This item is the archived peer-reviewed author-version of:

The bright and the dark side of national databases for research output

The bright and the dark side of national databases for research output

Linda Sīle

Centre for R&D Monitoring (ECOOM), University of Antwerp, Belgium

Linda.sile@uantwerpen.be

https://orcid.org/0000-0003-1435-999X

Abstract

National databases for research output are increasingly seen as a remedy for many data-related problems in research evaluation. A central argument in favour of data collection initiatives at the national level is the need for more comprehensive coverage of research output in, among else, social sciences. At the same time, the air of optimism within the debates about national databases forecloses space for more critical considerations of risks that come along the operation of a national database for research output. This, I argue, is a crucial limitation of the current debates around data sources for research evaluation. In this chapter, I offer a more symmetrical view of the use of national databases in evaluation of social sciences and highlight key positive and negative aspects of these databases with respect to research evaluation.

Keywords: bibliographic databases, research information, infrastructure, context

Introduction

National databases for research output are increasingly seen as a remedy for many data-related problems pertaining to evaluation of research in social sciences, especially when considering specifics of communication practices. Scholars in different knowledge domains carry out research in sometimes radically different manner. Also the means to communicate research are not always the same. In natural sciences, journal articles have been key media for scholarly communication since the establishment of *Philosophical Transactions of the Royal Society* and other early journals in the 17th century. It is still the case today. In social sciences (among others), on the contrary, journal articles are just one of the many artefacts researchers use to communicate (Hicks, 2004). Among the more common media, scholars use monographs, chapters in edited volumes, conference papers, essays and opinion pieces in newspapers, to name a few. For instance, social scientists have noted that

monographs, even though few in number (Engels et al., 2018), are deemed a central carrier of knowledge in a number of academic disciplines (Crossick, 2015) and manifest in citation analyses (Chi, 2014; Glänzel et al., 2016). A monograph is a space to develop more substantial arguments that cannot be easily expressed in such limited spaces as articles or even contributions to edited volumes (see also Sīle et al., 2021). Besides publications, novel forms of scholarly communication—movies, exhibitions, performances—are used in some disciplines within the social sciences. At the same time, only a small share of these activities are included in the often used international data sources such as Web of Science and Scopus. Therefore, a substantial part of the research output (and its citations) authored by social scientists remains invisible (Kulczycki et al., 2018; Martín-Martín et al., 2020).

This diversity in communication media is a characteristic shared with academic disciplines in the humanities and other knowledge domains with specific communication practices (e.g., engineering and computer science). Thus, arguments presented here are applicable to other knowledge domains where communication practices go beyond the mainstream focus on journal articles.

A feature of research and communication practices that is more often encountered within the social sciences concerns interaction with governmental bodies. This interaction can result in policy briefs, governmental reports, and other less tangible ways of communication—consultation, participation in a debate, etc. Finally, it is important to point out that the diversity within the social sciences is not easily captured in relation to specific disciplines. Research exploring within-discipline publication patterns shows that also within a single discipline there can be researchers whose publishing practices are more alike those in the natural sciences (i.e. focus on articles indexed in Web of Science) along with researchers who favour book publications—a characteristic typically attributed to the humanities (Verleysen & Weeren, 2016). All these characteristics of social sciences are a challenge for databases used in research assessment as it is not an easy task to determine what counts as research output for social sciences and how it should be represented in a database. National databases are not an exception. Even though there are likely similarities in communication practices across different national contexts, in many cases, extensive work is required to identify the key characteristics in a particular national context.

Awareness of this challenge is mirrored in recent debates and initiatives on research assessment that highlight the need to take into account the diversity of research—the different understandings of research, different research practices, and different scholarly communication practices (Curry et al., 2020; DORA, 2012; Hicks et al., 2015; Moher et al., 2020; Wilsdon et al., 2015; Working group for the responsible evaluation of a researcher, 2020). For example the term 'Responsible Research Assessment' denotes "approaches to assessment which incentivise, reflect and reward the plural characteristics of high-quality research, in support of diverse and inclusive research cultures"(Curry et al., 2020, p. 4). Similarly, the term 'research metrics'

denotes metrics (i.e. quantitative indicators) that address the following dimensions: robustness, humility, transparency, diversity, and reflexivity (Wilsdon et al., 2015). In relation to the data used in research assessment, a number of academics and authorities have expressed the need for high quality data sources (Mahieu et al., 2014; Martin et al., 2010; Science Europe, 2016; Sivertsen, 2010). Particularly, it has been argued that "[p]ublication metrics should be based on data that is relevant for the unit of assessment" (Working group for the responsible evaluation of a researcher, 2020, p. 12).

In this context, national databases are often seen as a way to address the diversity of research and counter the narrow focus on journal articles perpetuated by the use of commercial data sources such as the Web of Science and Scopus (Sivertsen & Larsen, 2012). National databases, as the argument goes, can include book publications and other forms of research output important for social scientists thus making available for research assessment data that are relevant for this knowledge domain. The point about comprehensive coverage, has by now become the go-to argument when discussing how national databases facilitate research evaluation. A report from 2010, discussing the need for a European database for research output in the social sciences (and humanities), called for a database "that brings together, in a consistent and comparable form, data on the main research outputs of SSH [i.e. social sciences and humanities] . . . and also provides an indication of the impact of those research outputs not only on fellow academic researchers but also more widely" (Martin et al., 2010, p. 2). Mahieu and colleagues pursued a study that aimed "to analyse the desirability and feasibility of creating a transnational system for collecting and monitoring research performance data (on inputs, outputs and productivity) in order to improve policymaking and to identify relevant research policy options" (Mahieu et al., 2014, p. 1). Similarly, in its position statement Science Europe called for "research information systems that are capable of effectively supporting this constant and unpredictable change with the intelligence and insights needed to perform strategic, analytical and management functions" (Science Europe, 2016, p. 3). All these examples point to the increasing awareness of the role of data played in research assessment and the benefits a national research information infrastructure could bring.

In the last two decades, the construction of national databases and information systems has gained momentum. In Europe alone, the number of databases and information systems operated at institutional, regional as well as at the national level has been rapidly increasing (Ribeiro et al., 2016; Sīle et al., 2017). On the one hand, this expansion of data sources collecting data on research outputs in social sciences (and science in general) offers resources that can be used in research assessment, monitoring, and for understanding of research practices in a broader sense. On the other hand, the air of optimism dominating the debates around the value of national databases in research assessments distracts from the more reflective and critical considerations about these data sources. For instance, over the last few recent years, we can witness an accumulation of critique of quantitative approaches to research

evaluation (Barré, 2019; Curry et al., 2020; de Rijcke et al., 2015; Rafols, 2017; Rafols & Stirling, 2020; Wilsdon et al., 2015). And yet, new systems, often with the production of evaluative metrics as their primary use, are being constructed and implemented without serious consideration of implications this critique has for systems—already existing ones and those yet to be implemented. This, I argue, is a crucial limitation of the current debates around data sources for research evaluation.

In this chapter, I offer a more symmetrical (Bloor, 1976) view of the use of national databases in evaluation of social sciences. First, I discuss how these databases can facilitate research evaluation—the bright side. I focus on two aspects—flexibility and transparency—both of which have the potential of rendering national databases particularly useful for the assessment of social sciences. Second, I bring to attention risks pertaining to the use of national databases for research evaluation purposes—the dark side. In this part of the paper, I limit the discussion to the following two points: (i) the risk of national databases becoming too locked into local practices and (ii) the implications for research evaluation that follow from the strong focus on research output that dominate the use of databases for research output in research assessment. Finally, I discuss directions that can be taken in order to make considerate and informed choices pertaining to the design and use of national databases in the context of research evaluation.

Framing databases for research output

Before that—some definitions and conceptual clarifications. By national databases of research output I mean structured collections of bibliographic metadata that refer to research output authored by researchers. This is a broad definition. This definition includes databases maintained for a variety of purposes—research evaluation or information retrieval, for example. At the same time, it saves the essential characteristic of databases for research output which is its content; a database for research output contains metadata of research output (see an example of such a record in Figure 1). The qualifier 'national' here denotes the scope of a database. A national database has a collection from the whole country, in whichever way defined (see more on this in Sīle et al., 2018). In terms of the technical setup of a database, I make no distinction between in-house projects and tailor-made solutions provided by commercial entities. Also, it does not matter, whether a database is a stand-alone database, a networked system, or a module in a broader research information system. The points discussed in this chapter are applicable to all these alternatives.

CRISTIN
Current Research Information System In Norway

Search for projects, results and persons

Cristin result ID: 1674156

Last modified: February 6, 2019, 3:28 PM

RESULT

Academic monograph
2019

Designing Social Science Research

Oddbjørn Bukve

PUBLISHER/SERIES

PUBLISHER
Palgrave Macmillan
NVI-level 2

Find in the Norwegian register for publishers

ABOUT THE RESULT

Academic monograph
Year of publication: 2019
ISBN: 978-3-030-03978-3
Revised edition

LINKS

ORIA

Search in ORIA with
978-3-030-03978-3

CLASSIFICATION

FIELD OF EXPERTISE (NPI)
Field of expertise: Interdisciplinary Social
Sciences
- Academic discipline: Social Science

DESCRIPTION

★ ENGLISH

TITLE
Designing Social Science Research

Figure 1. Metadata of a scholarly monograph in social sciences. User interface of the Norwegian database Cristin. Screenshot made on November 29, 2020, from https://app.cristin.no/results/show.jsf?id=1674156

The terms 'research output' and 'social sciences' I approach pragmatically. While, by 'research output' I mean publications as well as other artefacts (e.g. critical editions, corpora, datasets, code) authored by researchers, with the term 'social sciences' I refer to those academic disciplines that are recognised as such in a given context, national or other. Thus, if a discipline like anthropology is considered as a discipline in the social sciences, then it does belong to social sciences even though in other contexts it may be assumed to belong to the humanities.

Thinking about research evaluation I use the constructivist framing of evaluation (Dahler-Larsen, 2011). For Dahler-Larsen (2011), evaluation is a "liminal" and "assisted" form of "sense making" (p.13). This definition foregrounds that evaluation is "a situation where we stop and reflectively consider our experiences in the midst of a specific social practice. . . . An evaluation is . . . an occasion for reassessment and reinterpretation. . . . Meanings emerge or are negotiated in evaluation" (Ibid.). The term 'liminal' emphasises that evaluations are something other than the everyday life; it refers to transition situations "where one is both inside and outside at the same

time" (p.14). And the term 'assisted' highlights the artificial nature of evaluation; they are constructed and made to exist. It is not natural to evaluate.

Following the constructivist trail of thought, I conceptualise databases for research output using insights from infrastructure studies (Bowker et al., 2010; Star, 1999; Star & Ruhleder, 1996). This body of knowledge views infrastructures broadly and *in relation* to practices that the infrastructure is infrastructure *of*. Thus, when databases for research output are used to evaluate research and when research evaluation is the practice that drives database implementation projects, these databases become part of research evaluation infrastructure. This infrastructure consists of databases, systems, software and hardware *and* of people, organisations and practices, priorities and value-orientations that characterise them. Research evaluation is what shapes this infrastructure and, inversely, this infrastructure gives shape, direction and sets limitations for research evaluation. Viewed through this more social and relational framing of infrastructure, we have to think about databases for research output in a much broader sense—i.e. in the wider context they are operated.

Combined with the ideas on research evaluation from Dahler-Larsen (2011), databases for research output give materiality to the otherwise purely conceptual space for the sense-making of research and its value. These databases form and play a central role in this 'knowing space' (Law, 2017) that "set[s] more or less permeable boundaries to the possible and the accessible" (p.47). These insights help us to stay close to the understanding of research evaluation as an occasion that enables reflections upon our experiences and the ways databases for research output facilitate, hinder, and give shape to this process. With this thought in mind, I proceed to the elaboration on the bright side of national databases for research output.

The bright side

Flexibility

In the context of the numerous calls for research assessment to pay due attention to the diversity of research, the main value of national databases is bound to the flexibility in database design and the organisation of the database implementation and maintenance. The choice of data to be collected, auxiliary data sources to be used, and people and organisations to involve in the operation of a national database is much greater for national databases, if contrasted with commercial international databases upon which users have typically very little influence (if any). Moreover, each design and organisation decisions can be made to suit the needs of the specific context and envisioned uses. Consequently, if deemed necessary, a national database can include a wide variety of research outputs produced by researchers in social sciences and other knowledge domains thus addressing the problem of coverage—the most often acknowledged problem with respect to commercial international data sources (Sivertsen & Larsen, 2012).

The scope of a national databases, due to the flexibility, can be designed considering the characteristics of research within a given context. Thus, now there are databases in Europe that not only collect information on journal articles that fall beyond the scope of Web of Science and Scopus, but also are broad in scope (e.g. RIV in the Czech Republic: https://www.rvvi.cz/riv, COBISS in Slovenia: https://cobiss.si/). For example, CRISTIN in Norway (https://app.cristin.no/) include more than 70 different output types using a classification developed in collaboration with scholars working in different knowledge domains.

Furthermore, this flexibility enables an alignment of database functionalities with the needs of different stakeholders operating at the national level and beyond. For research evaluation, this can mean that there are modules that enable automated overviews of research at different levels of aggregation (see, for example, BFI in Denmark: https://bfi.fi.dk/Publication/NationalAnalysis or SICRIS in Slovenia: https://www.sicris.si). In addition to research evaluation, national databases are often used to report research overviews to governmental and/or research funding organisations. National databases can be used to create these reports. In addition, national databases can be continuously adapted and adjusted as new requirements emerge.

This flexibility to create fit-for-purpose designs and broaden out the scope thus doing justice to the diversity within research is a key asset of a national database for research output in the social sciences. This is not to say that national legislation or simply local practices cannot be obstacles for more contextualised database designs. Nevertheless, the scope for flexibility is much more considerable when compared to commercial data sources.

Transparency

The second aspect I wish to highlight is transparency. Transparency is, first of all, openness about the way how a database is run (i.e. who does what) and what are the processes behind the data on research output. Second, transparency addresses the decision-making processes that feed the database practices. Who takes decisions? How a particular decision is justified? In case of alternative views, how are they accommodated and/or reconciled? Transparency of this type requires opening up the inner workings of a database for interested stakeholders. Third, transparency also means reflexivity about the conceptual underpinning of databases for research output (i.e. what image of research is embedded in a database?).

The different initiatives paving way for new forms of research assessment and responsible use of metrics (DORA, 2012; Hicks et al., 2015; Moher et al., 2020; Wilsdon et al., 2015) has highlighted the importance of transparency within research assessment practices. In relation to data for research assessment, for example, the Leiden Manifesto enlists the following principles: "Keep data collection and analytical processes open, transparent and simple" and "Allow those evaluated to verify data and analysis"(Hicks et al., 2015). Both these points emphasise the need transparency

and both highlight the close and direct relation between data and insights produced in research evaluation settings. Incomplete data will produce incomplete insights. The wrong choice of data will produce inappropriate accounts. Data of poor quality will lead to poor quality representations of research. Besides this, high quality data for one academic discipline are not necessarily data of quality also for other disciplines (see also Working group for the responsible evaluation of a researcher, 2020). In other words, regardless how well developed a database is, there will be limitations for its use. Thus, for research evaluation to serve as a space for reflection, these limitations need be taken into account and openly communicated. This is why transparency is an important aspect of databases for research output.

Transparency of processes and practices 'behind the scenes' can be achieved by making publicly available information about practices behind a national database. This can be in a form of a description of the processes carried out or an establishment of a communication venue between users and database administrators.

Transparency of decision-making similarly can be greatly enhanced by making publicly available the minutes of meetings where crucial decisions on database design are taken. Ideally, however, transparency is adopted as a general working mode that acknowledges that for every decision made there are multiple alternatives, some of which may have important consequences when these databases are used in research metrics. This stance requires continuous readiness to inform, clarify, openly discuss, and eventually alter every aspect of a national database for research output.

Finally, transparency at the conceptual level requires awareness and reflexivity upon the kind of research that is enacted within a (national) database for research output. Following the advent of citation analysis, both Edge (1979) and Woolgar (1991) argued that such methods imply a rationalised view of research that is at odds of research as it unfolds in practice. For social scientists in some research traditions, the very idea that research output can meaningfully depict research practices is hardly acceptable. If social scientists tend to be aware of these conceptual clashes between research as experienced in one's life and research as depicted in a database, the database designers and proponents tend to be ignorant of the performativity the national databases carry. In contrast, awareness and transparency of the conceptual underpinnings of a national database for research output is a necessary first step towards an open discussion between proponents of contrasting ideas about what research is and how it could be depicted in a digital form.

National databases with their more considerable flexibility are well positioned to be transparent. At the same time, there is a considerable variation in the extent to which transparency is embraced in the currently operational national databases (in Europe). At the moment, I note, transparency of national databases is more a *potential* rather than an existing asset.

The dark side

Too national

Both points highlighted above can be flipped around; both flexibility and transparency can become risks when some of the directions are taken too far or their implementation clashes with existing institutionalised practices. All information systems carry a tension between *global* and *local* (Bowker, 2008). On the one hand, the intention is to be *local*, which means context-specific, understandable, and consequently, easy to work with, for those who engage with these systems locally. If, for example, anthropology is seen as part of the disciplines in the humanities, it will be easier to use a national database if this understanding of anthropology is mirrored also in the disciplinary classifications employed in a national database. On the other hand, there is the requirement to be *global* so that information contained can be related to and transferred across multiple local contexts. An example of this in relation to disciplinary classifications, is the tendency to implement the OECD Fields of Research and Development classification (OECD, 2015) in national databases for research output. The use of this international standard contributes to the interoperability of different national databases and to rendering them understandable across national borders. Yet at the same time it can mean that local versions of the OECD FoRD classification are created (Guns et al., 2018). The setup of an information system is the result of this global-local tension.

For national databases for research output in general terms this means that the attempt to take into account specifics of social sciences in a certain country, for example, can lead to a database that is too closely bound to local traditions and power relations. A rich example is the variety of classifications of research output types that sometimes include outputs not common elsewhere. In some countries, these classifications include outputs as diverse as maps, critical editions, and textbooks. The reasons for their inclusion might be absolutely clear locally, but without additional elaboration of their role in the local research practices, some of the research output categories can be puzzling for outsiders. This local-global tension becomes a political matter when a debate takes a more normative direction and cast doubt to the scientific status of some of these artefacts, and the worth they should be attributed in a research evaluation context. Can maps be regarded as research outputs? Are textbooks and reports to government research outputs? Are essays in newspapers research outputs? Is the process of organising a conference a research output? Answers to these questions depend on one's views of research. As soon as we are reminded that there are multiple views on research, answers to these questions acquire a normative tone. Whose views ought to be followed? Who has the right to decide which views will be translated in database designs? If multiple views will be taken into account, how will they be reconciled in database design? These questions can, certainly, be opened for a discussion. At the same time, it is possible that the resulting compromises are too local and hard to comprehend internationally.

The same applies for transparency. When national databases are intertwined with local power struggles, the database easily can become opaque and inaccessible. Decisions

about the database design then tend to be taken behind closed doors and queries about data collection and data processing practices remain unanswered. It is a challenge to keep a national database transparent when the environment in that particular context does not facilitate such a direction. Thus, the key values of national databases—flexibility and transparency—can be their weakness when the contexts where they operate hinders the process of making the best use of these databases.

Research output is just one side of the story

The last line of argumentation requires some zooming out. To remind, databases are best seen in relational terms as noted above. This framing keeps at its centre the acknowledgement that databases are not neutral representations of research *out there*. Instead, the kinds of representation we acquire with the help of national databases, result from a constellation of a myriad of social, cultural, political, and, surely, also technical considerations and negotiations. Databases mirror the circumstances in which they run. The next key question when databases are seen this way is: how is research seen through national databases? Here I will highlight a key characteristic of representations made using data from (national) databases: representations based on *research output*.

Since the rise of the New Public Management in 1990s, concepts like 'performance', 'indicators', 'targets' have become the buzzwords of research assessment. Along with this new vocabulary, comes the view of research as a production-like activity that can be described and measured in terms of inputs and outputs and metaphorically described as 'McDonaldisation' of higher education (Nadolny & Ryan, 2015; Parker & Jary, 1995) and economisation of research (Hallonsten, 2021). Consequently, predominantly quantitative approaches have found their ways into funding allocation mechanisms, institutional evaluation procedures and the sidelines of decision-making around research grant proposals and academic promotion (Curry et al., 2020; Hammarfelt & Rushforth, 2017; Jappe, 2020; Jonkers & Zacharewicz, 2016).

For social sciences just as in other knowledge domains, the discontent has often been focused on the usage of data that does not do justice to the richness of research in this knowledge domain—the insufficient coverage of data. National databases, as argued above, are then seen as the remedy for this problem since a wide variety of research outputs can be included. At the same time, reasoning of this type nevertheless is focused on research output. To get to the core of the limits of databases for research output, we have to return to the questions about the image of research that is produced when we limit our knowing of research to databases of research output. And this last point applies to databases of any scope, institutional, national, disciplinary, or global. There are social scientists who feel at ease with the view that research is a linear and clearly demarcated process. Research as a step-by-step process that follows graphs, tables, and overviews detailed in research grant proposals. Research that is isolated from students, activists, journalists and anyone else who might have a role to play in determining which route research might take. In some

disciplines, research may indeed lend itself to this linear model and the questions on output of this research would be obvious and easy to answer. The story becomes quite different as soon as we move to the more interpretative or conceptual side of research in the social sciences where it is much harder to disentangle research from all the other activities scholars pursue. Where is the line between research and teaching, discussion, thinking, or simply living? Am I doing research when I plan an academic workshop? Am I doing research when telling about my research to my friends over dinner (or is this already about societal impact?)? Does doing research includes also the thinking I do about my research when going for a walk? These questions have been raised ever since the databased thinking of research has found its way into knowing spaces for research and science policy. Scholars in higher education studies have been vocal for decades about the problematic sides of overly categorical, technicist, instrumental and typically enumerable ways of thinking about what counts as research productivity and output. And yet, research assessment procedures turn to quantitative measures, and to that end, new databases are being implemented.

National databases for research output are much more than just digital containers for traces of research. They are conceptual devices or knowing spaces. As soon as large investments have been made to set up a national database and various reward mechanisms have been linked up with the data, the ideas about research that the national databases carry, has the potential to become *the* way to think about research. This is less problematic when a database is run in a continuous collaboration with all relevant stakeholders. In contrast, when the design is based on views on one, typically the more powerful, group, there is a risk that it is hardly possible to use a national database to produce meaningful insights in research across the whole range of its diverse forms and practices.

Future directions

In this paper I have offered a discussion of two positive and two negative characteristics that are key for making the best use of national database for research output in the assessment of research in social science. While some of them have been actively discussed in the literature around national databases (for example, flexibility), others have been less explicitly present or altogether absent from the current debates. The points discussed here are by no means exhaustive; they are also not based on a systematic analysis. Instead, this is a selection of ideas that are in a need for discussion on new routes for national databases that facilitate evaluation of research in social sciences. In this last section of the paper, I offer a number of such directions.

Over the last decades, debates in research evaluation have moved beyond the simplistic and quantitative focus on research performance and scholarly impact that dominated in 1980s and 1990s. Discussions has branched out to include societal impact and take into account specifics such as interdisciplinary research as shown by several chapters in this handbook. There is also increasing awareness of the limits that quantified insights of the value of research offer. As a solution, a number of scholars

have proposed new modes of thinking about research evaluation. For example, de Rijcke and colleagues (de Rijcke et al., 2019), introducing the framework 'evaluative inquiry', call for plural answers to questions on *what* should be evaluated and *how*; pursuing evaluative inquiry, key concepts are developed in process, using a variety of means (e.g. data, consultations, document analysis) and in collaboration with stakeholders (see also Fochler & de Rijcke, 2017). Along the same lines, Rafols and Stirling (2020) elaborate a framework that allows to employ research evaluation methods that *open up* and not *close down* debates on pressing issues. Similarly, Marres and de Rijcke (2020), discussing interdisciplinarity, emphasise the need to move from indicators to *indicating*, a process-focused approach to understanding research (and its value).

All these new directions can be related also to national databases for research output. The use of national databases does not need to be equated with metrics and approached conceptually by the New Public Management; tools and approaches developed in sociological, historical or literary studies (among else) serve as examples on alternative uses of bibliographic data in knowledge making (e.g., Sapiro & Bustamante, 2009). One direction is to think of new quantitative ways to depict research activities. Another is to use qualitative or mixed methods to make sense of research using data from national databases for research output.

It is undeniable that national databases for research output can facilitate assessment of research in the social sciences. The more comprehensive coverage of outputs in social sciences, the typical advantage of national databases over international databases, alone is a strong argument. The closer national databases for research output are to the research practices, the more nuanced evaluation can take place (see also Oancea, 2019). National databases have the potential to be flexible and transparent—the bright side of databases. Both these characteristics are preconditions for research evaluation that is meaningful as well as acceptable in a particular national context. At the same time, it is also important to take into account the less positive side of national databases—the dark side. There is a risk for national databases to be too close to local traditions and power struggles that hinder the employment of the full potential national databases can offer. Also, it should not be forgotten that national databases shape the way research is seen and consequently, evaluated. All these points together, although not exhaustive, offer points to focus on when thinking on the best ways national databases for research output can serve research assessment in social sciences.

References

Barré, R. (2019). Les indicateurs sont morts, vive les indicateurs! Towards a political economy of S&T indicators: A critical overview of the past 35 years. *Research Evaluation*, *28*(1), 2–6. https://doi.org/10.1093/reseval/rvy029

Bloor, D. (1976). *Knowledge and social imagery*. Routledge & K. Paul.

Bowker, G. C. (2008). *Memory practices in the sciences* (1. paperback ed). MIT.

Bowker, G. C., Baker, K., Millerand, F., & Ribes, D. (2010). Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment. In J. Hunsinger, L. Klastrup, & M. Allen (Eds.), *International Handbook of Internet Research* (pp. 97–117). Springer Netherlands. https://doi.org/10.1007/978-1-4020-9789-8_5

Chi, P.-S. (2014). Which role do non-source items play in the social sciences? A case study in political science in Germany. *Scientometrics*, *101*(2), 1195–1213. https://doi.org/10.1007/s11192-014-1433-1

Crossick, G. (2015). *Monographs and Open Access. Report to HEFCE* (p. 77). HEFCE. https://webarchive.nationalarchives.gov.uk/20170712122802/http://www.hefce.ac.uk/media/hefce/content/pubs/indirreports/2015/Monographs,and,open,access/2014_monographs.pdf

Curry, S., Rijcke, S. D., Hatch, A., Pillay, D., Weijden, I. V. D., & Wilsdon, J. (2020). *The changing role of funders in responsible research assessment: Progress, obstacles and the way ahead* (p. 2449096 Bytes). Research on Research Institute. https://doi.org/10.6084/M9.FIGSHARE.13227914

Dahler-Larsen, P. (2011). *The Evaluation Society*. Stanford University Press.

de Rijcke, S., Holtrop, T., Kaltenbrunner, W., Zuijderwijk, J., Beaulieu, A., Franssen, T., van Leeuwen, T., Mongeon, P., Tatum, C., Valkenburg, G., & Wouters, P. (2019). Evaluative Inquiry: Engaging research evaluation analytically and strategically. *Fteval Journal for Research and Technology Policy Evaluation*, *48*, 176–182. https://doi.org/10.22163/fteval.2019.386

de Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2015). Evaluation practices and effects of indicator use—A literature review. *Research Evaluation*, *25*(2), 161–169. https://doi.org/10.1093/reseval/rvv038

DORA. (2012). *San Francisco Declaration on Research Assessment*. https://sfdora.org/read/

Edge, D. (1979). Quantitative Measures of Communication in Science: A Critical Review. *History of Science*, *17*(2), 102–134. https://doi.org/10.1177/007327537901700202

Engels, T. C. E., Istenič Starčič, A., Kulczycki, E., Pölönen, J., & Sivertsen, G. (2018). Are book publications disappearing from scholarly communication in the social sciences and humanities? *Aslib Journal of Information Management*, *70*(6), 592–607. https://doi.org/10.1108/AJIM-05-2018-0127

Fochler, M., & de Rijcke, S. (2017). Implicated in the Indicator Game? An Experimental Debate. *Engaging Science, Technology, and Society*, *3*, 21–40. https://doi.org/10.17351/ests2017.108

Glänzel, W., Thijs, B., & Chi, P.-S. (2016). The challenges to expand bibliometric studies from periodical literature to monographic literature with a new data source: The book citation index. *Scientometrics*, *109*(3), 2165–2179. https://doi.org/10.1007/s11192-016-2046-7

Guns, R., Sīle, L., Eykens, J., Verleysen, F. T., & Engels, T. C. E. (2018). A comparison of cognitive and organizational classification of publications in the social sciences and humanities. *Scientometrics*, *116*(2), 1093–1111. https://doi.org/10.1007/s11192-018-2775-x

Hallonsten, O. (2021). Stop evaluating science: A historical-sociological argument. *Social Science Information*, 053901842199220. https://doi.org/10.1177/0539018421992204

Hammarfelt, B., & Rushforth, A. D. (2017). Indicators as judgment devices: An empirical study of citizen bibliometrics in research evaluation. *Research Evaluation*, *26*(3), 169–180. https://doi.org/10.1093/reseval/rvx018

Hicks, D. (2004). The Four Literatures of Social Science. In H. F. Moed, W. Glänzel, & U. Schmoch (Eds.), *Handbook of Quantitative Science and Technology Research* (pp. 473–496). Kluwer Academic Publishers.

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, *520*, 429–431. https://doi.org/10.1038/520429a

Jappe, A. (2020). Professional standards in bibliometric research evaluation? A meta-evaluation of European assessment practice 2005–2019. *PLOS ONE*, *15*(4), e0231735. https://doi.org/10.1371/journal.pone.0231735

Jonkers, K., & Zacharewicz, T. (2016). *Research Performance Based Funding Systems: A Comparative Assessment* (JRC Science For Policy Report). European Commission: Joint Research Centre. https://rio.jrc.ec.europa.eu/en/file/9514/download?token=-8JG6aKx

Kulczycki, E., Engels, T. C. E., Pölönen, J., Bruun, K., Dušková, M., Guns, R., Nowotniak, R., Petr, M., Sivertsen, G., Istenič Starčič, A., & Zuccala, A. (2018). Publication patterns in the social sciences and humanities: Evidence from eight European countries. *Scientometrics*, *116*(1), 463–486. https://doi.org/10.1007/s11192-018-2711-0

Law, J. (2017). STS as Method. In U. Felt, R. Fouché, C. A. Miller, & L. Smith-Doerr (Eds.), *The handbook of science and technology studies* (Fourth edition, pp. 31–57). The MIT Press.

Mahieu, B., Arnold, E., & Kolarz, P. (2014). *Measuring scientific performance for improved policy making. Summary of a study* (p. 16). European Parliament; Directorate-General for Parliamentary Research Services. http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/527383/IPOL-JOIN_ET(2014)527383_EN.pdf

Marres, N., & de Rijcke, S. (2020). From indicators to indicating interdisciplinarity: A participatory mapping methodology for research communities in-the-making. *Quantitative Science Studies*, *1*(3), 1041–1055. https://doi.org/10.1162/qss_a_00062

Martin, B., Tang, P., Morgan, M., Glänzel, W., Hornbostel, S., Lauer, G., Lenclud, G., Lima, L., Oppenheim, C., van de Besselaar, P., & Žic-Fuchs, M. (2010). *Towards a Bibliometric Database for the Social Sciences and Humanities—A European Scoping Project* (p. 55). https://globalhighered.files.wordpress.com/2010/07/esf_report_final_100309.pdf

Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & López-Cózar, E. D. (2020). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations. *ArXiv:2004.14329 [Cs]*. http://arxiv.org/abs/2004.14329

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., Coriat, A.-M., Foeger, N., & Dirnagl, U. (2020). The Hong Kong Principles for assessing researchers: Fostering research integrity. *PLOS Biology*, *18*(7), e3000737. https://doi.org/10.1371/journal.pbio.3000737

Nadolny, A., & Ryan, S. (2015). McUniversities revisited: A comparison of university and McDonald's casual employee experiences in Australia. *Studies in Higher Education*, *40*(1), 142–157. https://doi.org/10.1080/03075079.2013.818642

Oancea, A. (2019). Research governance and the future(s) of research assessment. *Palgrave Communications*, *5*(1), 27. https://doi.org/10.1057/s41599-018-0213-6

OECD. (2015). *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development,*. OECD Publishing. https://doi.org/10.1787/9789264239012-en

Parker, M., & Jary, D. (1995). The McUniversity: Organization, Management and Academic Subjectivity. *Organization*, *2*(2), 319–338. https://doi.org/10.1177/135050849522013

Rafols, I. (2017). *S&T Indicators in the Wild Acting with indicators in an uncertain world*. Science, Technology an Innovations Indicators -STI 2017, Paris, France. http://www.ingenio.upv.es/sites/default/files/adjuntos_noticias/rafols-parissti2017.pdf

Rafols, I., & Stirling, A. (2020). *Designing indicators for opening up evaluation. Insights from research assessment* [Preprint]. SocArXiv. https://doi.org/10.31235/osf.io/h2fxp

Ribeiro, L., de Castro, P., & Mennielli, M. (2016). *Final report: EUNIS-EuroCRIS Joint Survey on CRIS and IR*. http://digitalcommons.unl.edu/scholcom/5/

Sapiro, G., & Bustamante, M. (2009). Translation as a Measure of International Consecration. Mapping the World Distribution of Bourdieu's Books in Translation. *Sociologica*, *2–3*, 1–46. https://doi.org/10.2383/31374

Science Europe. (2016). *Position Statement on Research Information Systems*. https://www.scienceeurope.org/media/qbziuyj2/se_positionstatement_ris_web.pdf

Sīle, L., Guns, R., Sivertsen, G., & Engels, T. C. E. (2017). *European Databases and Repositories for Social Sciences and Humanities Research Output* (p. 25). ECOOM & ENRESSH. https://doi.org/10.6084/m9.figshare.5172322.v2

Sīle, L., Guns, R., Zuccala, A. A., & Engels, T. C. E. (2021). Complexity-Sensitive Book Metrics for Monographs in National Databases for Research Output. *Journal of Documentation*, https://doi.org/10.1108/JD-06-2020-0107

Sīle, L., Pölönen, J., Sivertsen, G., Guns, R., Engels, T. C. E., Arefiev, P., Dušková, M., Faurbæk, L., Holl, A., Kulczycki, E., Macan, B., Nelhans, G., Petr, M., Pisk, M., Soós, S., Stojanovski, J., Stone, A., Šušol, J., & Teitelbaum, R. (2018). Comprehensiveness of national bibliographic databases for social sciences and humanities: Findings from a European survey. *Research Evaluation*, *27*(4), 310–322. https://doi.org/10.1093/reseval/rvy016

Sivertsen, G. (2010). A performance indicator based on complete data for the scientific publication output at research institutions. *ISSI NEWSLETTER*, *6*(1), 22–28.

Sivertsen, G., & Larsen, B. (2012). Comprehensive bibliographic coverage of the social sciences and humanities in a citation index: An empirical analysis of the potential. *Scientometrics*, *91*(2), 567–575. https://doi.org/10.1007/s11192-011-0615-3

Star, S. L. (1999). The Ethnography of Infrastructure. *American Behavioral Scientist*, *43*(3), 377–391. https://doi.org/10.1177/00027649921955326

Star, S. L., & Ruhleder, K. (1996). Steps toward an ecology of infrastructure: Design and access for large information spaces. *Information Systems Research*, *7*(1), 111–134. https://doi.org/10.1287/isre.7.1.111

Verleysen, F. T., & Weeren, A. (2016). Clustering by publication patterns of senior authors in the social sciences and humanities. *Journal of Informetrics*, *10*(1), 254–272. https://doi.org/10.1016/j.joi.2016.01.004

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I., Wouters, P., Hill, J., & Johnson, B. (2015). *The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management*. https://doi.org/10.13140/RG.2.1.4929.1363

Woolgar, S. (1991). Beyond the citation debate: Towards a sociology of measurement technologies and their use in science policy. *Science and Public Policy*, *18*(5), 319–326.

Working group for the responsible evaluation of a researcher. (2020). *Good practice in researcher evaluation. Recommendation for the responsible evaluation of a researcher in Finland*. The Committee for Public Information (TJNK) and Federation of Finnish Learned Societies (TSV). https://doi.org/10.23847/isbn.9789525995282