

An institutional implementation of the new European reform of research assessment

Kathrine Bjerg Bennike*, Poul Meier Melchiorson**, and Gunnar Sivertsen***

* kbb@aub.aau.dk

0000-0002-3498-2677

VCN Team, Aalborg University Library, Aalborg, Denmark

** pmm@aub.aau.dk

0000-0002-3751-9821

VCN Team, Aalborg University Library, Aalborg, Denmark

*** gunnar.sivertsen@nifu.no

0000-0003-1020-3189

Nordic Institute for Studies in Innovation, Research and Education (NIFU), Oslo, Norway

Abstract

The Agreement on Reforming Research Assessment (Coalition for Advancing Research Assessment, 2022) addresses the assessment of researchers, research proposals, and research performing organizations and units. Our contribution is relevant for the assessment of units within organizations. We describe how the main principles and requirements of the Agreement have inspired a new institutional-level framework at Aalborg University (DK) with the aim to deliver data, statistics, and indicators to support internal strategic development, research assessment and resource allocation. The framework was designed on request from the rectorate and in collaboration with the deans of the four faculties with feedback from their project steering group and the university's representative bodies.

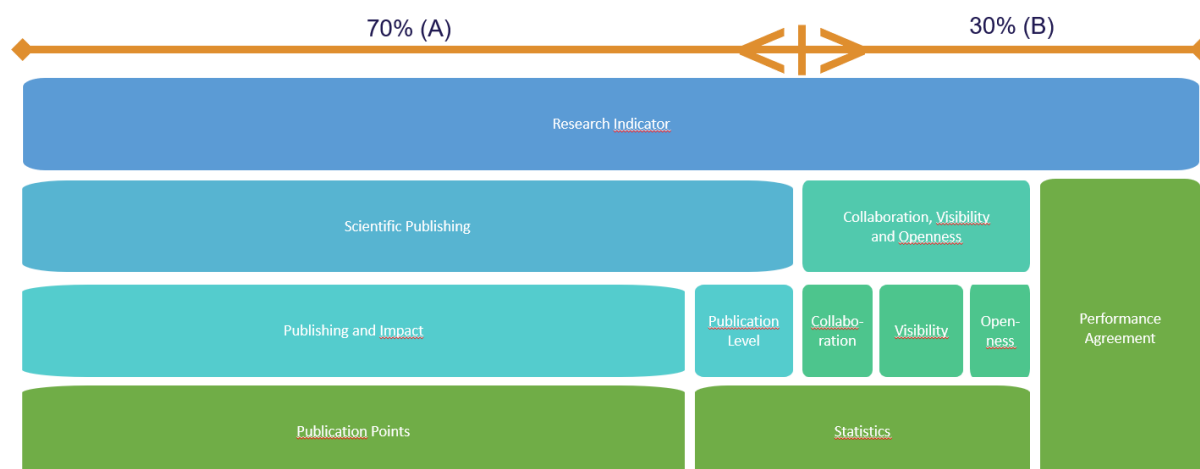
1. Introduction

We describe an institutional-level framework with the aim to deliver data, statistics, and indicators to support internal strategic development, research assessment and resource allocation within a university, in our case Aalborg University in Denmark. The units of assessment are the faculties and their departments, not the individual researchers.

The new framework is inspired by the Agreement on Reforming Research Assessment (Coalition for Advancing Research Assessment, 2022 – hereinafter: the Agreement), which seeks to “induce a research culture that recognises **collaboration, openness, and engagement with society**” by recognizing “the **diverse** outputs, practices and activities that maximise the **quality and impact** of research”. The words thereby emphasized indicate the main elements of the framework as visualized in *Figure 1*.

The left-hand side with “Publication points” represents a bibliometric indicator that balances the measurement of publication output and citation impact across all fields. It is comparable and may serve internal resource allocation, and it is designed to ensure “responsibly used quantitative indicators where appropriate”, as called for in the Agreement. The right-hand side of *Figure 1* represents the broader range of activities and competences called for in the Agreement. Here, the emphasis is on societal interaction: collaboration with society, visibility in society, and open research practices. We start by explaining this latter part of the framework.

Figure 1. The Aalborg Framework to promote research quality, collaboration, openness, and engagement with society.



2. Statistics on societal interactions and open research practices

Societal interaction (often called ‘impact’) is not easily measured or documented (Bornmann, 2013; Sivertsen & Meijer, 2020; Thelwall, 2021). Open research practices have the same problem. Examples and narratives need to be based on evidence (Hill, 2016). Institutional or national research information systems are promising solutions to this problem (Sivertsen, 2019; Sivertsen & Rushforth, 2022). Aalborg University applies *Pure*, an institutional research information system originally developed on our campus and then commercialized by Elsevier for the world market.

Pure is designed to register a broad range of research activities among the employees in categories prioritized by their departments. Our new framework prioritizes relevant categories for the documentation of societal interaction and open research. We have also extended the system by establishing a quality-assured database of media coverage of the university’s achievements in research as well as a database of patents originating from the university.

It must be recognized, however, that even when partly solving the documentation problem, it is difficult to provide *comparable* indicators of societal interaction and open research at department level. Media coverage and patenting, for example, vary among fields and neither of them may provide the complete picture of societal interaction (Thelwall, 2021). With whom, how, and for what purposes societal interaction can be expected to occur, also basically differs among fields of research (Sivertsen & Meijer, 2020). For this reason, the term ‘*statistics*’, not ‘*indicators*’, is used on the right-hand side in *Figure 1*. Statistics may inform strategy development, performance agreements and annual reports at the department level without being comparable across units.

The Agreement on Reforming Research Assessment uses the terms ‘*indicators*’ and ‘*metrics*’ throughout, but not the term ‘*statistics*’. With the Agreement’s ambition to guide organizational-level as well as individual-level assessments, the term is probably needed (Sivertsen & Rushforth, 2022).

3. The bibliometric indicator

The left-hand side with “Publication points” represents a further development of the bibliometric indicator that served national and internal performance-based resource allocation in Denmark until 2022 (Aagaard, 2018). Unlike most other bibliometric indicators, but similarly to its Finnish (Pölönen, 2018) and Norwegian (Sivertsen, 2018) counterparts, the Danish bibliometric indicator was developed and maintained in collaboration with the academic communities, it was inclusive in fully representing all scholarly publication formats and languages used in all fields of research, and it was balanced by providing a comparable measurement of research activity across all fields. It was, however, criticised for not covering citation impact, and, with reference to the San Francisco Declaration on Research Assessment (DORA, 2012), for focusing on the publication channel instead of the value of the individual publication.

Our solution introduces citation impact to the publication indicator and removes the focus on the venue of the publication. The solution is based on a proposal first presented at the STI Conference in Valencia in 2016 (Sivertsen, 2016) and in a report to the Danish Ministry of Research and Education (Uddannelses- og forskningsministeriet, 2019). Field-normalized citations counts (Waltman & van Eck, 2015) are added with an influence on the score reflecting the degree of coverage of the publications of the field in the citation database. The indicator is still inclusive by covering all scholarly publication practices in all fields. The method for calculating the contributions of co-authors ensures comparability across all fields of research (Sivertsen et al., 2019). The construction of the indicator makes it applicable only at the department level or at higher levels of aggregation. For technical reasons, it is not possible to calculate the indicator at the level of individual researchers. Hence, the indicator is designed to ensure “responsibly used quantitative indicators where appropriate”, as called for in the Agreement.

4. A modification of existing practices

The new framework does not impose fundamental changes on Aalborg University. For several reasons, the inspiration from the Agreement has been welcomed throughout the organization, and there has been consensus about the framework so far in the implementation process.

The vision of the Agreement “is that the assessment of research, researchers and research organisations recognises the diverse outputs, practices and activities that maximise the quality and impact of research”. This vision corresponds well to the self-defined official aims of Aalborg University as “a mission-oriented university”: “Through partnerships with the outside world, we strive to contribute to missions where we have strong research insight. We strive to play an active role in finding solutions to current – and future – global challenges and make a difference to the world around us.”

Aalborg University (as the other seven Danish universities) receives a relatively high proportion of block funding which also includes resources for research to be spent internally. The degree of autonomy is also relatively high. How to further allocate the resources from the rectorate via the faculties down to the departments is up to the university to decide. The allocation is partly based on indicators, partly on performance agreements, and partly on strategic priorities decided by formally responsible bodies at different levels within the university. The new framework does not change these procedures. With the inspiration from the Agreement, it only modifies the existing procedures by further developing the bibliometric indicator so far used for reallocation, and by providing new sources and forms of statistics for performance agreements and strategic development.

None of the modified annual procedures for resource allocation and strategic development at the organizational level will include peer review by external experts. This decision is mainly practical, not principal, but it could be perceived as contradicting the second of the four core commitments of the Agreement: “Base research assessment primarily on qualitative evaluation for which peer review is central, supported by responsible use of quantitative indicators.” It is normal practice in the Nordic countries to follow this principle in all contexts where individual researchers are applying for recruitment or promotion and in all contexts where there is an assessment of individual applications for external funding.

At the organizational level, however, the tradition in the Nordic countries is to use research assessment based on external peer review for *formative* purposes only without implications for resource allocation. The funding of organisations is instead based on *summative* assessments with the use of performance indicators, other explicit criteria, and performance agreements (Sivertsen, 2023). The Nordic countries thereby differ from the United Kingdom’s unique combination of research assessment and resource allocation in the Research Excellence Framework (Sivertsen, 2017). There is need for further discussion of how the second of the four commitments in the Agreement applies to summative versus formative research assessment at the *organizational* level (Sivertsen & Rushforth, 2022).

The new framework at Aalborg University is an early implementation of the Agreement which can be expected to create experiences and discussions about how the core principles will work in practice. Some early-stage reflections will conclude this paper.

5. Bibliometric measurement

In the process of measuring research, some elements of research practice are lifted out of context and examined while other elements are overlooked. James Vincent writes in his book on the history of measurement: “Measurement is a tool that reinforces what we find important in life, what we think is worth paying attention to” (Vincent, 2022; p. 19). On the other hand, measurement and evaluation can be experienced as “not a noble presence but a repressive tyrant” (p. 18). Excessive measurement and management are also well described by Berg & Seeber (2016) in the book “The Slow Professor”, but mostly with relevance for the individual level which we avoid by designing a framework for the department level and an indicator that cannot be used below this level.

Still, measurement is a tool of control which imposes limits on life. Excessive measurement and inappropriate measurement (Vincent, 2022, p. 321) are the problems now addressed by the Agreement. It calls for the abandonment of “inappropriate uses in research assessment of journal- and publication-based metrics” in the third core commitment. How does this principle apply on the Aalborg framework?

As explained in section 3, inspired not only by the Agreement, but also by DORA, our solution introduces citation impact to the publication indicator and removes the focus on the venue of the publication, thereby changing the weight from a journal-based to a publication-based indicator. The level of the journal (as assessed by expert panels, not by Journal Impact Factors) is no longer part of the measurement and only provided as statistical information for strategic development in the right-hand side of *Figure 1*. There was agreement on the continuing need for this information, not for assessing individual publications or researchers, but for assessing the varying quality of journals and their editorial procedures. At the same time, we have developed an indicator for direct measurement and comparison which is inclusive and balanced.

We leave it open for discussion whether there is *inappropriate* use of journal- and publication-based metrics in the new framework. Our model is dynamic and may be adjusted if needed. The Agreement encourages that practises are evaluated when new insights and evidence is present. This principle is also laid out in another guide to research evaluation, the Leiden Manifesto: “Scrutinize indicators regularly and update them” (Hicks et al, 2015).

6. Local implementation of an international agreement

DORA, the Leiden Manifesto, and recently the Agreement, are examples of international initiatives and agreements aimed at changing research assessment practices in local contexts. Such agreements often rely on broader normative ideas and ambitions (Finnemore and Sikkink 1998; Zwingel 2017), in this case, e.g., the focus of open science, the FAIR principles, and the sharing of knowledge (European Commission. Directorate General for Research and Innovation. 2017). The Agreement’s focus on alternative methods and a broader diversity of outputs demonstrates the EU’s ambition to challenge the use of solely quantitative indicators and in general to introduce a more responsible use of metrics in research assessments. However, as indicated by Sivertsen & Rushford (2022), the more concrete implications of the normative requirements of the Agreement still need to be developed, and the translation of the norms from the supranational to the local level might be challenging.

Finnemore and Sikkink (1998) argue that “*new norms never enter a normative vacuum but instead emerge in a highly contested normative space where they must compete with other norms and perceptions of interest*” (Finnemore and Sikkink 1998, 989). In other words, normative contestation or resistance is important for how we relate the “logic of appropriateness” to norms and the creation of new norms. Hence, to promote a new norm, the process needs to “*take place within the standards of “appropriateness” defined by prior norms*” (Finnemore and Sikkink 1998, 989). Further, to challenge existing logics of appropriateness, the new normative ideas may even need to be explicitly “inappropriate”.

The argument could be turned to cover the opposite direction of implementation: If accepted and implemented at Aalborg University, would it be possible to use the ideas and practices of the new framework (and the inspiration from the Agreement) for an implementation at a national level where comparable research assessments are needed as well? Could the local applicability of the framework be translated to a more aggregated level? After the governmental dismissal of the Danish bibliometric indicator at the national level, there is still need for legitimate criteria for funding allocation and for new sources and forms of statistics that cover a broader array of research activities including societal interaction and open research.

Again, relying on norm translation theory, the answer to the question probably depends on the acceptance at the national level of normative ideas that were originally translated from the supranational level where their development was supported by the European Commission, Science Europe, and the European University Association. Danish organizations are active at the European level and will readily recognize the Agreement’s main principles of broadening the basis for research assessment and introducing responsible use of indicators. The fact that the new framework has been received and implemented at Aalborg University without much resistance is an indication that a more general organizational-level implementation of the Agreement is within reach.

7. Conclusion

As pointed out by Sivertsen & Rushforth (2022), the Agreement’s approach to research assessment at the organizational level is so far less developed than for the contexts for

individual level assessment. Statistical information and indicators are addressed by the Agreement with scepticism, but they can be more adequate at the organizational level than at the individual level. A careful use of indicators and statistics can even be regarded as more responsible and transparent than decisions taken in closed rooms. The Aalborg framework might be an early test of how the Agreement can be further developed in this direction.

Author contributions

All three authors collaborated on all parts of the paper. We also worked together on the development of the new Aalborg framework and on the report in Danish which is presently the basis for the local discussion, decision, and implementation process. We have no competing interests in the development of the framework or in writing the final report or this paper.

Open science practices

The framework was designed on request from the rectorate and in collaboration with the deans of the four faculties with feedback from their project steering group and the university's representative bodies.

References

Aagaard, K. (2018). Performance-based Research Funding in Denmark: The Adoption and Translation of the Norwegian Model. *Journal of Data and Information Science*, 3(4), 20–30.

Berg, M., & Seeber, B. K. (2016). *The Slow Professor: Challenging the Culture of Speed in the Academy*. University of Toronto Press

Bornmann, L. (2013). What is societal impact of research and how can it be assessed? a literature survey. *Journal of the Association for Information Science and Technology*, 64: 217–233.

Coalition for Advancing Research Assessment. 2022. “Agreement on Reforming Research Assessment”. The European Commission. <https://coara.eu/agreement/the-agreement-full-text/>.

DORA. 2012. “San Francisco Declaration on Research Assessment”. Declaration on Research Assessment. <https://sfdora.org/read/>

Finnemore, M. & Sikkink K. (1998). “International Norm Dynamics and Political Change.” *International Organization* 52 (4): 887–917. doi.org/10.1162/002081898550789.

Hicks, D., Wouters, P., Waltman, L. et al. (2015) Bibliometrics: The Leiden Manifesto for research metrics. *Nature* 520, 429–431. doi.org/10.1038/520429a

Hill, S (2016). Assessing (for) impact: future assessment of the societal impact of research. *Palgrave Communications*. 2:16073. doi: 10.1057/palcomms.2016.73.

Pölönen, J. (2018). Applications of, and Experiences with, the Norwegian Model in Finland. *Journal of Data and Information Science*, 3(4), 31–44.

Sivertsen, G. (2016). A bibliometric indicator with a balanced representation of all fields. In Ràfols, I., Mollas-Gallart, J., Castro-Martínez, E., Woolley, R. (Eds.), *Proceedings of the 21ST*

International Conference on Science and Technology Indicators (910-914). Valencia: Editorial Universitat Politècnica de València.

Sivertsen, G. (2017). Unique, but still best practice? The Research Excellence Framework (REF) from an international perspective. *Palgrave Communications*, 3: 17078, doi:10.1057/palcomms.2017.78

Sivertsen, G. (2018). The Norwegian Model in Norway. *Journal of Data and Information Science*, 3(4), 2-18.

Sivertsen, G. (2019). Developing Current Research Information Systems (CRIS) as Data Sources for Studies of Research. In Glänzel, W., Moed, H.F., Schmoch, U., Thelwall, M. (Eds.), *Springer Handbook of Science and Technology Indicators* (pp. 667-683). Cham: Springer.

Sivertsen, G., Rousseau, R., Zhang, L. (2019). Measuring Scientific Production with Modified Fractional Counting. *Journal of Informetrics*, 13(2): 679-694.

Sivertsen, G., & Meijer, I. (2020). Normal versus extraordinary societal impact: how to understand, evaluate, and improve research activities in their relations to society? *Research Evaluation*, 66–70.

Sivertsen, G. & Rushforth, A. (2022). The new European reform of research assessment. R-QUEST Policy Brief no. 7 <https://www.r-quest.no/policy-briefs/>

Sivertsen, G. (2023). Performance-based research funding and its impacts on research organizations. In: Lepori B., Jongbloed B., and Hicks D. (Eds). *Handbook of Public Research Funding*, Edward Elgar Publishing, pp. 90-107.

Thelwall, M. (2021). Measuring societal impacts of research with altmetrics? Common problems and mistakes. *Journal of Economic Surveys*, 35: 1302-1314.

Uddannelses- og Forskningsministeriet. 2019. Fremtidssikring af forskningskvalitet - Ekspertudvalget for resultatbaseret fordeling af basismidler til forskning. København K.: Uddannelses- og Forskningsministeriet.

Vincent, J (2022). *Beyond Measure: The Hidden History of Measurement*. London. Faber & Faber Limited

Waltman, L., van Eck, N.J. (2015). Field-normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Informetrics*, 9(4), 872-894.

Zwingel, S. (2017). “Women’s Rights Norms as Content-in-Motion and Incomplete Practice.” *Third World Thematics: A TWQ Journal* 2 (5): 675–90. doi.org/10.1080/23802014.2017.1365625.