invent the library for the future by “enabling open science, leading in changing scholarship and shaping innovative research” [10]. The members of the Working Group represent a variety of research libraries and research institutions. They are considered experts in the field because they have worked extensively with or studied traditional (e.g., citation-based) as well as alternative, i.e. social media-based, metrics. The goal of the Metrics Working Group is to exchange best practices that can be used to develop recommendations for the optimal use of metrics in research libraries and information infrastructures.

Focused on the evaluation of scholarly publications, many of the principles of the Leiden Manifesto (Hicks *et al.*, 2015) provide a solid base for the beginner library as well as more



advanced principles for experienced libraries. However, the LIBER Working Group agrees with the conclusion of the fifth meeting of the Danish Research Indicator Network that for libraries “the manifesto is not straight forward [sic!] to implement” (FIN, 2016, p. 1). Thus, the Working Group set out to build upon the Leiden principles and translate them specifically for libraries. The Working Group discussed them critically in order to offer practical recommendations for the implementation of the principles [11]. In addition, the priority of the single recommendations are discussed, as well as, who (inside a library or institutional environment) should be responsible for its execution and who might request it. The criticisms and suggestions of FIN have also been incorporated, and expanded upon. Similar to Bornmann and Haunschild (2016) and their discussion of the Leiden Manifesto for altmetrics, the Working Group combines an evaluation of the high-level recommendations with practical examples by taking the heterogeneous state of knowledge of the libraries into account. Bornmann and Haunschild (2016) concluded after their extensive discussion of the manifesto for altmetrics that “the principles also have a great importance for altmetrics and should be taken into account in their application”. Consequently, we will review the principles for all metrics and will not focus on a particular kind of metrics.

By taking an inclusive approach the recommendations of the LIBER Working Group on Metrics will address beginners as well as advanced metrics users and will provide librarians with practical ideas, tools, and starting points for the efficient use of metrics, regardless of the size of the library, its budget, or location. In the future, the Working Group will discuss possibilities for open metrics, as reflected by the altmetrics and open science movements, and what this can entail.

Evaluation of Leiden Manifesto for Library Work

The following will assess the relevance of the recommendations of the Leiden Manifesto for the development and provision of metrics services in libraries.

#1 Quantitative evaluation should support qualitative, expert assessment

This is an essential principle that is of high priority for all libraries to assimilate into their daily services regardless of whether the library is in the start-up phase of their metrics services or whether it is more advanced. Quantitative metrics should not replace but should inform and strengthen expert assessment. Oftentimes quantitative and qualitative evaluations mutually inform each other (FIN, 2016). This recommendation facilitates transparency of assessment practices, enables scrutinization of indicators, and allows for inclusion of disciplines and data sources that are not standard (e.g., specific databases from humanities; FIN, 2016). Bornmann and Haunschild (2016) point out that altmetrics are a valuable resource adding to the ‘basket of metrics’ (Colledge and James, 2015) and expert assessment. Although they neither replace traditional metrics nor peer review altmetrics can almost immediately provide a broader picture of engagement with scholarly products provided they are equipped with persistent identifiers, e.g., DOIs (digital object identifiers).

According to FIN (2016) the downside of this recommendation is that qualitative evaluations are expensive and often time-consuming. Also, quantitative evaluations need time. Gorraiz *et al.* (2016) estimate that a bibliometric report for an individual researcher takes about four days for a bibliometrician. However, because to date there is no culture of assessment that embraces this principle, currently qualitative reports must be concise to be read, and should reflect the multidimensionality of research and its impact.

The need for human judgement in the final step of quantitative as well as qualitative assessments makes it paramount that librarians and support staff learn about metrics, their uses, and their limitations (for example via [12]). Librarians need to work closely with fellow local and international groups to increase their expertise (for example via [13] and Chin-Roemer and Borchardt, 2015; Holmberg, 2016; Konkiel, 2016; Tattersall, 2016).

When asked to provide quantitative information on researchers or scientific products, libraries should always provide a narrative accompanying the numbers with the narrative preferably cross-validated with peers (for example with the researchers under review; Gorraiz et al., 2016). For the quantitative evaluation of individuals Gorraiz et al. (2016) recommend including the following themes to guide the narrative:

- Methodology, i.e. description of used databases and indicators;
- Coverage in databases, i.e. demonstrate selection of adequate databases and coverage of research products (how much can be found in Web of Science, Scopus, Google Scholar, or other disciplinary databases?);
- Activity analysis for publications, i.e. presenting output in terms of most important publication and document types;
- Affiliation and funding analyses, i.e. information on main funding agencies and relationship between productivity and change of affiliations;
- Co-authorship analysis, i.e. indication of collaboration patterns and role of the author under review (e.g., first author vs. last author);
- Visibility analysis, i.e. number of items in top journals (e.g., determined via Journal Impact Factor);
- Impact analysis, e.g., citation analysis of publications, share of self-citations (<20% regular) and determination of h-Index;
- Citing analysis, i.e. to demonstrate the outgoing knowledge flows, e.g. which journals cite the author under review;
- Network analysis, e.g., cooperations with international, national, and domestic partners;
- Reference analysis, i.e. to demonstrate the incoming knowledge flow, e.g., on which disciplines the author builds his research;
- Research focus, i.e. summary of major research topics;
- Summary, i.e. main findings and specifications need to be known for analysis
- Annex including all data used in analyses.

Such a bibliometric report as this is the first step in research assessment and aims at revealing “potentially meaningful symptoms and to discuss them with the researcher under evaluation. [The researchers] are invited to clarify them in their own additional evaluation report” (Gorraiz et al., 2016, p. 910).

Libraries should learn more about scholars’ aims and challenges when conducting and publishing the research (e.g., their goal may have been to address practitioners with their articles. The resulting high number of articles in non-peer reviewed journals for professionals may not have been a perfect outcome in terms of the impact factor, that is considered as the quality criterion for journals all researchers strive for. But for the researcher under review the publication effort was truly intended and successful).

As well, close collaboration with research administrations to prepare data annexes for informed peer review will increase awareness of the importance of the narrative as well as the effectiveness of the assessment exercise. Before starting analyses the principal must communicate the purpose of the assessment to the executors and its targets (FIN, 2016). Hence, when preparing the quantitative evaluations libraries need to obtain answers to the following questions:

What is the unit of comparison? The assessment of individual researchers requires data and indicators other than those of departments or entire universities (see also principle #3).

Should the bibliometric report describe the current state of research productivity and/or impact or should it be used for comparison (with other researchers, department, proposals, points in time



etc.)? Comparative reports always need a carefully chosen point of reference, e.g., mean of the disciplines' productivity or last ten years.

#2 Measure performance against the research missions of the institution, group or researcher

In order to effectively use metrics for the evaluation of academic output it is necessary that libraries know the scientific culture and practices (i.e. citing, publishing, assessment, journal policies, requirements of funders) of their target group. Metrics are used to measure merits and performance in relation to specific research goals. Therefore, the goals have to be made explicit and matched to adequate indicators. Because no one-size-fits-all solution exists assessments can (and should) be highly individual even though. Libraries will repeatedly be confronted with similar goals and challenges (e.g., questions about adequate field normalization or citation window). Hence, libraries should aim at exchanging best practices and developing processes that, to a certain extent, can be standardized. These processes should be linked to use cases (i.e. research missions) such as journal selection, impact assessment of research projects, increasing visibility of research products, or tenure cases [14].

FIN (2016) pointed out, however, that in most assessment practices the influence on choosing or designing the metrics is limited because no standards are applied and the choice of metrics and methods strongly depends on the principal of the research assessment. In addition, the purpose (i.e., mission) of the assessment is often unclear, too general, or not available at all.

This principle is relevant for libraries and must be considered when developing user-oriented services. While it is a low priority for libraries, it requires the work of management and analysts in an ongoing manner. More intensive periods are required upon request or prior to visitation or inspection. Libraries should aim at being included in the development phase of the assessment so that purpose, methods, and indicators can be aligned (FIN, 2016). To facilitate it, the fields or groups that will be measured should be defined in collaboration with researchers and research administration. Also, it is important to ensure that goals and indicators of success are defined early.

#3 Protect excellence in locally relevant research

This principle is of medium priority and relevance for libraries. It is, according to FIN, of greater importance for those libraries serving humanities and social sciences as well as those serving practice based researchers such as public health (FIN, 2016). It requires the work of management and analysts in order to provide tools to demonstrate diversity and (societal) impact of the local organization as well as to be as inclusive as possible by aiming at high quality. It requires an entire library team to develop the service that provides this information when required. To do so, practitioners should highlight the diverse nature of research at the institution, ask for appropriate denominators, and act as the advocate for researchers, informing others about the diversity of research products and researchers (for example, art exhibitions vs. software as research products).

To illustrate this, the local university can be used as a means of comparison. Libraries could use the content of their local repository or current research information system (CRIS) to compare departments against the average performance of similar departments or to help researchers at their university in self-assessment exercises (e.g., comparing their performance against colleagues of similar academic age). This provides a more realistic picture of the performance and will level out factors such as location, which may have positively influenced researchers at different places. In the long run, repositories and CRIS data could even be used to build discipline- or nation-specific databases to serve as alternatives to proprietary bibliographic databases and provide complete data for benchmarking. Before doing so, guidelines or standards for the contents to be included in CRIS's have to be established in order to prevent relying on subjective data that was

entered by the subjects under review. Those guidelines should be collaboratively developed with the administration and researchers.

It is important to remember that all the data is needed for calculating relative size independent indicators. However, a lack of standardisation for data means that this can be a challenging task for certain areas of research. It is imperative to clearly and realistically recognise what can be measured (FIN, 2016).

#4 Keep data collection and analytical processes open, transparent and simple

This is an essential principle for all libraries when developing processes, studies and services. Transparency and openness concerning data collection and methods for indicator development and assessment reports is fundamental since it simultaneously provides authority of the analyses and allows for scrutinization and adjustment.

The level of fulfillment of this principle strongly depends on the evaluation task and the local situation (FIN, 2016). FIN (2016) is critical of the concepts surrounding this principle (i.e., open, transparent, simple) that it is not well-defined and leaves room for interpretation. A further complication is the heavy reliance on third parties for data collection and use of indicators, which limits transparency and openness (FIN, 2016). Data is not static and procedures as well as data sources included in the research assessment may change during the analyses. This competes with the notion of simplicity (FIN, 2016). Moreover, often indicators and metrics are neither simple nor easy to understand, especially when they are not in-house-products (FIN, 2016). FIN (2016) did acknowledge, however, that third party data with a certain reputation can serve as quality assurance in bibliometric exercises.

In order to do this effectively the user must:

- Develop a culture of open methodology so that one is prepared and open for scrutinization of data and methods by internal and external parties;
- If openness (as defined in the open definition [15]) is not possible because of third-party restrictions, aim for transparency in data, codes, methods and processes, studies, and results;
- Aim at defining concepts and methods used in research assessments (FIN, 2016);
- Strive to re-use or implement standards in study design (at least on the local level);
- Re-use and enhance open software that can be used to study scholarly communication including literature and data (e.g., [16]);
- Collaborate with colleagues to develop standards and make data and metadata transparent, open, and accessible wherever possible;
- Document procedures, methods and decisions made and be as detailed about the data and its caveats as possible (as for example in the NISO Code of Conduct for Altmetrics Providers [14]; FIN, 2016);
- Master open tools and datasets. Support staff, analysts, web managers and librarians alike must find their footing here.

#5 Allow those evaluated to verify data and analysis

Highly relevant and of high priority, this principle requires collaborative work between libraries and researchers. Complying with this principle ensures integrity of the research assessment and the people who execute it (FIN, 2016). As FIN (2016) pointed out this principle is strongly linked to principle #4 and therefore reflects similar problems as, for example, reliance on third parties. These data providers are oftentimes 'black boxes' that require an extensive amount of trust and limit transparency (Cochrane, 2017).

However, on a local level, libraries can collaborate with researchers for data import and can adjust their services to researchers' needs (e.g., addition of unconventional research products, such as software or presentation slides, in CRIS). Also, allowing researchers to regularly update and check data collected for them builds trust and ensures data quality. Certain areas of CRIS or other information systems might grant editing permissions to several stakeholders (e.g., researchers, librarians, university administration) so that they can manipulate data and/or reuse it on their own.

To overcome the third-party-data dilemma, libraries (as has been recently described for academia.edu [17]) should collect as much scholarly data and usage information (e.g., downloads, library loans) as needed and appropriate for its purposes (see #2). This data should be openly available to enable its reuse and to allow for its scrutinization. This is also true for the methods and tools used to collect and analyse the data. Libraries should use as much open software as feasible (e.g., [16]).

The following recommendations will help to make this cooperative effort as effective as possible:

- Be critical in the use of the data and verify its correctness (FIN, 2016);
- Prior to beginning the process, explain the need for data and the required format to those that fill in data bases;
- Actively request feedback and approval from the beginning of the start up of service and analysis, to the end when delivering results, including preliminary findings early on. Do not secretly perform analyses and distribute results, other than to the subjects under investigation before having them checked;
- Provide a “hotline” service via the library;
- Find ways to provide data on several aggregation levels (i.e. personal, institutional, national, global) for easier verification and analysis;
- Provide added value that serves as incentive for data provision, e.g., researchers can reuse their CRIS- and/or local repository entries for publishing their publication lists on the website;
- Ensure excellent data quality but keep the verification by researchers easy;
- Assist in the scrutinization of data analyses by documenting any errors found; and
- Make others aware that data can be political (think about an ethics of research assessments; FIN, 2016).

#6 Account for variation by field in publication and citation practices

Knowing about the scientific culture and practices (i.e. citing, publishing, assessment, journal policies, requirements of funders) of the target group, as well as consulting with the target group about the selection of adequate indicators or the creation of additional indicators is key to this principle. This principle is both highly relevant and a high priority for libraries for the initiation of metric services.

Although this principle is considered easy to implement FIN (2016) emphasizes that correct field normalization is a difficult process, denominators have to be chosen carefully, and they are subject to constant change. Furthermore, the field normalization can be responsible for strongly influencing the result of the quantitative assessment, even more than the actual performance of the field.

To account for variation, a library team should work on a set of reasonable denominators (e.g., mean citation rate in a discipline, number of department members), data source (e.g., Web of Science with only a limited number of social science journals or Google Scholar which includes a broad range of scholarly sources) and publication types (e.g., data sets, journal articles, working papers, books, software) to acknowledge field differences. Libraries should be aware of rankings (e.g., for conferences, journals, or universities) that might be of importance to their target groups.

Such rankings may affect publication behavior of the researchers. In addition, the ranking should provide context for each data point included in bibliometric analyses to enable understanding and comparison. Collaboration and exchange with the target group will increase the quality of the service and its acceptance rate. Similar to principle #3 but with a global perspective, the diverse nature of research at the institution as well as in the field should be highlighted, and appropriate denominators and indicators requested. Using a range of appropriate metrics from which at least two indicators are chosen for assessment is a reasonable approach (Colledge and James, 2015).

#7 Base assessment of individual researchers on a qualitative judgement of their portfolio

This principle is of low relevance and priority for libraries but high for researchers, management, and the managing team of the CRIS system since it is strongly linked to Principle #1. However, extensive collaboration between libraries and university administrations is needed here to establish the necessary links between the different types of information (e.g., publications, amount of funding, education, and academic age of the researchers). Researchers may be added to this list if social media profiles are important. Trust in information systems, privacy, and data protection as well as transparency of goals and methods are key success factors in this regard.

FIN (2016) emphasizes that compliance to this principle is easier in internal evaluations since they can flexibly react to the principal's needs (as has also been described in Gorraiz *et al.*, 2016). External evaluations, on the other hand, are serious barriers for following this principle because they often ask for specific metrics, which leaves little room for choice or additional information (FIN, 2016).

FIN (2016) reports little experience in qualitative portfolios, but emphasizes the role of the CRIS that offers the infrastructure that includes diverse non-traditional research products which can then demonstrate broader impact. What is in the CRIS is the most relevant output for consideration. The inclusion of non-traditional research products requires new expertise for those who carry out qualitative research assessments.

Hence, this principle demands extensive understanding on two levels. In order to facilitate and provide more accurate bibliometric analyses that properly reflect (personal) contexts it is essential to: 1) learn as much as possible about the researchers being served, research products provided, and disciplines analysed and 2) learn appropriate CRIS design and bibliometric analyses to reflect those specificities. Doing so can facilitate, among other things, comparison between researchers by taking their life situation into account (e.g., academic age, career stage, or parental leave). For instance, when assessing individual researchers for 6 month uninterrupted periods, the researcher might be asked for their five most important publications or a completed personal portfolio (see the ACUMEN Portfolio [18]).

#8 Avoid misplaced concreteness and false precision

This principle is of high priority for libraries. It is essential for the entire process from the initial interview with participants and is embedded in the end deliverables. This issue is most likely to occur when only a single indicator is presented in the report (FIN, 2016).

Therefore, in order to facilitate avoiding this misinformation:

- A variety of metrics should be used as it can be assumed that more indicators increase the robustness of analyses (the basket of metrics approach);
- Individual librarians, as editors of supporting documents and/or annual reports, should understand metrics, data, and methods to avoid false precision when displaying the results;
- Indicators need to be presented in a way that makes their limited precision clear to the user (e.g. multiple citation indicators with reference to the respective data source);

- Avoid misleading extrapolation (e.g., pars pro toto) or general conclusions based on small samples (FIN, 2016);
- Always disclose and discuss the limits of the study (even if they are not required; FIN, 2016); and
- Avoid composite indicators that appear to be simple but conceal the meaning of metrics.

#9 Recognize the systemic effects of assessment and indicators

The use of a single indicator invites gaming and unintended behaviour where the measurement becomes the goal. This effect increases the closer the assessment gets to the individual level (FIN, 2016). Therefore, this principle is a priority and highly relevant for libraries. It is also paramount that researchers and administration be educated about the responsible use of metrics (see for example the Responsible Metrics movement [19]) and gaming effects (as described in Goodhart's Law: "When a measure becomes a target, it ceases to be a good measure").

Given that research assessment always is related to performance (i.e., output) and impact (i.e., reputation measure via metrics) gaming also applies to performance, i.e. researchers have to publish before they can have impact. Libraries should draw on their expertise and assist researchers in finding publication venues of high quality by, for example, taking journal indicators into account, and other characteristics such as editorial boards. Librarians can also regularly review publication outlets to efficiently detect predatory journals or conferences.

Individual librarians, analysts in assessment exercises, and management can support counteracting the systemic effect by making the limitations and shortcomings of the methodology and data of research assessment explicit. At some point, participants can be educated about the assessment and the measure and data used. It is also advised by FIN (2016) that local conditions should be taken into account and knowledge be available about external drivers (e.g., funders that demand particular outcomes) that shape behaviour of the target group. This knowledge must also be assessed and interpreted (e.g., if funded projects must comply to an open access mandate, high numbers of open access publications are not signals for a legitimate increase in open access publishing but in the results of regulation).

#10 Scrutinize indicators regularly and update them

This is an ongoing fundamental role for the metrics analyst and their team and a principle of high importance. Libraries must be prepared for constant and/or immediate change on several levels: new users or databases can enter the field; the latter can be updated; a new evaluation task can be requested; or the needs of the users may change (FIN, 2016). Constant change means that evolutionary comparisons of assessment exercises are difficult to pursue and that there is a possibility that new indicators are not available for historic data (FIN, 2016). Adaption to change requires time and money, and allocation of these resources depends on formal decision making processes (FIN, 2016). It takes time to explain changes to the target group (FIN, 2016).

It is recommended to:

- Keep the overall goals of the research assessment up to date and re-evaluate the choice of metrics accordingly;
- Observe trends and changes in the target group and collaborate with the target group in service building as well as in assessing new emerging metrics and evaluating existing metrics that may be superseded or out of date;
- Observe trends and changes in metrics and scientometrics;
- Assist with regulating, updating, maintaining a bibliographic database of the research output of the institution and target group;



- Avoid flawed data (FIN, 2016) but provide data to support scrutinization and comparison of indicators;
- Document the changes in data and processes carried out and any errors found (FIN, 2016).

Conclusion

As the role of metrics in the evaluation process changes, the traditional role of the library must parallel it. As the evaluation of scholarly publications becomes of greater importance so must the role of the library. Both traditional metrics, based on, for example, citations, and new metrics, based on, for example, social media activity, can be utilized for creating library services which assist and support different types of principles of research assessments (from individuals to larger groups). Those libraries not yet offering metrics services to researchers need to make this a priority. Those already offering services need to assess what they are currently offering and adapt to meet the changing requirements to develop into bibliometric experts. Modern scientific libraries must face the new challenges of the digital era and its potential for research assessment (e.g., open peer review) sooner rather than later.

Recommendations for responsible research assessments presented in the Leiden Manifesto (Hicks *et al.*, 2015) and elsewhere are essential reads for everyone involved in such exercises. Every effort should be made to meet the requirements formulated in those principles regardless of whether libraries are requested to work with more traditional forms of research assessment, e.g., citation analysis, or with altmetrics (Bornmann and Haunschild, 2016). It is important to acknowledge that these efforts may vary in intensity and quantity, and that it is an evolutionary process that has to be undergone rather than an instant turnaround. Libraries are an important hub of bibliometric practices as well as indispensable executors and drivers of responsible research assessments.

This article aims to translate the principles of the Leiden Manifesto into more practical recommendations that indicate the priority of each principle and the role of those involved. The analysis indicates that full compliance to the principles is time-consuming, expensive, and requires a significant increase in bibliometric expertise with respect to both staffing and skill level. The need for time, funding and expertise is increasing given that the environment of research assessment and scholarly work is subject to constant change as are bibliometric methods and data sources (Gornitzki, 2015). Additional obstacles to adopting the principles are the current system of research evaluation and principles as well as third party involvement in research assessment that limit, if not hinder, transparency, openness, simplicity, and alignment of evaluations.

The overall recommendation that can be derived from these findings is that libraries need to collaborate if they are to successfully face the new challenges of responsible assessments. Rather than performing individual analyses or developing in-house solutions that are of limited use libraries should work together and develop modular, broadly applicable services and standards that can be used throughout the library community. Ideally, libraries will increase acceptance of these exercises, educate about strengths of evaluations and weaknesses of 'irresponsible' evaluations, and be actively involved in the design and development of proper and responsible systems, data sources, metrics, and methods.

To facilitate this, libraries need to build awareness first and then advertise their services. As responsible partners in the research assessment exercises they should also clarify what services they offer, what results are to be expected, who might be customers of the services, and with what guidelines they comply, e.g. Leiden Manifesto or DORA. Good examples are provided by University Library Vienna [20] or Loughborough University [21].

In addition, libraries should emphasize the diverse fields where metrics are valuable tools and which go far beyond research assessment. They could, for example, demonstrate that bibliometric and altmetric methods also assist in discovering literature and research collaborators (e.g., via



co-citation analysis or ranking mechanisms that take into account the number of citations) or in increasing the visibility of research products (via showcasing successful examples). At the university level, information taken from CRIS's about newly published research products could be distributed to colleagues to build awareness about the excellence of local research (see #3) as well as possible collaborators on site.

This cooperative service development approach and the successful reuse of those services requires interoperability of systems and standards (e.g., for exchange or import of data or bibliographic records). This allows libraries to build on their experiences in developing such mechanisms and quality guidelines (as for example demonstrated by [14] and [22]), at the same time as they can actively make the case for open standards (e.g., MODS [23]), data sources, and software (e.g., [16]). This guarantees service development and its provision outside from proprietary solutions, allows for uncomplicated travel of data, services, and software, and enables easy reuse. In fact, this last benefit may be the primary incentive for researchers to provide information on the research products.

In conclusion, the Leiden Manifesto provides a solid foundation on which academic libraries can assess the implementation of metrics in their institutions. Through the prioritizing of its principles and the concrete and practical recommendations of the LIBER WG on Metrics, libraries already engaged in metrics can build the necessary skills, use the right tools, and make the right collaborative efforts to enrich the academic world. Through these tools libraries can remain relevant. They can move with today's growing popularity of (quantitative) research assessment and contribute to responsible metrics that increase equity in evaluation.

Acknowledgements

We would like to thank Helen Coombs for editing the text and the LIBER Working Group on Metrics whose expertise has led to the recommendations outlined in this article: Kasper Abcouwer, Isidro F. Aguillo, Nathalie Cornée, Ellen Fest, Juan Gorraiz, Stefanie Haustein, Kim Holmberg, Najko Jahn, Peter Kraker, Ania López, Marc Martinez, Alenka Princic, Susan Reilly, and Birgit Schmidt.

Endnotes

- [1] http://www.niso.org/topics/tl/altmetrics_initiative
- [2] <https://www.springer.com/gp/authors-editors/journal-author/journal-author-helpdesk/article-level-metrics/23400>
- [3] <http://www.ascb.org/dora>
- [4] <https://responsiblemetrics.org>
- [5] <http://www.leidenmanifesto.org/blog/altmetric-on-the-lm>
- [6] <https://www.cwts.nl/news?article=n-q2x294>
- [7] <http://www.lboro.ac.uk/research/scholcomms/assessment/respmetrics>
- [8] <https://www.cwts.nl>
- [9] <http://libereurope.eu/strategy/strategic-direction-1-enable-open-science/wg-metrics>
- [10] <http://libereurope.eu>
- [11] As the development of recommendations is still a work in progress we encourage readers to share their views with us. Please email the authors for feedback.
- [12] https://libraryconnect.elsevier.com/sites/default/files/ELS_LC_metrics-poster_June-23-2016.pdf
- [13] Blog (<https://thebibliomagician.wordpress.com>) and mailing list (LIS-BIBLIOMETRICS@JISCMail.AC.UK)
- [14] NISO provides a code of conduct and use cases for altmetrics which might serve as starting point for further development: <http://www.niso.org/publications/rp/rp-25-2016>



- [15] The open definition: “Open data and content can be freely used, modified, and shared by anyone for any purpose” (<http://opendefinition.org>)
- [16] For example: <https://ropensci.org>, <http://www.harzing.com/resources/publish-or-perish>, <http://homepage.univie.ac.at/juan.gorraiz/bibexcel/>, <http://www.vosviewer.com/>, <http://mrvar.fdv.uni-lj.si/pajek/>
- [17] <http://www.forbes.com/sites/drsarahbond/2017/01/23/dear-scholars-delete-your-account-at-academia-edu>
- [18] <http://research-acumen.eu/wp-content/uploads/MS5-ACUMEN-portfolio.pdf>
- [19] <https://responsiblemetrics.org>
- [20] http://bibliothek.univie.ac.at/bibliometrie/for_administrators.html
- [21] <http://www.lboro.ac.uk/research/scholcomms/assessment/bibliometrics/>
- [22] <https://www.snowballmetrics.com/>
- [23] <http://www.loc.gov/standards/mods/>

Bibliography

Bornmann, L. and Haunschild, R. (2016), “To what extent does the Leiden manifesto also apply to altmetrics? A discussion of the manifesto against the background of research into altmetrics”, *Online Information Review*, Vol. 40 No. 4, pp. 529-543.

Cochrane, A. (2017), “How Many Grains of Salt Must We Take When Looking at Metrics?”, available at: <https://scholarlykitchen.sspnet.org/2017/02/08/49186/> (accessed 10 February 2017).

Colledge, L. and James, C. (2015), “A “basket of metrics”—the best support for understanding journal merit”, *European Science Editing*, Vol. 41 No. 3, pp. 61-65.

Cronin, B. and Sugimoto, C. R. (Eds.) (2016), *Scholarly metrics under the microscope*, 1st ed., Medford, NJ, Information Today, Inc./ASIST.

EC (2016), *Open Innovation, Open Science, Open to the World - A Vision for Europe*. DG Research and Innovation. 1st ed. Luxembourg: Publications Office. DOI:10.2777/061652

Gornitzki, C. (2015), “The Leiden Manifesto - Ten principles for a better research evaluation”, available at: <https://kib.ki.se/en/whatsup/blog/leiden-manifesto-ten-principles-better-research-evaluation> (accessed 10 February 2017).

Gorraiz, J., Wieland, M. and Gumpenberger, C. (2016), “Individual bibliometric assessment at University of Vienna: From numbers to multidimensional profiles”, *El Profesional de la Información*, Vol. 25 No. 6, pp.901-914.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S. and Rafols, I. (2015), “Bibliometrics: The Leiden Manifesto for research metrics”, *Nature*, Vol. 520 No. 7548, pp.429-431.

Holmberg, K. (2016), *Altmetrics for Information Professionals. Past, Present and Future*, Chandos Publishing, Waltham, MA, USA.

Konkiel, S. (2016), “Altmetrics for librarians: 100+ tips, tricks, and examples”, available at: <https://www.altmetric.com/libraries-ebook> (accessed 10 February 2017).

NISO (2016), “NISO RP-25-2016: Outputs of the NISO Alternative Assessment Project”, available at: <http://www.niso.org/publications/rp/rp-25-2016/> (accessed 10 February 2017).

Pickton, M. (2015), “The Metric Tide: Are you using bibliometrics responsibly?” available at <https://researchsupporthub.northampton.ac.uk/2015/12/09/the-metric-tide-are-you-using-bibliometrics-responsibly/> (accessed 10 February 2017).



Flavors Of Research Impact Through #Altmetrics”, available at: <https://researchremix.wordpress.com/2012/01/31/31-flavours/> (accessed 10 February 2017).

Priem, J., Taraborelli, D., Groth, P. and Neylon, C. (2010), “Altmetrics: A manifesto”, available at: <http://altmetrics.org/manifesto> (accessed 10 February 2017).

Roemer, R. and Borchardt, R. (2015), *Meaningful metrics*, 1st ed., Association of College and Research Libraries, a division of the American Library Association, Chicago.

Shuai, X., Pepe, A. and Bollen, J. (2012), “How the Scientific Community Reacts to Newly Submitted Preprints: Article Downloads, Twitter Mentions, and Citations”, *PLoS ONE*, Vol. 7 No. 11, p.e47523.

Tattersall, A. (Ed.) (2016), *Altmetrics: A practical guide for librarians, researchers and academics*, 1st ed., Facet Publishing, London, UK.

Wilsdon, J. et al. (2015), “The metric tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management”, DOI: 10.13140/RG.2.1.4929.1363.