Artificial Intelligence in the Life Sciences 1 (2021) 100002

Contents lists available at ScienceDirect

Artificial Intelligence in the Life Sciences

journal homepage: www.elsevier.com/locate/ailsci

Editorial

Reproducibility, reusability, and community efforts in artificial intelligence research

Jürgen Bajorath[a,*], Connor W. Coley[b], Melissa R. Landon[c], W. Patrick Walters[d], Mingyue Zheng[e]

[a] Department of Life Science Informatics and Data Science, B-IT, Rheinische Friedrich-Wilhelms-Universität, Friedrich-Hirzebruch-Allee 6, D-53115 Bonn, Germany
[b] Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, United States
[c] Cyclica, 207 Queens Quay West, Suite 420, Toronto, Ontario, M5J 1A7, Canada
[d] Relay Therapeutics, 399 Binney Street, 2nd Floor, Cambridge, MA 02139, United States
[e] State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Rd., Zuchongzhi, Shanghai, China

Artificial Intelligence in the Life Sciences (*AILSCI*) supports open science, data sharing, and good practices to ensure reproducibility of scientific investigations. Thus, for publication of original research in *AILSCI*, primary data that are not in the public domain must be made available as well as custom code that is essential for the main results of a study. However, proprietary data constraints in the industry are also appreciated and should not preclude publication of interesting science from commercial environments. Therefore, *AILSCI* offers a variety of article categories.

Currently, we consider the following types of manuscripts (for further details, see the Guide for Authors):

1 *Research Articles* communicate original research yielding novel findings. Reports of novel computational methods with life science impact as well as studies combining AI and experimental work are highly encouraged.
2 *Communications* report preliminary research that is novel and scientifically sound and may give rise to a larger-scale investigation.
3 *Reviews* comprehensively cover an area of AI with high relevance for the life sciences or AI applications in a particular life science discipline or interdisciplinary research.
4 *Perspectives* have the character of mini reviews of a given topic or area of research with an explicit personal outlook.
5 *Conceptual Analysis* articles cover AI concepts for the life sciences or re-evaluate existing theories from a life science viewpoint.
6 *Methods & Protocols* concisely introduce a method, workflow, software, or study protocol that is readily applicable and fully reproducible, including open source software. Presenting exemplary applications is encouraged.
7 *Controversial Views* discuss research results or scientific issues that may be controversially rated such as current methodological limitations, frequently overlooked problems, or provocative concepts charting new territory, even if preliminary. Such contributions might also be solicited by the editors to present opposing viewpoints of different investigators side-by-side.
8 *Opinions* are commentaries or reflections on a particular topic, finding, or trend in the field, with emphasis on personal views and/or recommendations.

Articles such as *Perspectives, Conceptual Analysis, Methods & Protocols, Controversial Views*, or *Opinions* provide a variety of formats for communicating new scientific approaches, case studies, and personal views or experiences, without the need for data disclosure. Hence, these types of manuscripts can be readily considered for presenting new scientific developments in the industry in different ways.

Naturally, there might occasionally be borderline situations requiring decisions on a case-by-case basis. For example, in a *Methods* paper, a new algorithm developed in the industry might be presented as pseudo-code, without providing the software, if the description is sufficiently detailed to enable re-implementation by others. However, if a software vendor would aim to benchmark proprietary software against publicly available programs and claim superior performance, the study would not be publishable in *AILSCI*, regardless of the manuscript type, unless the software is made available to the public.

In the following, we specify further opportunities for different types of contributions that fall within the scope of *AILSCI*.

Reusability and adaptability of algorithms and code

The use or extension of published methods and source code for different applications is important for the further development of the field and also supports our quest for reproducibility.

Inspired by the recent introduction of *Reusability Reports* in *Nature Machine Intelligence* [1,2], publication of such studies as *Methods & Protocols* contributions in *AILSCI* is strongly encouraged. It is essential that calculations and workflows reported in these papers are fully reproducible. Hence, they should be provided as open source scripts, Jupyter notebooks, or in another suitable format.

Furthermore, we note that computer science and the life sciences have different publication cultures. In computer science, new algorithms and methods are preferentially published in conference proceedings that

* Correspondence author.
E-mail address: bajorath@bit.uni-bonn.de (J. Bajorath).

https://doi.org/10.1016/j.ailsci.2021.100002
Received 15 March 2021; Accepted 16 March 2021
Available online 24 March 2021

are typically not considered in the life sciences. If AI approaches originally reported in conference proceedings have potential for life science applications *AILSCI* invites re-publication of such contributions in modified form as *Conceptual Analysis* or *Methods & Protocols* papers. In such cases, the life science relevance of the reported algorithm or method should be briefly discussed in the manuscript and additional explanations and/or a summary should be provided that render the work accessible to an interdisciplinary life science-oriented audience.

We also note that *Methods & Protocols* manuscripts cover studies published as *Application Notes, Software*, or equivalent reports published by other journals.

Educational and community efforts

AILSCI also offers the opportunity to publish thematic *Article Collections* resulting from community initiatives such as method evaluations or blind test predictions. *Article Collections* may involve all of *AILSCI's* manuscript categories and are introduced by invited *Editorials*. This opportunity is also available to scientific organizations or societies for publishing papers originating from conferences or organization-specific initiatives.

For an evolving scientific community engaging in interdisciplinary AI research, paying attention to educational efforts, at different levels, is also highly relevant. For example, a new teaching concept integrating AI and life science components may be presented as a *Conceptual Analysis.* Moreover, educational viewpoints might also aim to provide guidance for addressing critical issues in a fast-moving scientific field such as, for example, the gap between hype and reality in AI applications, for which *Controversial Views* provide an excellent discussion forum.

Open access and preprint servers

In the spirit of promoting open science, *AILSCI* articles are published either under the *Creative Commons Attribution (CC BY)* license or under the *Creative Commons Attribution-NonCommercial-NoDerivs (CC-BY-NC-ND)* license. This will ensure that all articles are immediately and permanently free for everyone to read, download, copy, and distribute. For further details, see the Open Access information on the Journal website.

We additionally recognize the value of preprint servers such as arXiv, bioRxiv, and ChemRxiv for rapid dissemination of research results in parallel with peer review. For all manuscript categories, authors can share their preprints anywhere at any time. For further details, see sharing policies of the Journal.

We look forward to receiving many interesting and thematically diverse contributions that will help to further support the development of AI-driven interdisciplinary science in various ways!

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Research, reuse, repeat (editorial). Nat Mach Intell 2020;2:729.
[2] Mohapatra S, Yang T, Gómez-Bombarelli R. Reusability report: designing organic photoelectronic molecules with descriptor conditional recurrent neural networks. Nat Mach Intell 2020;2:749–52.