BRAINIACS
BRAIN IMAGING AND COMPUTING SCIENCES

2024 Volume 5
Issue 1 Edoc G7ECAEAD9

The Multimedia FAIR Metrics Grand Challenge*

Adam Craig and Carl Taswell†

Abstract

Brain Health Alliance (BHA), a US 501c3 nonprofit organization, will host the first annual Multimedia FAIR Metrics Grand Challenge on 9th September 2024. The FAIR Metrics, with acronym FAIR for the phrases *Fair Attribution to Indexed Reports* and *Fair Acknowledgment of Information Records*, quantify how well a scholarly work cites and discusses prior literature and the extent to which it remains devoid of plagiarism or misrepresentation of previously published work. Unlike lexical plagiarism detection, FAIR Metrics semantic analyses require identifying statements with equivalent meaning. Recording the comparison of documents in searchable records of FAIR Metrics analyses strengthens the integrity of scholarly publishing by providing a more transparent and systematic way to trace the origins of concepts, ideas, and creative contributions to the historical record of published literature.

FAIR Metrics analysis by a human expert remains subject to debate because semantic similarity of concepts and ideas can be more difficult to quantify than lexical edit distance. This grand challenge will solicit automated tools to perform one or all stages of FAIR Metrics analysis with a focus on their use for plagiarism detection. Because FAIR Metrics analysis currently depends on human judgment with various opinions about which statements in a document are substantive claims and which are equivalent in meaning, we cannot currently declare a set of FAIR Metric values to be *uniquely correct* for a given document. Instead, we must establish *expert consensus* and will evaluate automated tools based on uniform formatting of the plagiarism analysis records and the ability to differentiate plagiarizing from non-plagiarizing documents. For the competition, there will be a total of four separate sets of published reports used in the data repository of documents for development versus evaluation purposes in the competition on plagiarizing and retracted versus non-plagiarizing and non-retracted.

In this year's grand challenge, we will focus on the core media types found in most scholarly articles: prose text, figures (images), and tables. However, a growing number of digitally published works include supplementary multimedia content, such as video, audio, source code, and numerical data in a wide variety of formats and data repositories. Future iterations of this grand challenge will focus on extracting claims from different types of media and placing them in a shared semantic format that allows contrast and comparison.

Keyphrases

Research integrity, citational justice, publishing ethics, plagiarism detection, FAIR Metrics, multimedia data, artificial intelligence.

Contents

Importance

The integrity of the historical record of scholarly published literature remains essential to the progress of science in every field because it enables both inquiry into whether information comes from reliable sources and the recognition of those scholars who contribute truthfully (LaFollette 1992). However, the current climate with overemphasis on citation metrics as measures of research productivity causes consequential adverse impacts with perverse incentives for researchers to avoid citing potential rivals (S. K. Taswell et al. 2020). With generative artificial intelligence making it so easy for non-experts to paraphrase and propagate plagiarized content and even fabricate entire research articles (Elali and Rachid 2023), the requirement for better tools to help fight against misinformation and disinformation in these information wars will only continue to increase in relevance and importance. In a world where authors now use diverse kinds of media to communicate the results of their work, the multimedia research community represents the best pool of multi-disciplinary know-how to help trace the flow of ideas from one medium to the next.

*Presented 2024-06-22 at EPSMS 2024 with slides and video.
†Correspondence to ctaswell@bhavi.us.

Research Tasks

In our previous work demonstrating FAIR Metrics analysis of published scholarly articles, both retracted and non-retracted, we defined a workflow for focused analysis examining a target article for presence of ideas and information plagiarized from a specific comparison article (Craig, Athreya, et al. 2023b). The goal of this grand challenge will be to automate this workflow. Here we describe the steps in more detail:

1. Identify and obtain the text, figures, tables, and other media of the target work and comparison work. For this grand challenge, the media for a single work will be in a single portable document format (PDF) file. Obtaining the text of a work will consist of loading a file stored on the file system of the local machine. The software may open a file selection dialogue, prompt the user to type in file paths, or accept the file names as command line arguments. We recommend the last of these options, as it allows use of automated evaluation scripts.
2. Identify the works to which the target work attributes claims (reference works), and obtain their text, figures, tables, and other media.
3. Identify the media types present in each work. In this iteration, we will focus on conversion of figures and tables present in the paper into text, but future iterations will address working with other media, such as audio and video recordings.
4. Convert figures, tables, and other media into text descriptions.
5. Split the text, including portions derived from non-text media, into individual statements.
6. Search for equivalent statements within a text to eliminate repeated statements.
7. Distinguish the claims of a work from other statements. Claims are statements that substantively contribute to the argument that the authors are presenting. For the purposes of this grand challenge, a claim should have one of two features: either an attribution to a reference work or associated language in the text indicating that it is novel to the target work.
8. For each claim with an attribution, search the reference work for an equivalent statement.
9. If one is present, classify the claim as Attributed. Otherwise, classify the claim as Misattributed.
10. For each claim with linguistic cues that the authors are representing it as novel, search the the comparison work for an equivalent statement.
11. If one is present, classify the claim as Plagiarized. Otherwise, classify the claim as Novel.
12. Count the number of claims in each category, storing the values in count variables A for Attributed, M for Misattributed, P for Plagiarized, and N for Novel.
13. Compute the FAIR metrics from the counts according to the following formulae:
    - $F_A = A/(A + M + P)$
    - $F_M = (A - M)/(A + M + P)$
    - $F_P = (A - P)/(A + M + P)$
    - $F_N = (A - N)/(A + M + P + N)$
14. Render a decision as to whether the target work has plagiarized from the comparison work based on whether F_P falls below some threshold set by the entrant either in the source code itself, in a launch script, or in a configuration file.
15. Record the results of the analysis, including the following:
    - the individual claims considered
    - their attributions as applicable
    - any matches found for them in reference or comparison works
    - the claim category counts $A, M, P,$ and N
    - the FAIR Metrics values $F_A, F_M, F_P,$ and F_N
    - the plagiarism/non-plagiarism decision threshold
    - the plagiarism/non-plagiarism decision

Note that we have updated and revised our terminology for this grand challenge. Previously, we used the terms “Quoted” and “Misquoted” (Craig, Ambati, et al. 2019; Craig, Athreya, et al. 2023a) whereas now we use the terms “Attributed” and “Misattributed”. Rephrasing a claim and attributing it to its proper source does not necessarily involve reproducing an exact quote of the claim. Attribution does not necessarily mean inclusion of a formal citation but rather any use of language that identifies a prior work as the source of a claim. An ideal automated solution would be able to identify any such attribution regardless of form. However, for the purposes of this grand challenge, we will only require that an entry identify formal citations.

State of the Art

Current state of the art techniques for detecting plagiarism tend to rely on naive lexical measures like edit distance or black-box machine learning techniques that do not provide insight or accountability. A growing number, including (Kaur et al. 2023) and (Abdi et al. 2015), make use of semantic features to identify passages with similar meaning rather than just similar text. Some use complex hybrid lexical-semantic machine learning methods, as in A. Altheneyan and Menai 2020; A. S. Altheneyan and Menai 2020. Advanced machine learning techniques can even achieve human-level performance at identifying plagiarism obfuscated with automated techniques (Wahle et al. 2022), but such approaches lack transparency and interpretability. The framework for machine learning-assisted plagiarism detection in (Quidwai et al. 2023) does include both sentence-level and document-level metrics, but it is only suitable for distinguishing human-written from machine-generated text in a generic question-answering scenario, not for detecting plagiarism in scholarly publications featuring a mix of original ideas and citation and discussion of the ideas in prior work.

Others ignore the text of the article entirely and focus on similarity in the choice and order of citations, as in (Meuschke, Gipp, et al. 2012), or use citation analysis in conjunction with other methods, as in (Meuschke, Stange, et al. 2019) and (Gipp et al. 2014). Some efforts, such as (Seaward and Matwin 2009) and (AlSallal et al. 2019), do not attempt to compare a work to other works at all and instead look to shifts in writing style within a text as clues that indicate plagiarism, while others look beyond the text of the work to data describing the social network in wich the authors participate (Zrnec and Lavbič 2017).

Another facet of the plagiarism detection that has received increased attention within the past decade is obfuscation of plagiarism via automated text generation. Detection of automatically generated text remains an open problem at which commonly used plagiarism detection tools such as Turnitin and PlagiarismCheck still perform poorly (Weber-Wulff et al. 2023; Odri and Yoon 2023; Khalil and Er 2023). While conventional plagiarism detection techniques typically report that text generated by large language models is original and free of plagiarism, both human reviewers and automated agents often distinguish such content from human-written text by falling back on stylistic peculiarities of the machine-generated text (Santra and Majhi 2023; Gao et al. 2022). As artificial intelligence technologies become more advanced, it is likely that developers will train subsequent generations of algorithms to write more convincingly human-like prose, rendering such approaches ineffective. Instead, to fight plagiarism, it will be necessary to look at the originality of the ideas expressed, regardless of whether the wording comes from a human or a software agent.

Comparatively few researchers have proposed methods to detect plagiarism of non-text content, such as figures in scholarly articles. Some that do attempt to address this problem avoid analysis of the images themselves and focus on the surrounding text, as in (Eisa et al. 2017). Others, such as (Meuschke, Gondek, et al. 2018), attempt to measure image similarity using generic measures of image similarity. Far fewer attempt to derive semantic information from figures, as done in (Eisa 2022). To make the task more manageable, some focus on specific kinds of figures, such as bar charts (Al-Dabbagh et al. 2014) or flow charts (Arrish et al. 2014).

This grand challenge will help focus efforts on explainable, transparent approaches to plagiarism detection that are resistant to paraphrasing and offers both those accused of plagiarism and those whose works have been plagiarized an opportunity to make their cases by performing their own FAIR Metrics analyses. The review of plagiarism detection methods in (Foltýnek et al. 2019) identified the lack of methodologically thorough performance evaluations as a major gap in the plagiarism detection literature. FAIR Metrics analysis addresses this by providing a framework within which developers can apply a wide variety of plagiarism detection techniques so long as they record the results of the comparisons in a way that makes clear which claims of a document match which claims in other documents and whether the authors properly cite those matched claims or falsely represent them as novel. A shared format for recording the results of analyses makes it possible to identify where different comparison methods agree or disagree with each other, how those differences impact the final score of a work, and whether the resulting scores support or contradict human judgments of plagiarism or non-plagiarism. The review in (Foltýnek et al. 2019) also emphasizes the importance hybrid methods will play in the future of plagiarism detection. FAIR Metrics analysis, by providing a common framework for analysis of different media, such as both text and images, will improve the effectiveness of combining different plagiarism detection tools.

Resources

Brain Health Alliance's reference implementation of the Nexus-PORTAL-DOORS-Scribe cyberinfrastructure provides a means of curating and publishing collections of descriptions of resources (C. Taswell 2007; C. Taswell 2010; Craig and C. Taswell 2021). These descriptions can include both lexical and semantic components and have a flexible structure that lends itself to grouping multiple kinds of media by problem domain rather than by data type (Craig and C. Taswell 2021). The BHA website portaldoors.net features multiple live instances of Nexus record repositories. Our practice up until the present has been to store records of known or strongly suspected plagiarism cases in the Fidentinus repository and distribute records for other articles to repositories according to their problem domains. In particular, we have set aside the Martialis repository for resources relevant to fighting plagiarism.

To simplify finding the records of interest to entrants, at the start of the grand challenge, we will replace the contents of these two repositories with records for the entrants to use when training or self-evaluating their entries. Fidentinus will contain records for 12 example articles retracted for plagiarism that have remained retracted for at least 10 years, while Martialis will contain records for 12 example articles published at least 10 years ago that have not been retracted for plagiarism. In general, a decade should be sufficient time for any concerned parties to submit accusations of plagiarism, for journal editors to investigate those accusations and decide whether they warrant retraction, and for authors to voice any objections to retractions. For reference, a survey of genetics articles published between 1970 and 2018 found that the median time taken to retract an article for plagiarism was 2.3 years with a 95% confidence interval of 1.7 to 3.0 years (Dal-Ré and Ayuso 2019).

Because these articles are under copyright of their respective publishers, we will not redistribute PDFs of the articles themselves. Instead, we will select open access articles that themselves only cite other open access articles so that entrants can legally download for themselves the texts of the articles from the publishers. The record for each example article will include the URLs from which to download its PDF, the PDF of the comparison article, and the PDFs of the reference articles in the provenance metadata facet. It will also include a correctly formatted FAIR Metrics analysis record curated by a BHA contributor in the description metadata facet. After reporting the results of the grand challenge, we will make available a second set of 12 example plagiarism cases in Fidentinus and 12 example non-plagiarism cases in Martialis that we will have used to evaluate the entries. These, too, will include URLs from which to download the target, comparison, and reference article PDFs and human-curated FAIR Metrics analysis records.

Evaluation

A single entry must consist of source code written by the entrant in a language of their choice plus English-language instructions for how to run the code. If the entrant includes third-party code with their entry, the documentation must make clear which code the entrant wrote and which came from a third party. The instructions must include URLs from which to download any required third-party software not included with the entry. Any use of this third-party software must not require payment of money to a commercial for-profit vendor. For the sake of cross-platform compatibility, we encourage entrants to use containers or virtual machines as described in the guidelines for the ACM Multimedia Open-Source Software Competition. The entry must be able to run entirely on the local machine without making calls to any remote web services. This requirement will assure that we can examine all parts of the code submitted. Running the entry on the 24 open examples must require no more than 8 hours of CPU time, 8 GB of disk storage, 8 GB of RAM, and 8 GB of GPU memory.

The entry must take as input the target paper, the comparison paper, and all reference papers as PDF files and output a correctly formatted FAIR Metrics analysis report of the target paper. A correctly formatted FAIR Metrics analysis report must have the following properties:

1. It is a valid eXtensible Markup Language (XML) document.
2. It adheres to the FAIR Metrics analysis report XML schema that we will provide.
3. It lists one or more claims of the target work, providing the text of each claim. If the claim is expressed in a figure, it should indicate which figure, e.g., "Figure 2". If the claim is expressed in a table, it should indicate which table and the row(s) and column(s) of the relevant cell(s).
4. It classifies each claim as Attributed, Misattributed, Novel, or Plagiarized.
5. For each Attributed or Misattributed claim, it includes the ID of the work to which the target work attributes the claim. For the purposes of the grand challenge, the ID of the work is the file name of the PDF input file without the ".pdf" file extension.
6. For each Attributed or Plagiarized claim, it includes the text of the matched claim and the ID of the work containing it.
7. It includes the four claim counts: A, M, P, and N.
8. Each count matches the number of listed claims of the appropriate type.
9. It includes the four FAIR Metrics: F_A, F_M, F_N, and F_P.
10. The four FAIR Metric values match what we get by evaluating their formulas with the given values of the four claim counts.
11. It includes the threshold value for F_P used to distinguish plagiarism cases from non-plagiarism cases.
12. It includes the final decision as to whether or not the target work is a plagiarism case.
13. The decision is "Plagiarism" if the value of F_P is below the threshold and "Not plagiarism" if it is above the threshold.

As described above, at the start of the grand challenge, we will make available two repositories, each containing metadata records for 12 target papers with URLs from which to obtain the articles and example valid FAIR Metrics evaluation records. In each case, the reference works will be those that the target work includes in its list of references. For plagiarism cases, the comparison paper will be one of the works from which the target paper plagiarized according to the retraction notice. For non-plagiarism cases, the comparison paper will be a prior work on a related topic. We will also provide the FAIR Metrics Ontology as an OWL ontology, an XML Schema document for a correctly formatted RDF-XML FAIR Metrics analysis report, and a single-page web app that will validate a FAIR Metrics analysis report according to the criteria described above. The FAIR Metrics Ontology is a sub-module of the larger PDP-DREAM ontology and provides semantics for the elements of a FAIR Metrics analysis record (Craig, Athreya, et al. 2023b). By contrast, the XML Schema file will enable the entrant to validate the syntax of the records using standard XML editor software.

After we close submissions for the grand challenge on the training pair of sets of 12 plagiarizing and 12 non-plagiarizing papers, we will evaluate each entry on the test pair of independent sets of 12 plagiarizing papers retracted for plagiarism and 12 non-plagiarizing papers not retracted for plagiarism with their associated comparison works and reference works. We will follow the instructions provided for running each submitted entry. If the instructions are overly long or complex or require excessive downloads and installations of software to the point where they take more than 8 hours to complete, we will disqualify the entry. If we are unable to run the software, we will disqualify the entry.

We will allow the entry a total compute time of 8 hours to complete calculation of FAIR metrics on all 24 examples in the test pair of sets of plagiarizing and non-plagiarizing papers. For each example for which it produces a FAIR Metrics analysis report, we will award 1 point for correct formatting of the report and 1 point for correct classification of each case as plagiarism or non-plagiarism. Thus, the maximum number of points we can award is 48. After evaluating all entries, we will provide each entrant with a report of any issues encountered while running their code, the output obtained for each case, and a table summarizing whether the entry produced correctly formatted output and whether it produced the correct final decision for each case. Because we do not have an objective ground truth for where one statement begins and another ends in a paper, which statements are claims, or which claims belong to which category, we will not use these as criteria for evaluation. After informing entrants of the results, we will make available the metadata records for the cases we used to evaluate the entries. All entrants whose entries produce correctly formatted records and correct decisions for all 24 cases in this test sets will be co-winners of the grand challenge.

Persistence

Brain Health Alliance has maintained publicly available open access at PORTALDOORS.org to the PORTAL-DOORS Project, various multimedia data and metadata repositories across diverse fields of science, and the PDP open source software since 2007. In particular, the more recent data repositories related to plagiarism, ie, the Fidentinus and Martialis repositories, have been maintained since 2018. During and after the grand challenge, we will maintain the records for all 48 test cases of the current grand challenge on plagiarism detection, and will continue to add new records to the Fidentinus and Martialis repositories with each new year of the grand challenge.

Citation

Brainiacs 2024 Volume 5 Issue 1 Edoc G7ECAEAD9
Title: "The Multimedia FAIR Metrics Grand Challenge"
Authors: Adam Craig and Carl Taswell
Dates: created 2024-02-03, received 2024-03-03, presented 2024-06-22, updated 2024-06-30, published 2024-06-30

Contact: ctaswell@bhavi.us
NPDS: LINKS/Brainiacs/Craig2024MMFMGC
DOI: 10.48085/G7ECAEAD9

Affiliations

Brain Health Alliance Virtual Institute, www.BHAVI.us, Ladera Ranch, California, USA.

References

[1] A. Abdi, N. Idris, R. M. Alguliyev, and R. M. Aliguliyev. "PDLK: Plagiarism detection using linguistic knowledge." *Expert Systems with Applications* 42.22 (2015), pp. 8936–8946 (cited p. 2).

[2] M. AlSallal, R. Iqbal, V. Palade, S. Amin, and V. Chang. "An integrated approach for intrinsic plagiarism detection." *Future Generation Computer Systems* 96 (2019), pp. 700–712 (cited p. 2).

[3] A. Altheneyan and M. E. B. Menai. "Evaluation of state-of-the-art paraphrase identification and its application to automatic plagiarism detection." *International Journal of Pattern Recognition and Artificial Intelligence* 34.04 (2020), p. 2053004 (cited p. 2).

[4] A. S. Altheneyan and M. E. B. Menai. "Automatic plagiarism detection in obfuscated text." *Pattern Analysis and Applications* 23 (2020), pp. 1627–1650 (cited p. 2).

[5] S. Arrish, F. N. Afif, A. Maidorawa, and N. Salim. "Shape-based plagiarism detection for flowchart figures in texts." *arXiv preprint arXiv:1403.2871* (2014) (cited p. 3).

[6] A. Craig, A. Ambati, S. Dutta, P. Kowshik, S. Nori, S. K. Taswell, Q. Wu, and C. Taswell. "DREAM Principles and FAIR Metrics from the PORTAL-DOORS Project for the Semantic Web." In: *2019 IEEE 11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)* (June 28, 2019). Pitesti, Romania: IEEE, June 2019, pp. 1–8. DOI: 10.1109/ECAI46879.2019.9042003. URL: https://portaldoors.org/pub/docs/ECAI2019DREAMFAIR0612.pdf (cited p. 2).

[7] A. Craig, A. Athreya, and C. Taswell. "Example evaluations of plagiarism cases using FAIR Metrics and the PDP-DREAM Ontology." *IEEE eScience 2023* 61 (Oct. 13, 2023), pp. 1–2. DOI: 10.1109/e-Science58273.2023.10254806 (cited p. 2).

[8] 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 pp. 2, 4).

[9] 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/links/brainiacs/Craig2021SIMDIR (cited p. 3).

[10] M. M. Al-Dabbagh, N. Salim, A. Rehman, M. H. Alkawaz, T. Saba, M. Al-Rodhaan, A. Al-Dhelaan, et al. "Intelligent bar chart plagiarism detection in documents." *The Scientific World Journal* 2014 (2014) (cited p. 3).

[11] R. Dal-Ré and C. Ayuso. "Reasons for and time to retraction of genetics articles published between 1970 and 2018." *Journal of Medical Genetics* 56.11 (2019), p. 734 (cited p. 3).

[12] T. A. E. Eisa. "Plagiarism detection of figure images in scientific publications." *International Journal of Data Mining, Modelling and Management* 14.1 (2022), pp. 15–29 (cited p. 3).

[13] T. A. E. Eisa, N. Salim, and S. Alzahrani. "Figure plagiarism detection based on textual features representation." In: *2017 6th ICT International Student Project Conference (ICT-ISPC)*. IEEE. 2017, pp. 1–4 (cited p. 3).

[14] F. R. Elali and L. N. Rachid. "AI-generated research paper fabrication and plagiarism in the scientific community." *Patterns* 4.3 (2023) (cited p. 1).

[15] T. Foltýnek, N. Meuschke, and B. Gipp. "Academic plagiarism detection: a systematic literature review." *ACM Computing Surveys (CSUR)* 52.6 (2019), pp. 1–42 (cited p. 3).

[16] C. A. Gao, F. M. Howard, N. S. Markov, E. C. Dyer, S. Ramesh, Y. Luo, and A. T. Pearson. "Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers." *bioRxiv* (2022), pp. 2022–12 (cited p. 3).

[17] B. Gipp, N. Meuschke, and C. Breitinger. "Citation-based plagiarism detection: Practicability on a large-scale scientific corpus." *Journal of the Association for Information Science and Technology* 65.8 (2014), pp. 1527–1540 (cited p. 2).

[18] M. Kaur, V. Gupta, and R. Kaur. "Semantic-based integrated plagiarism detection approach for english documents." *IETE Journal of Research* 69.9 (2023), pp. 6120–6136 (cited p. 2).

[19] M. Khalil and E. Er. "Will ChatGPT get you caught? Rethinking of plagiarism detection." *arXiv* (2023). doi:10.48550/arXiv.2302.04335. DOI: 10.48550/arXiv.2302.04335 (cited p. 3).

[20] M. C. LaFollette. *Stealing into print: Fraud, plagiarism, and misconduct in scientific publishing*. Univ of California Press, 1992 (cited p. 1).

[21] N. Meuschke, B. Gipp, C. Breitinger, and U. Berkeley. "CitePlag: A citation-based plagiarism detection system prototype." In: *Proceedings of the 5th international plagiarism conference*. Citeseer. 2012, pp. 1–10 (cited p. 2).

[22] N. Meuschke, C. Gondek, D. Seebacher, C. Breitinger, D. Keim, and B. Gipp. "An adaptive image-based plagiarism detection approach." In: *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries*. 2018, pp. 131–140 (cited p. 3).

[23] N. Meuschke, V. Stange, M. Schubotz, M. Kramer, and B. Gipp. "Improving academic plagiarism detection for STEM documents by analyzing mathematical content and citations." In: *2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. IEEE. 2019, pp. 120–129 (cited p. 2).

[24] G.-A. Odri and D. J. Y. Yoon. "Detecting generative artificial intelligence in scientific articles: evasion techniques and implications for scientific integrity." *Orthopaedics & Traumatology: Surgery & Research* 109.8 (2023), p. 103706 (cited p. 3).

[25] M. A. Quidwai, C. Li, and P. Dube. "Beyond black box AI-generated plagiarism detection: from sentence to document level." *arXiv preprint arXiv:2306.08122* (2023) (cited p. 2).

[26] P. P. Santra and D. Majhi. "Scholarly communication and machine-generated text: Is it finally AI vs AI in plagiarism detection?" *Journal of Information and Knowledge* (2023), pp. 175–183 (cited p. 3).

[27] L. Seaward and S. Matwin. "Intrinsic plagiarism detection using complexity analysis." In: *Proc. SEPLN*. 2009, pp. 56–61 (cited p. 2).

[28] 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. 3).

[29] C. Taswell. "A Distributed Infrastructure for Metadata about Metadata: The HDMM Architectural Style and PORTAL-DOORS System." *Future Internet* 2.2 (2010), pp. 156–189. ISSN: 1999-5903. DOI: 10.3390/FI2020156. URL: https://www.mdpi.com/1999-5903/2/2/156/ (cited p. 3).

[30] S. K. Taswell, C. Triggle, J. Vayo, S. Dutta, and C. Taswell. "The Hitchhiker's Guide to Scholarly Research Integrity." In: *2020 ASIS&T 83rd Annual Meeting* (Oct. 22, 2020). Vol. 57. Wiley, 2020, e223. DOI: 10.1002/pra2.223. URL: https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/pra2.223 (cited p. 1).

[31] J. P. Wahle, T. Ruas, T. Foltýnek, N. Meuschke, and B. Gipp. "Identifying machine-paraphrased plagiarism." In: *International Conference on Information*. Springer. 2022, pp. 393–413 (cited p. 2).

[32] D. Weber-Wulff, A. Anohina-Naumeca, S. Bjelobaba, T. Foltýnek, J. Guerrero-Dib, O. Popoola, P. Šigut, and L. Waddington. "Testing of detection tools for AI-generated text." *International Journal for Educational Integrity* 19.1 (2023), p. 26 (cited p. 3).

[33] A. Zrnec and D. Lavbič. "Social network aided plagiarism detection." *British Journal of Educational Technology* 48.1 (2017), pp. 113–128 (cited p. 3).

