



OPEN ACCESS

EDITED AND REVIEWED BY
Dietmar Wolfram,
University of Wisconsin–Milwaukee,
United States

*CORRESPONDENCE

Jean-Claude Burgelman
✉ jean-claude.burgelman@vub.be

SPECIALTY SECTION

This article was submitted to
Scholarly Communication,
a section of the journal
Frontiers in Research Metrics and Analytics

RECEIVED 02 March 2023

ACCEPTED 20 March 2023

PUBLISHED 24 April 2023

CITATION

Burgelman J-C and Wang K (2023) Editorial:
Data science and artificial intelligence for
(better) science.
Front. Res. Metr. Anal. 8:1177903.
doi: 10.3389/frma.2023.1177903

COPYRIGHT

© 2023 Burgelman and Wang. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License
\(CC BY\)](#). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted which
does not comply with these terms.

Editorial: Data science and artificial intelligence for (better) science

Jean-Claude Burgelman^{1*} and Kuansan Wang²

¹Department of Social Sciences and Solvay Business School, Open Science, Vrije Universiteit, Brussels, Belgium, ²Microsoft Research, Redmond, WA, United States

KEYWORDS

data science, artificial intelligence, open data, research life cycle, knowledge production

Editorial on the Research Topic

Data science and artificial intelligence for (better) science

The impact of data science and AI on science and knowledge production is an important and timely topic. The Frontiers Research Topic entitled “*Data science and artificial intelligence for (better) science*” has collated unique mixes of various contributions from experts, exploring a range of novel approaches to help solving problems facing scientists and advance scientific goals.

Research is also urgent to track the use of open data and to develop approaches to address the opacity of algorithms through open data. Scientific disciplines are called to make data in a way that is findable, accessible, interoperable, and reusable (FAIR), and crosses scientific boundaries.

Meaningful and explainable AI in research can only be fulfilled when as much data as possible is made FAIR (Findable, Accessible, Interoperable, and Reusable). How meaning is communicated in science “as precisely as possible” to machines when we formulate scientific concepts is a key question. Machine readability and interpretability is needed in order to make data and information “Fully AI-Ready” and support data-intensive research (Schultes et al.). The future of science is where there is only “one computer” and FAIR services see all FAIR data and effectively access a global FAIR database.

The single most important challenge is whether Data Science (and AI) can have a key role to improve the credibility and efficiency of research, one of the cornerstones on which science is built. When it comes to research software, caching (Schubotz et al.) can make experiments in research software reproducible. It is also a step forward toward making data related to research software FAIRer by extension.

Questions that science needs to raise with regard to Data Science, for instance, how to interact with data (which includes complex metadata), and how data science can facilitate the scientific cycle (exploration, analysis, interpretation, communication). Predictive models for Web-Enabled scientific discoveries are enabled by the surge of big (social) data. The social data are used to either discover or test scientific hypotheses. The process of scientific discovery is a cyclically sequence of exploration, prediction and validation. Data-driven

computational social network science (DD-CSNS) (Emmert-Streib and Dehmer), combining big social data with social networks will enable scientific discoveries.

Finally, the question is how to enable (better) open science. Increasingly relevant today than ever before is the greater reliance on access to data, artificial intelligence (AI) and machine learning (ML). Data access increasingly determines scientific discoveries and advancements. Data reuse is at the forefront of an emerging “third wave of open data” (Verhulst et al., 2020). But despite progress in implementing open data and FAIR principles, science data asymmetries (as in disparities in access to science data) are a growing problem and can undermine scientific progress. Comparative research is needed to document (Verhulst and Young) for instance, investigating the creation of new types of data asymmetries by, e.g., new private-sector investments in data platforms and knowledge repositories, how data portability and interoperability impact the practice of data collaboration, the relationship and interplay between existing asymmetries and technological and societal drivers. Finally, new methods for achieving a social license for data use and reuse toward the public good are needed, capturing multiple stakeholders’ acceptance of standard practices and procedures.

References

Verhulst, S., Young, A., Zahuranec, A., Aaronson, S. A., Calderon, A., and Gee, M. (2020). *The Emergence of a Third Wave of Open Data: How To Accelerate the Re-Use of Data for Public Interest Purposes While Ensuring Data Rights*

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

KW was employed by Microsoft Research.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

and Community Flourishing. Open Data Policy Lab. Available online at: <https://opendatapolicylab.org/images/odpl/third-wave-of-opendata.pdf> (accessed June 30, 2022).