Preprint abstracts in times of crisis: a comparative study with the pre-pandemic period

Frederique Bordignon[a], Liana Ermakova[b] and Marianne Noel[c]

[a]   *Ecole des Ponts ParisTech, Marne-La-Vallée, France*
[b]   *Univ Brest, HCTI, Brest, France*
[c]   *LISIS, CNRS, INRAE, Université Gustave Eiffel, Marne-la-Vallée, France*

Abstract
The urgency to respond to the COVID-19 outbreak has driven an unprecedented surge in preprints that aim to speed up knowledge dissemination as they are available much sooner than peer-reviewed publications. In this study we consider abstracts of research articles and preprints as main entry points that draw attention to the most important information of the document and that try to entice us to read the whole article. In this paper, we try to capture and examine shifts in scientific abstract writing produced at the very beginning of the pandemic. We made a comparative study of abstracts in terms of their informativeness associated with preprints issued in response to the COVID-19 pandemic and those produced in 2019, the closest pre-pandemic period. Our results clearly differ from one preprint server to another and show that there are community-centered habits as regards writing and reporting results. The preprints issued from the arXiv, ChemRxiv and Research Square servers tend to have more informative (generous) abstracts than the ones submitted to the other servers. In four servers, the ratio of structured abstracts decreases with the pandemic.

Keywords [1]
Scientific abstract, preprints, academic writing, informativeness, COVID-19

1.   Introduction

The urgency to respond to the COVID-19 pandemic (declared on March 11, 2020 by the WHO[2]) has driven an unprecedented surge in preprints [1] that aim to speed up knowledge dissemination as they are available much sooner than peer-reviewed publications. The International Committee of Medical Journal Editors stated that pre-publication dissemination of information critical to public health would not prejudice journal publication in the context of health emergencies declared by the WHO [2]. Although researchers respond quickly to these emergencies, as Zhang et al. [3] show in their comparative study of the response patterns of academia to the outbreaks of four viruses (Ebola, H1N1, Zika and SARS), most articles are published after an epidemic is over. This has been highlighted by a number of studies about the academic response to different epidemic outbreaks: Xing et al. [4] on the 2003 SARS epidemic, Rabaan et al. [5] on the MERS-CoV disease in Saudi Arabia from 2013 to 2015, and Kobres et al. [6] on the Zika outbreak. For Xing et al. [4], possible reasons for this publication delay include "the time taken by authors to prepare and undertake their studies, to write and submit their papers, and, possibly, their tendency to first submit their results to high profile journals".

A preprint is the version of an academic article before it has been submitted for peer review and has been accepted for publication. Preprints related to the COVID-19 crisis are characterized by the urgency of the expected response, unlike papers published several months after the end of a crisis. Here, therefore, we are interested in the study of a preprint rather than a published paper.

[2]   https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

Abstracts are main entry points into scientific articles and preprints. The notion of an abstract is part of everyday language, but it has multiple definitions: the term "abstract" is used loosely to refer to almost any brief account of a longer paper. Most definitions refer to ideal abstracts produced by professional summarizers. According to Orasan [7], it is unlikely that an abstract is intended to replace the whole document. In this paper we will use the definition of the abstract given in [8]: "a concise representation of a document's contents to enable the reader to determine its relevance to a specific information". Thus, the abstract is not only a "mirror" of the document; but it is intended to draw attention to the most important information of the document [7] and to entice us to read the whole article. Reflecting the content of the full text of the document refers to the **informativeness** of the abstract. But the abstract can also be considered as a promotional genre [9] as authors aim to "hook" the readers [10].

In this paper, we aim to capture and examine the shifts in scientific abstracts written at the very beginning of the pandemic. We propose a comparative study of the recent abstracts associated with preprints issued in response to the COVID-19 pandemic (from January 1 to April 12, 2020) with those from the closest pre-pandemic period, produced in 2019.

We aim to investigate whether the COVID-19 crisis had changed the content of the abstracts. This is why we will address the following research question: are abstracts of the preprints produced during the crisis more *informative* than before?

2. State-of-the-art
2.1. Preprints as a response to the crisis

Even though the problem of overfill is important with the COVID-19 pandemic, it has already arisen in the same terms with the previous health crises mentioned above. In an editorial published in 2010 entitled "Journals, Academics, and Pandemics", the PLoS Medicine Editors highlighted an "inherent limitation in the journal publication system with regard to rapid dissemination of results in a time of crisis" [11]. Johansson et al. [12] suggested that preprints could provide a solution: they showed that preprints posted online during the Ebola and Zika outbreaks proposed novel analyses and new data, and sped up knowledge dissemination, as most of those that were matched to later peer-reviewed publications were available more than 100 days before the publication. Less than 5% of Ebola and Zika journal articles were posted as preprints prior to publication in journals, and thus without being peer-reviewed. Although many have warned against considering such papers as more than just a "work in progress" [13, 14], an "interim research product" [15], preprint is the preferred solution researchers have chosen to contribute to the research on COVID-19. It also provides us with an opportunity to access a unique written output, produced *during* the crisis itself, and to seek to identify differences in writing practices. Researchers' guidance evolved with the data and science, the pace of which is rapid during a pandemic of a novel disease. This led to an interdisciplinary response with preprints deposited in various servers. At present, a range of platform types exist from either for-profit or non-profit entities. These include discipline-specific platforms (e.g., arXiv[3], bioRxiv[4], ChemRxiv[5], medRxiv[6]), and generic platforms (e.g., Preprints.org), the latter hosting articles from across a range of disciplines.

2.2. Abstract informativeness metrics

To answer the question of whether it is worth the effort to read full texts or whether the abstract (along with the title and keywords, all of which are freely available) could be sufficient to gain a clear idea of a scientific paper, researchers compare the content of the abstract with the content of the associated full-text. They have established two types of metrics to estimate the quantity of the information given in a summary: (1) questionnaire-based metrics and (2) overlap-based metrics [17].

In the case of questionnaire-based metrics, to compute the level of the retained information, a set of questions issued from the input texts is built and assessors answer these questions reading only the summaries [18]. Otherwise, an assessor may be asked to evaluate the importance of each

---

[3] https://arxiv.org/

[4] https://www.biorxiv.org/

[5] https://chemrxiv.org/

[6] https://www.medrxiv.org/

sentence/passage. One example of questionnaire-based measures is the Responsiveness metric introduced at the Document Understanding Conference (DUC) [19]. The expert nature of the metrics of this type makes further re-use impossible.

A Pyramid score is in the middle between the questionnaire based and overlap-based metrics since its idea is to calculate the number of repetitions of information units of variable length inside a sentence labeled by experts in their own words [20]. However, since the Pyramid score is based on manual assessment not only of the reference summaries, but also of the candidate ones, it can not be re-used to measure the information quantity in new summaries.

The main idea of overlap-based measures is to estimate the proportion of shared words between the gold-standard (i.e. reference) summary and the summary under consideration. One of the most widely used measures is the family of ROUGE (Recall-Oriented Understudy for Gisting Evaluation) metrics proposed at the DUC Conference [21]. While ROUGE is recall oriented, BLEU (Bilingual Evaluation Understudy) is a modified form of precision of a candidate against multiple references [22]. The METEOR (Metric for Evaluation of Translation with Explicit ORdering) metric is similar to BLEU but it is able to treat spelling variants and synonyms [23].

As argued in some papers [17, 24], metrics based on vocabulary overlap are not suitable for measuring the quantity of the information retained in a summary with regard to its corresponding full text since they do not measure the importance of the information presented in different article sections. These metrics are designed to rank candidate summaries (i.e. answer the question: Which summary is better?), but they fail to deal with the comparison of an isolated summary with the full text or with the comparison of metric scores for summaries of different documents. Thus, they are not able to answer the question: Is this summary a true representation of the content of the full text? The second reason why the overlap based metrics fail to answer this question is the lack of interpretability of their output values since in practice, the values tend to be small, e.g. usually ROUGE score is less than 0.2 (see [17] for more details).

Thus, to overcome these issues we introduced a metric called GEM (GEnerosity Measure) [25]. The GEM metric considers the importance of the different sections of a scientific paper based on the comparison of sections from the full paper and sentences from the abstract. The metric GEM is similar to the classification of sentences from medical publication abstracts proposed by Dernoncourt and Lee [26] who released a PubMed dataset for sequential sentence classification where "each sentence of each abstract is labeled with their role in the abstract using one of the following classes: background, objective, method, result, or conclusion". The automatic classification of sentences in medical scientific abstracts was also addressed in Jin and Szolovits' work [27]. In contrast to these two works, GEM is not limited to biomedical texts. Thus, we decided to use GEM for our study as (1) it is designed for the analysis of scientific abstracts and is not limited to medical research, (2) it provides interpretable results, and (3) is publicly available [16].

3.   Data and Methodology

We used a corpus of 23,957 preprints available online [28]. Indeed, this corpus of preprints is designed to allow the comparison of abstracts before and during the crisis. It is based on data indexed by Dimensions[7] and Lens[8]. These preprints come from the following seven servers: SSRN[9], arXiv, medRxiv, bioRxiv, Research Square[10], Preprints.org and ChemRxiv. Similar to Fraser [1], to extract the subset of preprints related to COVID-19 (**COVID-19 corpus**) the following query was used: *"coronavirus" OR "COVID-19" OR "sars-cov" OR "ncov-2019" OR "2019-ncov"*. This corpus contains 3,341 preprints (and their metadata) deposited since January 1, 2020 and retrieved on April 12, 2020, from the seven different preprint servers mentioned above. The control corpus (**Pre-pandemic Corpus**) also contained preprints taken from the same preprint servers published in 2019 with similar subjects and which we could be almost sure were written by the same communities. The control corpus was built so as to be comparable, i.e. to deal with comparable topics to those of the COVID-19 corpus

---

[7] https://www.dimensions.ai/dimensions-apis/

[8] https://www.lens.org/

[9] https://www.ssrn.com/index.cfm/en

[10] https://www.researchsquare.com/

(e.g.: virology, immunology, health policies). This makes it possible to exclude preprints about new or different topics arisen by the pandemic and to ensure that we are assessing similar written content. All queries are available online in the dataset [30].

We tried to use Unpaywall API[11] to find the URLs of the preprints full-texts, but too many files were not available via this tool, especially preprints on ChemRxiv, Research Square and SSRN. As a consequence, we semi-manually retrieved a maximum of full texts in HTML or PDF format.

In order to assess whether the COVID-19 crisis changed writing habits, we decided to evaluate the *informativeness* of the abstracts of the retrieved preprints, and therefore, in the present study used the GEM score : the methodology and the rules are fully described in [25] but to sum up, we can say that the GEM score is calculated as the sum of the weights of the section classes retrieved both from a summary (the abstract in our case) and from a full-text normalized over the total sum of weights of section classes of a full text. Thus, a higher GEM score (i.e. close to 1) corresponds to a higher level of abstract generosity (informativeness), while 0 corresponds to an ungenerous abstract and -1 is assigned when the GEM calculation is unreliable. As in [25], we consider that the GEM score is reliable if at least four out of the seven section classes (Introduction, Methods, Results, Conclusion, Objectives, Limits and Perspectives) are automatically identified in the full text using the GROBID tool [29] for section splitting and their classification algorithm. The section weights were obtained from an online survey conducted among the scientific community. For each sentence in an abstract, the section class is assigned according to the class of the sentence from the full-text with the maximal cosine similarity.

We computed GEM scores (i.e. GEM $\geq$ 0) for 74% of the preprints in the whole corpus, with differences in the proportions among the servers (see Table 1 and the dataset available for reuse [30]).

Table 1

Number and percentage of preprints for which a GEM score was computed

Servers	Pre-pandemic corpus	Covid-19 corpus	All (%)
arXiv	1 149	340	1 489 (81%)
bioRxiv	5 523	248	5 771 (71%)
ChemRxiv	3 296	74	3 370 (72%)
medRxiv	656	803	1 459 (68%)
Preprints.org	252	166	418 (69%)
Research Square	3 173	264	3 437 (90%)
SSRN	1 661	356	2 017 (68%)
All	15 580	2 196	17 776 (74%)

The comparison of abstracts with papers' sections is related to the analysis of their *structure*. Structured abstracts are an emerging trend since they tend to be informative [31] . A *structured abstract* is an abstract with distinct, labeled sections for rapid comprehension (Medline/Pubmed 2018). The IMRAD format (Introduction, Methods, Results, and Discussion) or the CONSORT guidelines for reporting randomized controlled trials (RCTs) are commonly used. Journal guidelines describe how to prepare contributions for submission. Some journals have precise guidelines for what an abstract must include and how it should be structured. Most journals ask for between 150 and 200 words for traditional abstracts (i.e., those without subheadings). Structured abstracts, which are divided into a number of named sections, can be longer than traditional ones [32].

For each abstract, we determine whether it is structured or not by considering the presence of one of the following words within the first 50 characters of the abstract: "background, purpose, objective, aim, introduction, rationale, importance". We assume that the existence of these words is a very reliable indication that the abstract is structured or not, but this method needs further evaluation.

4.   Results

As far as structure is concerned, our first overview of abstracts leads us to find that they are very different from one server to another, and therefore probably from one community to another (fig 1). There are less than 5% of structured abstracts in bioRxiv, arXiv and ChemRxiv and there is no big difference when comparing the two corpuses. For instance, most chemistry journals that feed from ChemRXiv require graphical abstracts rather than structured ones In contrast, abstracts of preprints deposited on Research Square are usually structured, which was the case for up to 97% of pre-pandemic abstracts. But since the crisis, there are less structured abstracts (69%). On medRxiv, SSRN and Preprints.org, there is also a decline in the number of structured abstracts. The decrease of the proportion of structured abstracts could be explained by the fact that authors tried to share their results as soon as possible and as such, may have privileged preprint servers. In contrast to journals, those venues do not request structured abstracts.
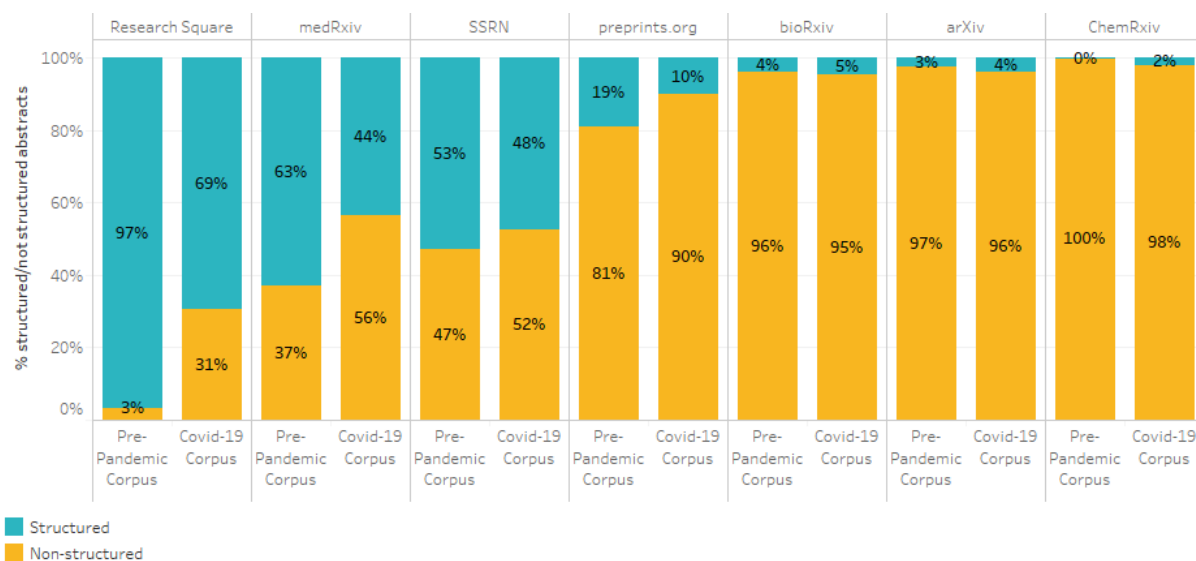

Figure 1: Ratio of structured abstracts depending on the preprint servers (Pre-pandemic and Covid-19 corpora)

Shah et al. [33] showed that even though abstracts display many keywords in a small space there is much more relevant information in the rest of the article. Thus, we decided to calculate GEM score along with abstract structure analysis. The structured abstracts can be viewed as an attempt to summarize each section of the document. In contrast, the GEM score shows which sentence is the closest in the full text. Fig 2 gives a comprehensive synthesis of our results showing differences among the servers that reveal writing habits specific to scientific communities. This trend is evident for four of the seven preprint servers. GEM score varies depending on the server, with a greater increase for the abstracts of arXiv preprints. In contrast to the results of the abstract structure analysis, the GEM score is higher for the COVID-19 preprints. One possible explanation is that the authors of the preprints' tried to share more information in order to attract potential readers to their full text. This contrasts with studies that consider structured abstracts to be generally more informative [31]. This contradiction needs further study to be explained.

The GEM score of abstracts varies a great deal between servers, as shown in Table 2. On Preprints.org, abstracts have been less generous since the pandemic started: -12,20%. On medRxiv, SSRN and bioRxiv, there is not much decrease in the GEM score. In contrast, on ChemRxiv and Research Square, the abstracts are clearly more generous (+7.07% and +12.28%) than they used to be, and on arXiv, we note a significant increase of +17.49%.

With the exception of SSRN, for which the GEM score remained stable, it can be seen that the servers with the lowest rates in 2019 tended to increase and those with the highest rates tended to decrease: the range of its distribution narrowed, going from [0.466; 0.679] to [0.526; 0.637]. This indicates shift towards the homogenization of abstracts so as to make them sufficiently informative for a larger number of readers and people beyond the scientific community.
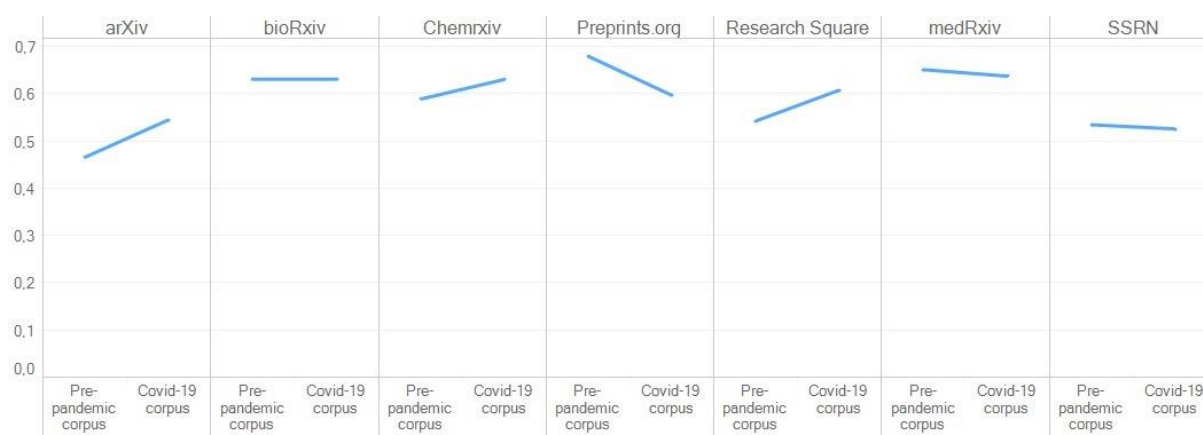
Figure 2: Average GEM score per preprint server before and during the COVID-19 crisis

Table 2

Average GEM score per period and per server

Servers	Pre-Pandemic Corpus		COVID-19 Corpus		Variation
	Avg GEM score	Standard deviation	Avg GEM score	Standard deviation	
Preprints.org	0.679	0.25	0.596	0.27	-12.20%
medRxiv	0.651	0.25	0.637	0.26	-2.14%
SSRN	0.534	0.30	0.526	0.29	-1.56%
bioRxiv	0.631	0.30	0.631	0.28	-0.03%
ChemRxiv	0.589	0.28	0.630	0.28	+7.07%
Research Square	0.541	0.32	0.608	0.26	+12.28%
arXiv	0.464	0.26	0.545	0.26	+17.49%

5.   Conclusion

We found that the general trend for preprint servers is a decrease in the ratio of structured abstracts during the pandemic. We suppose that authors privileged preprint servers to speed up knowledge dissemination. As a consequence, structured abstracts are less frequent. Indeed, these servers do not request structured abstracts in their submission guidelines.

Our study shows that the rate of abstract generosity ranged widely depending on the server. The highest increase was found on the arXiv server, whose readers were certainly limited to a community composed of scientists used to posting there. However, the COVID-19 crisis has attracted a range of new readers to arXiv (researchers developing modelling and predictive works, journalists tracking the news, etc.), the authors probably became aware of this at a very early stage and, as a precaution, made their abstracts more informative considering that it was probably the only piece of text that would be read by most of the readers. Our work also shows that the results are clearly different from one server to another. This is important for two reasons: it shows 1) that there are community-centered habits of writing and reporting results, and 2) it also presents preprints and, more precisely, preprint servers, as possible bases for other types of analyses that would examine communities or even disciplinary distinctions.

The strength of our study is that we considered the peak in production in the very early months of 2020, but also a limitation that calls for work to continue over a longer period of time and to await a return to normality once the crisis is over.

We obtain somewhat contradictory results from the abstract structure analysis and the GEM scores, and this requires further study. The difference in trends of the GEM score and the share of structured abstracts on various preprint servers also requires further analysis.

6. References

[1]   N. Fraser, J. K. Polka, M. Palfy, and J. A. Coates, 'Preprinting a pandemic: the role of preprints in the COVID-19 pandemic', BioRxiv Sci. Commun. Educ., 2020, doi: 10.1101/2020.05.22.111294.

[2]   V. Moorthy, A. M. Henao Restrepo, M.-P. Preziosi, and S. Swaminathan, 'Data sharing for novel coronavirus (COVID-19)', Bull. World Health Organ., vol. 98, no. 3, pp. 150–150, Mar. 2020, doi: 10.2471/BLT.20.251561.

[3]   L. Zhang, W. Zhao, B. Sun, Y. Huang, and W. Glänzel, 'How scientific research reacts to international public health emergencies: a global analysis of response patterns', Scientometrics, vol. 124, no. 1, pp. 747–773, juillet 2020, doi: 10.1007/s11192-020-03531-4.

[4]   W. Xing, G. Hejblum, G. M. Leung, and A.-J. Valleron, 'Anatomy of the Epidemiological Literature on the 2003 SARS Outbreaks in Hong Kong and Toronto: A Time-Stratified Review', PLoS Med., vol. 7, no. 5, pp. e1000272–e1000272, mai 2010, doi: 10.1371/journal.pmed.1000272.

[5]   A. A. Rabaan, S. H. Al-Ahmed, A. M. Bazzi, and J. A. Al-Tawfiq, 'Dynamics of scientific publications on the MERS-CoV outbreaks in Saudi Arabia', J. Infect. Public Health, vol. 10, no. 6, pp. 702–710, Nov. 2017, doi: 10.1016/j.jiph.2017.05.005.

[6]   P. Y. Kobres et al., 'A systematic review and evaluation of Zika virus forecasting and prediction research during a public health emergency of international concern', PLoS Negl. Trop. Dis., vol. 13, no. 10, pp. e0007451–e0007451, 2019, doi: 10.1371/journal.pntd.0007451.

[7]   C. Orasan, 'Patterns in scientific abstracts', in Proceedings of Corpus Linguistics 2001 Conference, Lancaster, 2001, pp. 433–443.

[8]   F. Johnson, 'Automatic abstracting research', Libr. Rev., vol. 44, no. 8, pp. 28–36, December 1995, doi: 10.1108/00242539510102574.

[9]   C. A. Berkenkotter and T. N. Huckin, 'Genre Knowledge in Disciplinary Communication: Cognition/Culture/Power', 1995, [Online]. Available: https://experts.umn.edu/en/publications/genre-knowledge-in-disciplinary-communication-cognitionculturepow.

[10] K. Hyland and P. Tse, 'Hooking the reader: a corpus study of evaluative that in abstracts', Engl. Specif. Purp., vol. 24, no. 2, pp. 123–139, Jan. 2005, doi: 10.1016/j.esp.2004.02.002.

[11] T. Pl. M. Editors, 'Journals, Academics, and Pandemics', PLoS Med., vol. 7, no. 5, pp. e1000282–e1000282, mai 2010, doi: 10.1371/journal.pmed.1000282.

[12] M. A. Johansson, N. G. Reich, L. A. Meyers, and M. Lipsitch, 'Preprints: An underutilized mechanism to accelerate outbreak science', PLOS Med., vol. 15, no. 4, pp. e1002549–e1002549, avril 2018, doi: 10.1371/journal.pmed.1002549.

[13] P. Desjardins-Proulx, E. P. White, J. J. Adamson, K. Ram, T. Poisot, and D. Gravel, 'The Case for Open Preprints in Biology', PLoS Biol., vol. 11, no. 5, pp. e1001563–e1001563, mai 2013, doi: 10.1371/journal.pbio.1001563.

[14] J. A. Teixeira da Silva, 'The preprint debate: What are the issues?', Med. J. Armed Forces India, vol. 74, no. 2, pp. 162–164, avril 2018, doi: 10.1016/j.mjafi.2017.08.002.

[15] D. Poremski et al., 'Moving from "personal communication" to "available online at": Preprint servers enhance the timeliness of scientific exchange', Child Adolesc. Psychiatry Ment. Health, vol. 13, no. 1, pp. 42–42, Oct. 2019, doi: 10.1186/s13034-019-0301-4.

[16] Ermakova, L. (2018). GEM: measure of the generosity of the abstract comparing to the full text. doi: 10.5281/zenodo.1162951

[17] L. Ermakova, J. V. Cossu, and J. Mothe, 'A survey on evaluation of summarization methods', Inf. Process. Manag., vol. 56, no. 5, pp. 1794–1814, Sep. 2019, doi: 10.1016/j.ipm.2019.04.001.

[18] Y. Seki, 'Automatic Summarization Focusing on Document Genre and Text Structure', ACM SIGIR Forum, vol. 39, no. 1, pp. 65–67, 2005.

[19] K. Owczarzak, J. M. Conroy, H. T. Dang, and A. Nenkova, 'An Assessment of the Accuracy of Automatic Evaluation in Summarization', in Proceedings of Workshop on Evaluation Metrics and System Comparison for Automatic Summarization, Stroudsburg, PA, USA, 2012, pp. 1–9, [Online]. Available: http://dl.acm.org/citation.cfm?id=2391258.2391259.

[20] A. Nenkova, R. Passonneau, and K. McKeown, 'The Pyramid Method: Incorporating human content selection variation in summarization evaluation', ACM Trans Speech Lang Process, vol. 4, no. 2, May 2007, doi: 10.1145/1233912.1233913.

[21] C.-Y. Lin, 'ROUGE: A Package for Automatic Evaluation of Summaries', 2004.

[22] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, 'BLEU: a method for automatic evaluation of machine translation', in Proceedings of the 40th annual meeting on association for computational linguistics, 2002, pp. 311–318.

[23] M. Denkowski and A. Lavie, 'Meteor 1.3: Automatic Metric for Reliable Optimization and Evaluation of Machine Translation Systems', Proceedings of the EMNLP 2011 Workshop on Statistical Machine Translation, pp. 85–91, 2011.

[24] A. Louis and A. Nenkova, 'What Makes Writing Great? First Experiments on Article Quality Prediction in the Science Journalism Domain', Trans. Assoc. Comput. Linguist., vol. 1, pp. 341–352, décembre 2013, doi: 10.1162/tacl_a_00232.

[25] L. Ermakova, F. Bordignon, N. Turenne, and M. Noel, 'Is the Abstract a Mere Teaser? Evaluating Generosity of Article Abstracts in the Environmental Sciences', Front. Res. Metr. Anal., vol. 3, mai 2018, doi: 10.3389/frma.2018.00016.

[26] F. Dernoncourt and J. Y. Lee, 'PubMed 200k RCT: a Dataset for Sequential Sentence Classification in Medical Abstracts', in Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers), Taipei, Taiwan, Nov. 2017, pp. 308–313, Accessed: Dec. 04, 2020. [Online]. Available: https://www.aclweb.org/anthology/I17-2052.

[27] D. Jin and P. Szolovits, 'Hierarchical Neural Networks for Sequential Sentence Classification in Medical Scientific Abstracts', in Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, Oct. 2018, pp. 3100–3109, doi: 10.18653/v1/D18-1349.

[28] F. Bordignon, L. Ermakova, and M. Noel, 'A corpus designed to study preprints produced during the Covid-19 crisis and to make comparative studies with the pre-pandemic period'. Mendeley Data, V1, 2021, [Online]. Available: DOI: 10.17632/rn9b93x5d4.

[29] P. Lopez, 'GROBID: Combining Automatic Bibliographic Data Recognition and Term Extraction for Scholarship Publications', 2009, pp. 473–474.

[30] L. Ermakova, F. Bordignon, and M. Noel, 'Data for "Preprint abstracts in times of crisis: a comparative study with the pre-pandemic period"'. Mendeley Data, V1, 2021, [Online]. Available: doi: 10.17632/nsr333t977.1

[31] P. Fontelo, A. Gavino, and R. F. Sarmiento, 'Comparing data accuracy between structured abstracts and full-text journal articles: implications in their use for informing clinical decisions', Evid. Based Med., vol. 18, no. 6, pp. 207–211, décembre 2013, doi: 10.1136/eb-2013-101272.

[32] J. Hartley, 'Current findings from research on structured abstracts.', J. Med. Libr. Assoc., vol. 92, no. 3, pp. 368–371, 2004, doi: 10.3163/1536-5050.102.3.002.

[33] P. K. Shah, C. Perez-Iratxeta, P. Bork, and M. A. Andrade, 'Information extraction from full text scientific articles: Where are the keywords?', BMC Bioinformatics, vol. 4, p. 20, May 2003, doi: 10.1186/1471-2105-4-20.