

Guest Editorial

New Developments In Science Publishing In Official Statistics (Open Data, FAIR Publishing, And The Challenges For Science Publishing To Stay Relevant)

Pieter Everaers*

The Journals liaised with the ISI's associations were offered to organize a Special Invited Paper Session (SIPS) at the Ottawa ISI World Statistics Conference. This is an excellent opportunity for the Journal and the Editor in Chief to promote a specific emerging theme and convey messages concerning the journal to a wider audience. This section in the Statistical Journal of the IAOS reports on this Special Session as held in July 2023 in Ottawa.

For the future of the Statistical Journal of the IAOS, developments such as the widening data availability and the increasing number of data stakeholders, the 'avalanche type' of new techniques and methodologies for example those related to Artificial Intelligence, and the IT-related developments in science publishing are an existential issue. The recent data and methodology developments are touched on in a variety of manuscripts published in the Journal as well as in an Editorial of the Statistical Journal of the IAOS, Volume 38/4.¹ IOS Press, the publisher of the Statistical Journal of the IAOS, takes a leading role in adapting to the trends in science publishing as was witnessed in the organization of the special seminar for the 35th anniversary of the company.²

The Special Session, organized by the Statistical Journal of the IAOS aimed to point to two important

developments in science publishing in official statistics. On the one hand, there are initiatives like the Open Data movement and FAIR publishing that aim to improve the transparency of the analysis, metadata, and data used by researchers as well as the accessibility of the results for users. On the other hand nowadays almost standard in online publishing practice, is the advanced use of hyperlinks to the data sets used, (dynamic) graphs and tables from other sources. The fast-growing variety of IT tools that allow comparing results from similar studies, creates opportunities for users to a more thorough assessment of the quality of the data and analysis as well as the added value of the specific analysis. Both developments are expected to ease the access and understanding of official statistics analysis and consequently impact statistical literacy. In the session at the conference, both developments were discussed, from the perspective of producers of official statistics as well as from the academic users' perspective.

The manuscripts in this section can of course only partially reflect the rich presentations and discussion, however, it is hoped that these manuscripts give a good flavor of the main trends in science publishing that will heavily impact publishing also in Official Statistics.

In the first article '*Ever More Transparent, Accessible, And Reproducible? The Impact Of Open Access, Open Data, And FAIR Publishing Principles On Data-Driven Research*' Gaby Umbach frames the Open Science, Open Data, Open Access, and Fair discussions and positions these developments within the requirements for openly accessible high-quality knowl-

*Corresponding author: International Association for Official Statistics, Salzburg, Austria. E-mail: pevssjiaos@gmail.com.

¹<https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji220107>.

²<https://www.iospress.com/news/celebrating-35-years-of-ios-press>.

edge as input into transparent and accountable decision-making and informed societal for evidence-informed policy-making (EIPM) and well-functioning societies in general. She argues that the wider concept of ‘Open Science’ supports this requirement. She shows in her manuscript how the ideas of Open Access, Open Data, and FAIR publishing principles, both as enablers and logical consequences of the wider paradigm of Open Science, revolutionize how academic research needs to be conceptualized, conducted, disseminated, published, and used. This ‘academic openness quartet’ is especially relevant for how research data are created, annotated, curated, managed, shared, reproduced, (re-)used, and further developed in academia. Greater accessibility of scientific output and scholarly data also aims at increasing the transparency and reproducibility of research results and the quality of research itself. In the applied ‘academic openness quartet’ perspective, they also function as remedies for academic malaises, like missing replicability of results or secrecy around research data. Against this backdrop, the present article offers a conceptual discussion on the four academic openness paradigms, their meanings, and their interrelations, as well as potential benefits and challenges arising from their application in data-driven research.

In the second manuscript, the focus is fully on the role of Open data and shows how especially during the last years the Open Data initiative, gained a strong momentum. In their manuscript, ‘*Data For Building Trust And Facilitating Use; The State Of Open Data In Official Statistics*’ Francesca Perucci and Eric Swanson explain how the COVID-19 pandemic and disasters related to climate change have demonstrated to us the critical importance of timely and open access to trusted data by underscoring the importance of open health and science data in managing crises. Open data principles and practices that facilitate data access and use, relevance to policy needs, and increase the impact and value of data are central to building trust in data. The authors argue that to improve the state of open data across the world more effort is needed. One underlying issue is the lack of political support leading to inadequate funding for NSOs to respond to the demand for more and better data systems. Another issue is the need to promote data use. However, the official statistics and data-for-development community still struggle with quantifying and measuring data use to truly understand the value of open data in official statistics.

The manuscript outlines four trends that present opportunities for expanding the adoption and use of open

data principles and practices and building data trust: the modernization of data governance; increased attention to the role of citizens in building trust and increasing the relevance of data and citizens’ contribution to data throughout the data value chain; the adoption of open data principles; and the work of watchdog organizations monitoring the progress of countries and agencies and identifying areas of data governance that still need attention.

The third manuscript illustrates how Eurostat in its position as a major provider of (European) official statistics, responds to the main challenges for open data dissemination taking into account its dissemination tradition. Christine Laaboudi, Martin Karlberg and Maja Islam in their manuscript ‘*Open data dissemination at Eurostat – state of the art*’, explain that the changes in the current information landscape make it increasingly challenging to provide high-quality statistics for Europe. The authors present the Eurostat dissemination approach (including the traditional dissemination vectors), and thereafter proceed to present the recent initiatives to make European statistics data and metadata available in the form of Linked Open Data (LOD). Following the European Commission and European Statistical System policies and strategic objectives and to improve the findability, accessibility, interoperability, and reuse of its data as per the FAIR Guiding Principles for scientific data management and stewardship (<https://www.go-fair.org/fair-principles/>), Eurostat has recently introduced several improvements to its data dissemination products. Presenting some of the main challenges for open data dissemination (complete reproducibility, availability of high-quality LOD, capacity to consume LOD and achieving meaningful mashups between official statistics LOD and other data sources), the manuscript concludes by noting the potential of LOD to foster transparency, reproducibility, collaboration, interdisciplinary research driving scientific advancements, and contributing to a broader understanding of complex scientific challenges.

The fourth contribution by Arofan Gregory ‘*Open And FAIR: Trends In Scientific Publishing And The Implications For Official Statistics*’, presents the emergence of the FAIR data principles as a major focus in the world of scientific research data. Recent trends in scientific publishing towards making data as open as possible and, more generally conformant to the FAIR Principles, impact not only the later stages of the scientific research cycle (i.e., communication and publication) but also earlier stages of project planning and data pro-

duction. While methodological solutions are emerging (and will likely become commonplace) commitments to open and FAIR data require fundamentally different approaches to data management within projects and long-term data stewardship after project funding ends. Wide-spread uptake of FAIR data and services will require the coordinated participation of numerous stakeholders including trusted and neutral organizations that can provide governance on both generic infrastructure and domain-relevant community standards.

The impact of the FAIR principles on official statistics, is only now starting to be visible. The author shows how the increased availability of research data, improvements in the area of machine-actionable metadata, and a focus on provenance information will lead to increased transparency and data. He argues that these are good reasons for the implementation of FAIR in official statistical organizations and that open and FAIR data will have a transformative impact on the role of Official Statistics in the data-intensive sciences

The Special Session at the ISI conference ended with a discussion by Pietro Gennari, *'Are Open Data And FAIR Publishing Principles Easy To Interpret And Apply?'* In his presentation Pietro Gennari, in his first appearance as the incoming editor-in-chief of the Statis-

tical Journal of the IAOS, argued how initiatives like the Open Data movement and FAIR publishing are changing the way Statistical Organizations and academic/Research Institutions disseminate and publish their data and research outputs. These initiatives improve the transparency and accessibility of the data as well as their reuse.

In his role as discussant (no manuscript), he reflected on his experiences in how these principles are applied by some selected producers of official statistics as well as by some academic institutions. The investigation assesses how closely existing data archives comply with the FAIR principles and how much effort is needed to adjust existing data repository structures to adhere to them. He concluded that the FAIR principles seem rather straightforward to interpret. In reality, their description presents several challenges as some facets are vague, while others appear to overlap; some are open-ended, while others require an external assessment. Even more challenging, he argued, is putting the FAIR guidelines into practice. For many data repositories, the degree of compliance about many facets is rather low. The Interoperable and Re-usable facets are, in particular, the most difficult to implement for both producers of official statistics and academic institutions.