BRAINIACS
BRAIN IMAGING AND COMPUTING SCIENCES

2024 Volume 5
Issue 2 Edoc QA6A795A3

From Open Review to Reproducible Review: FAIR Metrics Analysis of Open Peer Reviews for Brain Informatics Literature*

Adam Craig and Carl Taswell†

Abstract

Because the goal of brain informatics is to help researchers discover and derive new insights from existing data and metadata, documentation of new methods, platforms, and data sources in scholarly reports is especially important in this field. Expert evaluation of new scholarly works through peer review is essential to maintaining the integrity of these reports, but how best to evaluate the quality of the peer reviews remains an open question. We previously proposed the paradigm of reproducible peer review, in which a second reviewer should be able to draw on the same factual claims as the first in order to reach the same conclusion. We then introduced the Fair Attribution to Indexed Reports (FAIR) Metrics for peer reviews as metrics of how well reviewers attributed the claims substantiating their recommendations to sources. However, we previously only demonstrated these new FAIR Metrics on five example reviews. We here report the results of FAIR Metrics analyses of published peer reviews of 14 brain informatics articles. These analyses demonstrate the value of the FAIR Metrics by highlighting key ways the brain informatics community can improve the reproducibility of the peer review process, particularly by calling for references to or quotes from the specific passages of the work under review, indication of which standards of the publication venue the work meets or fails to meet, and citation of the literature when drawing on prior knowledge of the problem domain.

Keyphrases

brain informatics, peer review, reproducibility, FAIR Metrics, PDP-DREAM Ontology, NPDS Cyberinfrastructure, PORTAL-DOORS Project

Contents

Introduction

Peer review in brain informatics

As early as 2007, multiple works, including Zhong et al. 2007 and Taswell 2007, were highlighting both the potential of brain informatics as an emerging field that leverages artificial intelligence to aid humans in solving the problems related to brain health and also the dependence of that emergence on the availability of well-curated data and knowledge resources on the World-Wide Web. Since the 19th century, peer-reviewed journals have played an increasingly central role as sources of such resources Burnham 1990. However, few studies have systematically assessed the effectiveness of peer review at maintaining the quality of such resources Jefferson et al. 2002, and the lack of clear standards for accountability of editors and reviewers and justification for recommendations or decisions has become an impediment to fully realizing the social good that peer review can achieve Tennant and Ross-Hellauer 2020. Open peer review is one proposed solution to the limitations of the current peer review process: Since journals cannot guarantee that the quality peer review will be satisfactory for a given purpose, they should publish the reviews alongside the articles so that readers can decide for themselves Wolfram et al. 2020. While this move toward increased openness can create more opportunities for public discussion and debate of the merits of works and the quality of the review process, it alone is not sufficient to address the lack of systematic standards. We previously proposed a further step toward establishing clear and systematic community standards that we call reproducible peer review: The reviewer should support their recommendation with sufficient factual claims, each clearly attributed to a source, that a second reviewer can follow their reasoning and arrive at the same recommendation Craig, Lee, et al. 2022.

*Presented 2024-10-09 at Guardians 2024 with slides and video.
†Correspondence to ctaswell@bhavi.us.

FAIR Metrics for peer review

For reproducible peer review to gain traction as a standard of excellence, not just a catchy phrase, reproducibility needs to be measurable. To that end, we recently introduced FAIR Metrics for peer reviews in a poster presentation Craig and Taswell 2024a and a full paper Craig and Taswell in press. These metrics draw on the principles that we had previously used to guide the design of FAIR Metrics of adherence to good citation practices for scholarly articles Craig, Athreya, et al. 2023. Most importantly, the metrics should measure how well the authors support reproducibility by clearly attributing claims to their sources, and claims are equivalent when they convey the same meaning, regardless of wording Craig, Athreya, et al. 2023. Furthermore, it is not enough for the reviews themselves to be reproducible Craig, Athreya, et al. 2023. The evaluator must record the FAIR Metrics analysis in a systematic way Craig, Athreya, et al. 2023. To support this, we created a module of the PDP-DREAM Ontology with classes and properties useful for creating machine-readable semantic records of the analyses Craig, Athreya, et al. 2023. We subsequently extended this module to aid creation of records of FAIR Metrics analyses of peer reviews Craig and Taswell 2024a. The PDP-DREAM Ontology is a formal ontology that codifies the DREAM Principles, the design principles that guide the PORTAL-DOORS Project, but it also supports the inclusion of smaller modules for such specific purposes Craig and Taswell 2021.

The key distinction between FAIR Metrics for scholarly publications and for peer reviews is that we do not expect peer reviews to introduce novel ideas, and thus we only focus on how well a review supports its recommendations with factual claims properly attributed to their sources Craig and Taswell 2024a. Additionally, we separate claims into types based on the subject of the claim and thus the kind of attribution needed: claims about the work under review, about the publication venue, or about the problem domain to which the reviewed work or venue relates Craig and Taswell 2024a.

In the previous work, we demonstrated use of FAIR Metrics with five example reviews Craig and Taswell 2024a: one simple example review of a fictional paper, two reviews of a rejected submission to the ACM Multimedia call for grand challenges, a revised version of which is available as Craig and Taswell 2024b from *Brainiacs Journal*, and two published peer reviews of a recently published neuroscience article G. Lu et al. 2024. In the current work, we apply these same FAIR Metrics to evaluate examples of published peer reviews of works related to brain informatics. This analysis allows us to identify key areas where editors and readers can set higher standards to support reproducibility of peer review and, thereby, a more well-curated scientific record.

Methods

Literature search

We searched the websites of six publishers advertised as practicing open peer review. On each site, we searched with the two queries "brain informatics", "brain imaging data management" without quotes and selected the first four articles that appeared to be about brain informatics based on their abstracts and that had at least two published peer reviews. We considered only reviews of initial submissions, as reviews of revised versions rarely had new critiques and instead merely acknowledged that the authors had made the recommended changes. The reviews we evaluated are of the following 14 articles, grouped by publisher and journal:

- 2 from eLife Sciences Publications, both in *eLife*: Scheffer et al. 2020, Markiewicz et al. 2021
- 4 from F1000Research, all in *F1000Research*: Attendees 2016, Crusio et al. 2017, Navale et al. 2020, Guiet et al. 2021
- 2 from Open Research Europe (also part of the F1000 publishing group), both in *Open Research Europe*: Tarnanas et al. 2021, Ilias et al. 2023
- 1 from IOS Press, in *Semantic Web Journal* Sy et al. 2023
- 3 from Nature Research: 1 in *Nature* Oh et al. 2014, 1 in *Nature Communications* Collins et al. 2024, and 1 in *Nature Human Behavior* Li et al. 2024
- 2 from Oxford University Press, both in *GigaScience*: Craddock et al. 2015, O'Connor et al. 2017

We assessed two reviews per article for a total of 28 reviews.

FAIR Metrics calculations

For each peer review, we calculated FAIR Metrics according to the process described in Craig and Taswell 2024a. We can summarize this process in five steps: 1. Read the review, and identify the key factual claims that the reviewer is using to support their recommendation. 2. Classify each claim as pertaining primarily to either the work under review (target work), the publication or conference to which the authors submitted it for publication (venue), or additional information relevant to the problem domain of the venue or publication (domain knowledge). 3. Classify each claim as correctly attributed to a source or misattributed. A correctly attributed claim has a cited source and accurately reflects the meaning of one or more statements in that source. A misattributed claim either has no cited source or misrepresents the content of that source. 4. Tabulate six counts of classified claims:

A_T : correctly attributed statements about the target work

M_T : misattributed statements about the target work

A_V : correctly attributed statements about the venue

M_V : misattributed statements about the venue

A_D : correctly attributed statements about domain knowledge

M_D : misattributed statements about domain knowledge

5. Use these counts to calculate four ratio FAIR Metrics of peer review quality:

$f_T = \frac{A_T - M_T}{A_T + M_T}$ target ratio

$f_V = \frac{A_V - M_V}{A_V + M_V}$ venue ratio

$f_D = \frac{A_D - M_D}{A_D + M_D}$ domain ratio

$f_J = \frac{A_T + A_V + A_D - M_T - M_V - M_D}{A_T + A_V + A_D + M_T + M_V + M_D}$ combined justification ratio

Semantic records of FAIR Metrics analyses

As described in Craig and Taswell 2024a, we have extended the PDP-DREAM Ontology FAIR Metrics module with classes and properties that we use to make Resource Description Framework (RDF) documents recording FAIR Metrics analyses of peer reviews, which we have listed in 1, 2, and 3. A well-documented FAIR Metrics analysis of a peer review lists a unique URI to identify the work under review, all key claims of the review, each assigned the correct class corresponding to one of the six categories, the cited source of each claim if any, the six counts, and the four FAIR Metrics values.

Results

While explicit misrepresentations of works as making claims that they did not were rare (2 occurrences), misrepresentations of omission, in which reviewers falsely claimed that a work was missing information when it was present were common (10 occurrences). Most claims used to justify a recommendation were about the work itself with very few explicitly invoking requirements of the publication venue (4 occurrences) or knowledge from prior work (3 occurrences). In 2 instances, reviewers implicitly treated common practices, such as the division of a work into Introduction, Methods, Results, Discussion, and Conclusion sections, as requirements, even when the editorial policies of the journal or publisher made no such requirement.

Discussion

Holding peer review to a higher standard

While the infrequent references to venue requirements may reflect an implicit mutual understanding that a submission meets basic requirements of relevance and proper presentation unless otherwise noted, the near-absence of discussion of how the works under review fit into the larger context of prior research in brain informatics indicates that the reviewers have not adequately assessed the novelty or importance of the work. At a minimum, a reviewer should agree or disagree with the authors' assertion that the work fills some gap in knowledge or solves some unsolved problem. If they agree, then they can refer to the same sources the authors used to justify the claim. If they disagree, then they can reference other works that present results answering the same question or providing an existing solution.

The high incidence of misattributions of omission suggests a different problem, but the origin of this tendency is unclear. It may simply be due to a lack of attention to detail from reviewers when reading works, but that does not explain why they then assume that the information they insist is important to include is absent. One possible explanation is miscommunication between reviewers and authors due to differences in their understanding of the terminology of the field. Another is that reviewers simply use accusations of omission as stock criticisms with which to fill out the length of the review. To identify the root of the problem, we will need to do further research that will involve actively engaging with peer reviewers, presenting them with the passages that provide the information they demanded and recording their responses regarding whether they suffice and, if not, how they should provide more detail.

Easing adoption of FAIR Metrics

FAIR Metrics evaluation, whether of scholarly works submitted for publication or of peer reviews, requires systematic assessment of not only the text itself but of multiple related works. Tools for automating parts of the FAIR Metrics evaluation process will make their use more practical. To this end, Brain Health Alliance will be opening the first Multimedia FAIR Metrics Grand Challenge to submissions in 2025. This contest will award a cash prize to the team that develops the best software automating some or all of the steps of FAIR Metrics evaluation Craig and Taswell 2024b:

- Extract text and image data from different file formats.
- Separate text into discrete statements.
- Convert information in figures and tables into discrete statements.
- Distinguish substantive claims from other statements.
- Retrieve cited sources of claims.
- Search prior work for potential uncited sources of claims.
- Distinguish whether two claims are equivalent in meaning.

Conclusion

We performed FAIR Metrics analyses of published peer reviews of scholarly articles on brain informatics. Based on these analyses, we recommend that the brain informatics community hold peer review to a higher standard of reproducibility. Specifically, we recommend that reviewers make clear the sources on which they are basing their claims.

Table 1: Classes of the FAIR module of the PDP-DREAM Ontology relevant to the assessment of peer reviews
("New" means introduced specifically for the assessment of peer review.)

Classes			
Name	New	Parent	Explanation
PdpDreamFairEntity	No	owl:Thing	equivalent to owl:Thing, root class for the module
Document	No	PdpDreamFairEntity	a resource containing text and possibly other media
Review	Yes	Document	a document that is a review of another resource
Statement	No	PdpDreamFairEntity	a statement in some language
AttributedStatement	No	Statement	a statement correctly attributed to another resource, called "Quoted" in earlier versions
MisattributedStatement	No	Statement	a statement incorrectly attributed to another resource, called "Misquoted" in earlier versions
AttributedTargetStatement	Yes	AttributedStatement	a statement correctly attributed to the work under review
MisattributedTargetStatement	Yes	MisattributedStatement	a statement incorrectly attributed to the work under review
AttributedVenueStatement	Yes	AttributedStatement	an statement correctly attributed to an editorial policies document of the publication venue
MisattributedVenueStatement	Yes	MisattributedStatement	a statement incorrectly attributed to an editorial policies document of the a publication venue
AttributedDomainStatement	Yes	AttributedStatement	a statement correctly attributed to other prior work
MisattributedDomainStatement	Yes	MisattributedStatement	a statement incorrectly attributed to other prior work

Table 2: Object properties of the FAIR module of the PDP-DREAM Ontology relevant to the assessment of peer reviews
("New" means introduced specifically for the assessment of peer review.)

Object Properties			
Name	New	Parent	Explanation
hasPdpDreamFairObjectProperty	No	owl:ObjectProperty	root object property for the module
isReviewOf	Yes	hasPdpDreamFairObjectProperty	The subject is a review of the object.
hasStatement	No	hasPdpDreamFairObjectProperty	The subject resource includes the object statement.
hasAttribution	No	hasPdpDreamFairObjectProperty	The subject statement has an attribution (whether correct or not) to the object resource.
hasEquivalentStatement	No	hasPdpDreamFairObjectProperty	The subject and object statements are equivalent in meaning.
hasContradictingtStatement	Yes	hasPdpDreamFairObjectProperty	The subject and object statements contradicts each other.

Table 3: Data properties of the FAIR module of the PDP-DREAM Ontology relevant to the assessment of peer reviews ("New" means introduced specifically for the assessment of peer review.)

Data Properties			
Name	New	Parent	Explanation
hasPdpDreamFairDataProperty	No	owl:DatatypeProperty	root data property for the module
hasName	No	hasPdpDreamFairDataProperty	The object is a name for the subject.
hasText	No	hasPdpDreamFairDataProperty	The object is a textual representation of the subject.
hasFairMetricValue	No	hasPdpDreamFairDataProperty	root for properties that assign FAIR Metric values
hasFairMetricCount	No	hasFairMetricValue	root for data properties that assign FAIR Metric counts
hasFairATCount	Yes	hasFairMetricCount	The object is the A_T count value of the subject.
hasFairMTCount	Yes	hasFairMetricCount	The object is the M_T count value of the subject.
hasFairAVCount	Yes	hasFairMetricCount	The object is the A_V count value of the subject.
hasFairMVCount	Yes	hasFairMetricCount	The object is the M_V count value of the subject.
hasFairADCount	Yes	hasFairMetricCount	The object is the A_D count value of the subject.
hasFairMDCount	Yes	hasFairMetricCount	The object is the M_D count value of the subject.
hasFairMetricRatio	No	hasFairMetricValue	root for data properties that assign FAIR Metric ratio scores
hasFairFTRatio	Yes	hasFairMetricRatio	The object is the f_T score value of the subject.
hasFairFVRatio	Yes	hasFairMetricRatio	The object is the f_V score value of the subject.
hasFairFDRatio	Yes	hasFairMetricRatio	The object is the f_D score value of the subject.
hasFairFJRatio	Yes	hasFairMetricRatio	The object is the f_J score value of the subject.
raisesQuestion	Yes	hasPdpDreamFairDataProperty	The meta-reviewer is asking the reviewer this question about the subject.

Citation

Brainiacs 2024 Volume 5 Issue 2 Edoc QA6A795A3
Title: “From Open Review to Reproducible Review: FAIR Metrics Analysis of Open Peer Reviews for Brain Informatics Literature”
Authors: Adam Craig and Carl Taswell
Dates: created 2024-10-07, presented 2024-10-09, updated 2024-12-23, published 2024-12-23,

Contact: ctaswell@bhavi.us
URL: BrainiacsJournal.org/arc/pub/Craig2024FAIRMEPR
PDP: /Nexus/Brainiacs/Craig2024FAIRMEPR
DOI: /10.48085/QA6A795A3

References

[1] G. B. W. 2. Attendees. “Grand challenges for global brain sciences.” *F1000Research* 5 (2016), p. 2873 (cited p. 2).

[2] J. C. Burnham. “The evolution of editorial peer review.” *Jama* 263.10 (1990), pp. 1323–1329 (cited p. 1).

[3] E. Collins, O. Chishti, S. Obaid, H. McGrath, et al. “Mapping the structure-function relationship along macroscale gradients in the human brain.” *Nature Communications* 15.1 (2024), p. 7063 (cited p. 2).

[4] R. C. Craddock, R. L. Tungaraza, and M. P. Milham. “Connectomics and new approaches for analyzing human brain functional connectivity.” *Gigascience* 4.1 (2015), s13742–015 (cited p. 2).

[5] A. Craig, A. Athreya, and C. Taswell. “Managing Lexical-Semantic Hybrid Records of FAIR Metrics Analyses with the NPDS Cyberinfrastructure.” *Brainiacs Journal of Brain Imaging And Computing Sciences* 4.2 (Dec. 27, 2023). DOI: 10.48085/D5B2734F2 (cited p. 2).

[6] A. Craig, C. Lee, N. Bala, and C. Taswell. “Motivating and Maintaining Ethics, Equity, Effectiveness, Efficiency, and Expertise in Peer Review.” *Brainiacs Journal of Brain Imaging And Computing Sciences* 3.1, I5B147D9D (1 June 30, 2022), pp. 1–21. DOI: 10.48085/I5B147D9D. URL: https://BrainiacsJournal.org/arc/pub/Craig2022MMEEPR (cited p. 1).

[7] A. Craig and C. Taswell. “PDP-DREAM Software for Integrating Multimedia Data with Interoperable Repositories.” *Brainiacs Journal of Brain Imaging And Computing Sciences* 2.1, HA46280EF (1 Dec. 31, 2021), pp. 1–6. DOI: 10.48085/HA46280EF. URL: https://BrainiacsJournal.org/arc/pub/Craig2021SIMDIR (cited p. 2).

[8] A. Craig and C. Taswell. “FAIR Metrics for Motivating Excellence in Peer Review.” In: *2024 IEEE 20th International Conference on e-Science (e-Science)*. IEEE. 2024, pp. 1–2 (cited pp. 2, 3).

[9] A. Craig and C. Taswell. “The Multimedia FAIR Metrics Grand Challenge.” *Brainiacs Journal of Brain Imaging And Computing Sciences* 5.1 (June 30, 2024). DOI: 10.48085/G7ECAEAD9 (cited pp. 2, 3).

[10] A. Craig and C. Taswell. “FAIR Metrics for Motivating Ethics in Peer Review.” In: *The 16th Workshop on Natural Language Processing and Ontology Engineering (NLPOE2024), in conjunction with The 23rd IEEE/WIC International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT2024)*. IEEE. in press (cited p. 2).

[11] W. E. Crusio, C. Rubino, and A. Delprato. “Engaging high school students in neuroscience research-through an e-internship program.” *F1000Research* 6 (2017) (cited p. 2).

[12] R. Guiet, O. Burri, N. Chiaruttini, O. Hagens, and A. Seitz. “DEVILS: a tool for the visualization of large datasets with a high dynamic range.” *F1000Research* 9.1380 (2021), p. 1380 (cited p. 2).

[13] L. Ilias, G. Doukas, M. Kontoulis, K. Alexakis, A. Michalitsi-Psarrou, C. Ntanos, and D. Askounis. “Overview of methods and available tools used in complex brain disorders.” *Open Research Europe* 3.152 (2023), p. 152 (cited p. 2).

[14] T. Jefferson, P. Alderson, E. Wager, and F. Davidoff. “Effects of editorial peer review: a systematic review.” *Jama* 287.21 (2002), pp. 2784–2786 (cited p. 1).

[15] X. Li, N. Bianchini Esper, L. Ai, S. Giavasis, et al. “Moving beyond processing-and analysis-related variation in resting-state functional brain imaging.” *Nature Human Behaviour* 8.10 (2024), pp. 2003–2017 (cited p. 2).

[16] G. Lu, C. Gong, Y. Sun, X. Qian, et al. “Noninvasive imaging-guided ultrasonic neurostimulation with arbitrary 2D patterns and its application for high-quality vision restoration.” *Nature Communications* 15.4481 (May 2024). DOI: https://doi.org/10.1038/s41467-024-48683-6 (cited p. 2).

[17] C. J. Markiewicz, K. J. Gorgolewski, F. Feingold, R. Blair, et al. “The OpenNeuro resource for sharing of neuroscience data.” *Elife* 10 (2021), e71774 (cited p. 2).

[18] V. Navale, M. Ji, O. Vovk, L. Misquitta, T. Gebremichael, A. Garcia, Y. Fann, and M. McAuliffe. “Development of an informatics system for accelerating biomedical research.” *F1000Research* 8 (2020), p. 1430 (cited p. 2).

[19] D. O’Connor, N. V. Potler, M. Kovacs, T. Xu, et al. “The Healthy Brain Network Serial Scanning Initiative: a resource for evaluating inter-individual differences and their reliabilities across scan conditions and sessions.” *Gigascience* 6.2 (2017), giw011 (cited p. 2).

[20] S. W. Oh, J. A. Harris, L. Ng, B. Winslow, et al. “A mesoscale connectome of the mouse brain.” *Nature* 508.7495 (2014), pp. 207–214 (cited p. 2).

[21] L. K. Scheffer, C. S. Xu, M. Januszewski, Z. Lu, et al. “A connectome and analysis of the adult Drosophila central brain.” *elife* 9 (2020), e57443 (cited p. 2).

[22] M. F. Sy, B. Roman, S. Kerrien, D. M. Mendez, et al. “Blue Brain Nexus: An open, secure, scalable system for knowledge graph management and data-driven science.” *Semantic Web* 14.4 (2023), pp. 697–727 (cited p. 2).

[23] I. Tarnanas, P. Vlamos, R.-A. Consortium, et al. “Can detection and prediction models for Alzheimer’s Disease be applied to Prodromal Parkinson’s Disease using explainable artificial intelligence? A brief report on Digital Neuro Signatures.” *Open Research Europe* 1.146 (2021), p. 146 (cited p. 2).

[24] C. Taswell. “DOORS to the Semantic Web and Grid with a PORTAL for Biomedical Computing.” *IEEE Transactions on Information Technology in Biomedicine* 12.2 (2 Mar. 2007). In the Special Section on Bio-Grid published online 3 Aug. 2007, pp. 191–204. ISSN: 1089-7771. DOI: 10.1109/TITB.2007.905861 (cited p. 1).

[25] J. P. Tennant and T. Ross-Hellauer. “The limitations to our understanding of peer review.” *Research integrity and peer review* 5.1 (2020), p. 6 (cited p. 1).

[26] D. Wolfram, P. Wang, A. Hembree, and H. Park. “Open peer review: promoting transparency in open science.” *Scientometrics* 125.2 (2020), pp. 1033–1051 (cited p. 1).

[27] N. Zhong, J. Liu, Y. Yao, J. Wu, S. Lu, Y. Qin, K. Li, and B. Wah. “Web intelligence meets brain informatics.” In: *Web Intelligence Meets Brain Informatics: First WICI International Workshop, WImBI 2006, Beijing, China, December 15-16, 2006, Revised Selected and Invited Papers 1*. Springer. 2007, pp. 1–31 (cited p. 1).