Noname manuscript No.
(will be inserted by the editor)

A framework for assessing the peer review duration of journals: case study in computer science

Besim Bilalli · Rana Faisal Munir · Alberto Abelló

Received: DD Month YEAR / Accepted: DD Month YEAR

Abstract In various fields, scientific article publication is a measure of productivity and in many occasions it is used as a critical factor for evaluating researchers. Therefore, a lot of time is dedicated to writing articles that are then submitted for publication in journals. Nevertheless, the publication process in general and the review process in particular tend to be rather slow. This is the case for instance of Computer Science (CS) journals. Moreover, the process typically lacks in transparency, where information about the duration of the review process is at best provided in an aggregated manner, if made available at all.

In this paper, we develop a framework as a step towards bringing more reliable data with respect to review duration. Based on this framework, we implement a tool — Journal Response Time (JRT), that allows for automatically extracting the review process data and helps researchers to find the average response times of journals, which can be used to study the duration of CS journals' peer review process. The information is extracted as metadata from the published articles, when available. This study reveals that the response times publicly provided by publishers differ from the actual values obtained by JRT (e.g., for ten selected journals the average duration reported by publishers deviates by more than 500% from the actual average value calculated from the data inside the articles), which we suspect could be from the fact that, when calculating the aggregated values, publishers consider the review time of rejected articles too (including quick desk-rejections that do not require reviewers).

Keywords Peer review process · Review process duration · Review process quality

Besim Bilalli
Universitat Politècnica de Catalunya, BarcelonaTech, Barcelona, Spain
E-mail: bbilalli@essi.upc.edu

Rana Faisal Munir
Eurecat Centre Tecnológic
E-mail: rana.faisal@eurecat.org

Alberto Abelló
Universitat Politècnica de Catalunya, BarcelonaTech, Barcelona, Spain
E-mail: aabello@essi.upc.edu

This is a post-peer-review, pre-copyedit version of an article published in *Scientometrics*.
The final authenticated version is available online at: http://dx.doi.org/10.1007/s11192-020-03742-9

1 Introduction

In the world of science in general and of Computer Science (CS) in particular, a typical measure of productivity is the number of scientific articles published in the journals related to the field of study. The number of publications is an indirect measure of the impact and relevance of the scientist, since the higher it is, the more visible the author is. In addition, at times, this number is critical for getting a better position in universities or in different research departments (R&D). Therefore, the phrase "publish or perish" (sometimes paraphrased as "be cited or perish") is commonly used to describe the pressure researchers feel to publish their research findings in order to stay relevant and successful in the academic community[1].

The number of publications is a matter of concern for every scientist, but even more for young researchers and doctoral students, typically being funded for a fixed period of time (e.g., three years). During this short period, they need to publish scientific articles, most of the times as a requirement to finish their doctoral studies (e.g., minimum 1 journal article) [4, 20, 23], and sometimes for a position in academia or R&D. The same applies to postdoctoral researchers or assistant/associate professors that want to get their accreditation for tenure-track in the duration of their grant/contract (e.g., 4-5 years). Hence, a researcher typically needs to publish a certain number of articles in a given period of time. However, for many different reasons, this task is not an easy one. One of the main reasons is the peer review process and the delays caused due to that.

Generally, in a scientific article, new ideas, solutions, algorithms, or results are presented to the entire community or any other interested party, and prior to publication this undergoes a rigorous peer review process. That is, several other experts on the field (typically three) get to read the work and provide opinions on the matter. Their task is to evaluate the novelty, contributions, and originality of the work [24]. Hence, each one of them provides an answer with regards to whether the article has enough contributions to be published or not. Depending on the comments, the article may require further revision. That is, the authors may need to update or provide further clarifications about their contributions, and this can iterate several times (e.g., 2.5 times on average for CS [13]). This whole process is managed by the editors of the journals and their task consists of finding reviewers, and communicating the decisions to the authors. Furthermore, the editors have the final say based on the reviews received.

The review process is typically based on volunteer work, and thus a task that is supposedly doable in a matter of days gets prolonged to several weeks or months creating a lot of delay for authors whose time is generally limited. In general, there is evidence that this process is unnecessarily slow [17], and some studies have shown that, in some fields, the process is becoming even slower than in the past [1], "which represents a cost to the scientific community and general public" [14], as well as a burden on the academic careers of the authors [5].

The latter two are especially true for the discipline of CS, which has a very fast pace of evolution, where a topic that is hot today can become outdated quickly. Hence, a very typical problem faced by authors, given their quality and time

[1] https://www.researchtrends.com/issue-38-september-2014/publish-or-perish-the-rise-of-the-fractional-author

constraints, is to look for the right journal to submit. Regarding the latter, a survey with hundreds of authors showed that the speed of publication was the third most important factor affecting authors' choice of journal, after topical fit and the quality (in some cases as measured by the impact factor) of the journal [25].

To help with this, recently, several journals provide statistics with regards to the speed of the review process, where the rapidity of the process is meant to attract new submissions. However, rejected articles are also considered in the calculation of such aggregated values (e.g., journals from Elsevier). The latter tends to lower the average values, since rejected articles typically have shorter review times (e.g., desk rejections) [18, 19]. Hence, the values reported by journals tend to mislead the authors and seem to be in contradiction with the general perception of the review process and the studies performed so far [11].

In this work, we propose a framework and a tool — Journal Response Time (JRT), that extracts automatically the information about the review time from the actual published articles. Using JRT, we study the peer review duration of CS journals, however, JRT is publicly available[2] and can be further used as a data collection tool for studying the peer review duration in other disciplines.

In what follows, we firstly describe the multi-phase review process (Section 2). Then, we perform a literature review on peer review duration (Section 3). Next, we present a generic framework (Section 4), that can serve as a starting point for accessing review duration data and highlighting the problems and bottlenecks in the peer review process. We instantiate the framework in terms of a tool, which based on real data that is available inside the articles, allows researchers to compare the response times of different journals. We compare our findings against the response/publication times reported by journals (Section 5), and find out that the values reported are far from what it takes to publish in reality, illustrating the clear deflating effect of rejected articles.

2 Peer review process

Today's world is characterized by rapid advances in science and technology, which is typically by-product of the work performed by researchers, who usually do not develop ready-to-market artifacts, be it software or hardware, but mainly contribute through research articles. A research article. be it published in a conference proceeding or a journal, is the best way to present an idea or solution, since it endures, it is easily communicable, and the presented work is generally reproducible [27]. A published article is immediately on the reach of the community, and other works can then benefit and build on top of the presented ideas. Hence, article publication is critical for the development of a field of study. Yet, the pace of development, for example in CS is very fast — taking into account the progress of the last three decades. Therefore, delays in the peer review process of scientific articles may unintentionally "render the results of the research outdated and of little use by the time they are finally published" [5].

For example, some technologies or solutions are adopted only for a short period of time and then they seize to be so (e.g., the Map-Reduce programming model). Scientific works around such active topics cannot wait for long. Thus, a delay of

[2] `http://dtim.essi.upc.edu/jrt/index.php`

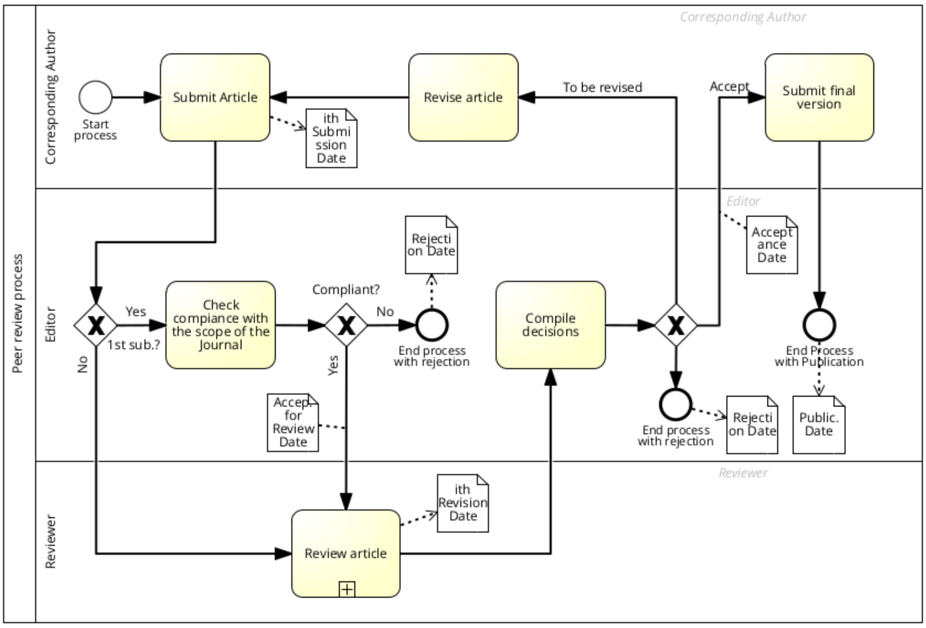


Fig. 1: High level BPMN representation of the peer review process

two years, common in some fields [5], would be difficult to accept for academics in the computer sciences.

2.1 Process

When a research article is submitted for publication, it typically goes through several phases which are depicted in a high level view in Figure 1. Note that each phase is marked with an occurrence timestamp.

First, the *editor* receives the article and performs a first evaluation to verify whether it fits within the scope of the journal. If so, the article proceeds for review, and if not it gets rejected. Next, the editor looks for potential *reviewers*, and once they are found, the article is sent for review. Depending on the reviews, there may be several cycles of revision until the final decision is made.

If accepted, the article gets to be published (in an electronic format and/or hard copy, depending on the journal type: *electronic* or *printed*), and this is done after the final revised version is submitted.

As previously mentioned, due to many factors (e.g., poor reviewer agreement [9], volunteering based process, long idle times in editorial desks, poor editor engagement), this process progresses slowly [17]. Moreover, the process, in terms of expected duration usually lacks in predictability. That is, the expected overall duration for a response is generally not public, and in case this is available, it is usually an aggregated value (e.g., average response time in terms of days) lacking the details of how it was calculated (e.g., what are the numerators and denominators, or whether 1^{st} revision date or n^{th} revision date was used). Furthermore, the lack of transparency manifests in other ways too, for instance in case the process is stuck (e.g., a reviewer is unresponsive), the authors are either not notified or they are notified late, and thus need to be pro-active in order to detect any bottlenecks in the process. All in all, even in the case when some data with respect to expected duration is available, this usually represents an underestimation, or is simply not respected in practice.

2.2 Actors

Different actors are involved in the review process and they are represented through lanes in Figure 1. The *authors* are the ones writing and submitting the scientific article, and typically one of them acts as the *corresponding author*. The *editor* is the person in charge of the editorial process of the journals, usually a leading expert in his/her field (most commonly but not always an academic) appointed and sometimes financially supported by the publisher.

The journal editor is there to receive articles from authors, to judge their relevance to the journal and to refer them to equally expert colleagues — *reviewers*, for anonymous peer review. Peer review is a methodological check on the soundness of the arguments made by the author, the authorities cited in the research and the strength or originality of the conclusions [27].

2.3 Data

In general, the review process generates data of different types ranging from, (i) data about the article itself (e.g., datasets used, code, etc.), to (ii) data about the reviews (e.g., the text of the reviews/comments sent by the reviewers), and (iii) data about the process itself (e.g., dates of when the different activities have occurred). The first is typically available in public repositories. The second remains inside the content management systems of the journals, and is usually considered confidential, only for internal use. Finally, the third one is in one way or another made partially public, and generally includes the *submission date*, *revision dates*, *accepted date*, *online publication date* (see Figure 1). However, these data are not provided in a structured way, hence they are typically either embedded inside the actual document of the article, or embedded inside the website where the article can be downloaded from (see Table 1). Either one or the other, they are basically not useful for machines (i.e., only for humans). That is, in none of the cases these data can be easily accessed through a computer program (e.g., for further analysis). Yet, we contend that such data need to be provided in a machine readable format, in a similar manner as the data for referencing a research article is provided (e.g., DBLP[3] provides such data in RDF/XML format).

3 Studies on peer review duration and transparency

It is estimated that there are more than 28,000 scientific journals worldwide [27]. Although the number of yearly published scientific articles increases, many analysis

[3] `https://dblp.org`

show that the peer review process keeps being slow [13, 17]. One of the main reasons for such a slow process is attributed to the reviewers, who do not give high priority to the work of reviewing, and thus fail to complete a review on time [9, 17, 21]. They are generally overloaded with other academic work and are not paid for the review work. Even financial incentives would not be effective when time constraints are prohibitive [26]. Other potential sources of delay include but are not limited to, poor reviewer agreement (where reviewers do not easily agree on the acceptability of manuscripts), the submission of badly formatted or low quality manuscripts, or even the non promptness of authors when asked to revise the paper in response to review. In general, mainly reviewers and authors have been blamed for delays on the peer review process [17]. However, the problem is more complex. Since authors of manuscripts, reviewers, and editors form a complex network of mutual connections, its structure has a direct influence on the effectiveness of the review process [18], and all of them have their own share of responsibility. When it comes to editors, in [18], they found out that the completion rate and speed is much higher when editors send invitations to known reviewers. Thus, a strategy to achieve shorter times could be to have smaller and tighter communities, but this could also result to be more endogamic.

Yet, regardless of the reasons for the extended review duration, a reaction of authors for the slow process, has been to, in some fields, "bypass the system by publishing their manuscripts in open web repositories prior to submission as working papers (economics) or preprints (computer science and physics, e.g., arxiv[4])" [5]. This has mainly risen from the need to speed up the dissemination of the results. However, the drawback is that sometimes these articles get cited, even though they have not been properly peer reviewed. The latter is raising the fear that the dramatic increase in opportunities to publish results online leads to a decline in the use of peer-reviewed articles of contrasted quality [12].

Although an estimated 2.5 million scientific articles are published annually [27], it is estimated that the number of times the average manuscript is rejected is between three to six times [3]. According to Thomson Reuters [27] there has been an increase in the rejection rate from 59% to 63% between 2005 and 2010. There is also an increase of rejection of up to three times [16] for some journals in the first phase, when the manuscript reaches the editor (the so called desk rejections). According to the authors this is attributed to increased publication pressure [16]. On the other hand, there are growing concerns by the editors that the quality [6] and duration of the review process is negatively affected, and they are challenged in finding sufficient numbers of reviewers in a timely manner [17]. This is not the end of the problem, since even when reviewers are found, other problems may emerge. The most typical ones are ethical problems [22], and as we mentioned earlier, poor reviewer agreementt [6, 9].

Aiming to improve the process, there have been attempts, like Scirev[5], that allow authors to share their experiences with their peers and in a way help others to have an idea about the journal response times, such that they can discriminate between journals that have an opaque and unpredictable review process. This work is based on "crowd-sourcing", that is a user needs to manually introduce the data, which, (i) cannot be easily verified, and (ii) can be at best useful in a

[4] E-prints posted in `http://www.arxiv.org` are not peer reviewed
[5] `https://www.scirev.org`

very active community (i.e., a big enough crowd). Furthermore, given the number of journals and potential authors, this is rather difficult to scale. For this reason, in three years, in Scirev they managed to receive only 3,500 reviews for journals of ten different disciplines [13]. However, extrapolating review durations in those few journals to others is not possible, since there is evidence that it varies not only from one journal to another, but also between different fields, sometimes even within the same broader discipline. For instance, [15] observed a significant difference between conservation biology journals (i.e., 572 days on average) and applied conservation ecology journals (i.e., 249 days).

In comparison, in this work we opt for a more automated process, where we can directly access the data provided by publishers and not wait for authors to feed in the data. Note that there are earlier studies that have collected the data included in published articles [5], but through very labour intensive processes, resulting in studies limited to only hundreds of articles. Our difference is that we automate the process and provide a publicly available tool for the researchers to be able to study the review process using thousands of articles. As a matter of fact, we were able to extract data for more than 136,064 articles in a short period of time (this number keeps growing). Clearly, due to its nature, our approach can be easily extended to cover journals from many different disciplines beyond CS. Yet, mind that the data gathered manually from authors are certainly different from the ones extracted automatically. The former are subjective, thus not 100% reliable (author's impressions of the review process are strongly dependent on the final editorial decision and the speed of the review process [10]), but obviously more extensive, whereas the latter are more reliable but less extensive (typically only dates).

4 Generic framework and tool for the transparent communication of journal response times

In general, we advocate for the need of solving the problem of the unpredictable slow review process, but since this is not in the hands of anyone to change, what we request is that, at least the process becomes more transparent, such that authors are aware of what is the expected review duration of a certain journal. Therefore, we propose a generic framework that would allow researchers to get a picture of how the review process evolves in practice, and how long it takes on average for a journal to provide a response. Based on the proposed framework we implement a tool — JRT, that automatically extracts data from articles. JRT is meant to be a tool of reference for anyone interested in the response times of journals from different domains, but initially the tool is used to scrutinize CS journals. JRT is based on the actual data provided inside the published articles, and thus cannot track the data about rejected articles. Yet, many works have studied rejected articles too, and an interesting finding has been the journal-dependent 'seasonal effect' [2] of the peer review process. In the following, we discuss the overall framework and the challenges faced on instantiating the functional components. Furthermore, we analyze the ethical considerations that may rise around the use of data by JRT.

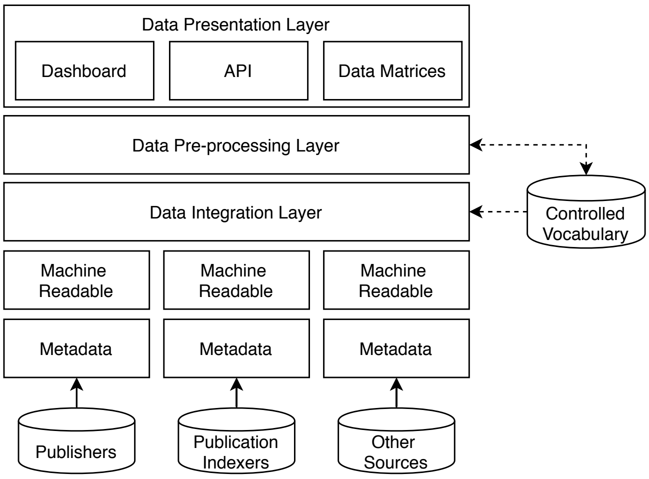


Fig. 2: Generic framework for assessing the peer review duration in journals

4.1 Framework architecture

Our proposed framework involves many actors, who are required to abide to some standards in publishing their data, such that these data can be easily reachable by those who need them. In Figure 2, we depict a high level architecture of the framework. The goals set here, may not be easily reached in the short term, since they require commitment from different parties, but this work is an initial step towards that. The architecture consists of the following layers: *Data sources*, *Data integration*, *Data pre-processing*, and finally the *Data presentation layer*. Different entities may act as sources of information, and thus provide the raw material for a system which aims at helping researchers to, (i) find the expected review times for the different stages of the peer review process, (ii) eventually identify the bottlenecks. In the following, we provide a description of the functionalities for each layer.

4.1.1 Data sources

In the world of scientific article publication, there are many sources that publish partial but different data about the articles or the review process in general. One needs to combine them to obtain a holistic view. For instance, *publishers* keep track of data about the journals, the actual papers, the review process (e.g., date of submission, date of revision, etc.). On the other hand, *article indexers* play the role of search engines for articles, and provide data about the articles from many different journals, from many different publishers (e.g., citation data). Yet, they do not provide data about the review process. Finally, *other sources* exist that may also provide relevant information (e.g., the rank of a journal according to the impact factor). All this information by the different actors is typically provided in a human readable format, hence not machine readable. To this end, not much data is semantically annotated (lack of standardization, interoperability and modularization [7]), which hinders any form of automation when it comes to

Table 1: Metadata provided by publishers (last accessed, September 2020)

Publisher	Reception Date	Revision Date	Acceptance Date	Publication Date	Source
ACM	•	•	•		PDF
Elsevier	•	•	•	•	HTML
IEEE	•	•	•	•	PDF
Oxford	•	•	•	•	HTML
Springer	•		•	•	HTML
Wiley	•	•	•	•	HTML

analyzing the data. We consider that all the actors should agree on the use of a shared vocabulary and provide annotations over the data, which would make them easily accessible. Indeed, this is challenging (many actors are involved), but it is a way to overcome heterogeneity, and allow for building interoperable systems crossing different data sources. However, this is not unattainable, since different actors already publish semantically annotated data (e.g., DBLP, Scopus, etc). The problem is that such data does not include information/metadata about the review process itself.

In our current implementation of JRT, we rely on data from three different sources. Data about journals is collected from Scopus. Data about research articles is collected from DBLP. Finally data about the peer review process of articles is extracted from the actual published articles. As mentioned in Section 2.3 however, the extracted data depends on what is available, see Table 1. Furthermore, the data provided by the publishers is embedded either inside the web page of the article, or in the actual electronic version of the published article. To cope with these two cases, we developed two modules, namely the *Crawler* module, and the *PDF Parser* module, which are explained below. Nevertheless, we contend that the data that needs to be crawled or parsed should be reachable in a much easier way (e.g., through an API), and publishers should provide the necessary tools to extract them.

Crawler. In this module, we read the URLs of the articles from the repository and crawl the entire web page to extract the dates we are interested in. These dates are embedded inside HTML documents and therefore there is no easy way to extract the information, except from parsing the entire document. We additionally use pattern matching to extract the dates and configure it based on the particularities of the journals (e.g., the attached labels may be different).

PDF parser. Publishers that do not embed the information about dates in the web pages of the articles (e.g., IEEE, ACM, etc.), embed such information (sometimes excluding the revision dates) inside the electronic versions of the articles (i.e., PDF files). For these, we configure the PDF Parser module based on the peculiarities of each publisher and extract the provided dates. However, this means that we need to have access to the PDF document of the article in question. Yet, not always we have the permissions to download the documents. Hence, for the cases where we do not have permissions, we allow users (that have access to the documents) to upload their documents such that we can extract the necessary information (the document is immediately deleted once the dates are extracted).

4.1.2 Data integration layer

As mentioned above, data that are semantically annotated can be almost automatically integrated [8]. Yet, when data lacks such annotations the integration process is costly, since it requires a lot of manual effort. In the case of the peer review process, data integration is critical since a researcher looking for a journal typically wants to know about the journal's rank, its impact factor, the publications with their review time, the average review time, and also how it all compares to other journals. Such information helps in finding a journal based on author's constraints. Since in practice data is not annotated using a shared vocabulary, in JRT we download the data from the data sources, and then manually integrate them to provide a global view to the users (i.e., not only review dates, but also journal's ranking, impact factor, publisher, etc.). JRT uses the International Standard Serial Number (ISSN) for identifying and crossing journals.

4.1.3 Data pre-processing layer

The data that has been previously integrated may be raw and lacking many important derived metrics. For instance, given the basic information about the dates of a review process, shown in Table 1, one may derive additional metrics that can be useful (see Table 2 for some examples). The task of this layer is thus to allow for computing such metrics. Yet, the metrics need not be limited only to simple statistics; one can think of more advanced ones, e.g., consuming the information in form of a graph and finding the most influencing author or article.

In JRT, we compute the following basic metrics: *1st Revision Time* - the time it takes until the reviewed version of a paper is submitted; *Acceptance Time* - the time it takes until the article gets accepted; and *Publication Time* - the time it takes until the article gets published. On top of that, we also compute the statistics listed in Table 2. These metrics can be extended with new ones, as soon as more information with respect to the peer review process is provided by publishers.

4.1.4 Data presentation

This layer is the one serving the end users and in JRT it allows for the data to be consumed through an interactive *Dashboard*[6], or through a *REST API*[7] (see Table 2 for some examples) — for users that may want to build applications on top. Furthermore, the third option is to access the data in batch mode, in the form of *data matrices*. Since data about the peer review process has typically been collected manually, the latter may act as a data source for researchers that are interested to study the peer review duration of different disciplines using hundreds of journals and thousands of articles.

Dashboard. We developed a web-based dashboard that can be used by researchers or anyone interested on the peer review process of different journals. Users can (i) search individual journals and see different visualizations for the average acceptance time of a journal in a given year, or average publication time in a given year; (ii) see visualizations which compare different journals based on the

[6] `http://dtim.essi.upc.edu/jrt`
[7] `http://dtim.essi.upc.edu/jrt/api`

Table 2: A subset of metrics provided by JRT

Metric	Description	API call
Article 1st Revision Time	Overall time it took for the 1st revision of an article	/article/[name]/revision_time
Article Acceptance Time	Overall time it took to accept an article	/article/[name]/acceptance_time
Article Publication Time	Overall time it took to publish an article	/article/[name]/publication_time
Avg. 1st Revision Time	Average 1st revision time for a journal	/journal/[name]/avg_revision_time
Avg. Acceptance Time	Average acceptance time of a journal	/journal/[name]/avg_acceptance_time
Avg. Publication Time	Average publication time of a journal	/journal/[name]/avg_publication_time
Min. 1st Revision Time	Minimum time it took for the 1st revision of a journal	/journal/[name]/min_revision_time
Max. 1st Revision Time	Maximum time it took for the 1st revision of a journal	/journal/[name]/max_revision_time
Min. Acceptance Time	Minimum time it took for the acceptance of a journal	/journal/[name]/min_acceptance_time
Max. Acceptance Time	Maximum time it took for the acceptance of a journal	/journal/[name]/max_acceptance_time
Min. Publication Time	Minimum time it took for the publication of a journal	/journal/[name]/min_publication_time
Max. Publication Time	Maximum time it took for the publication of a journal	/journal/[name]/max_publication_time

time it takes to revise, accept, or publish articles; or (iii) search and see metrics for individual articles of a journal, and as such can see the minimum, maximum, or average time it took for articles to be revised, accepted, or published (bar charts, box plots, or line charts help users in consuming such information).

REST API. Together with the metrics listed in Table 2, the API[7] can also be used for accessing other information regarding articles, journals, journal rankings, authors, as well as the dates corresponding to different phases of the review process for each individual article. That is, the information that needs to be provided by the publisher can be easily accessed through this API. Thus, one can think of the API as an accessibility veneer around the non-explicit review process data (e.g., dates).

Data matrices. JRT also allows for the integrated datasets to be downloaded in batch mode for further analysis. These datasets are updated frequently to provide up-to-date information to interested parties. They may be used to find bottlenecks in the review process or research fields in general. In fact, the analysis on top of these datasets may help to come up with potential solutions for improving the process in general.

4.2 Ethical considerations

The crawler generates Internet traffic on publishers' websites to extract dates. Thus, some publishers may consider these requests undesirable, as they potentially place a burden on the target. We took this issue into account and found a balance between the amount of interaction and the merits of the results. To this end, the requests of the crawler module are sent with delays (i.e., 1 request in 5 minutes) to avoid overloading publishers' websites. In addition, we make sure that the requests sent by JRT are all well-formed HTTP requests.

When it is necessary to access the actual documents to extract the dates, we need the electronic versions (i.e., PDF), however sometimes we do not have permissions to download articles from publisher's websites. Thus, we use a crowd-sourcing approach to ask researchers to manually upload PDFs, and then directly extract the dates from the articles. To avoid copyright issues, none of the articles are physically stored in our servers.

5 Lessons learnt

JRT has been put online in January 2019, and since then, based on the analytics of the site, it has attracted users from more than 87 different countries. 3-5 users have visited JRT each day, and approximately 10% of the users have used it several times. In what follows, we provide the results obtained by analyzing the data obtained by JRT about the peer review process of articles in different journals of CS.

We first start by checking the first revision times of all the articles crawled by JRT. We observe that the average of the medians for the articles published in the last eight years (2012-2019) is 6.45 months. Next, we zoom in on ten randomly selected journals and compare the values obtained by JRT with the values reported by publishers about the average first revision times, to find out that there is a huge discrepancy. Next, for the selected journals, we go into more details to check the distributions of first revision times for all of their articles, and see that there is a long tail of submissions (22%) whose first review time goes beyond 10 months, and furthermore, more than half of the papers (55%) take more than one year to be published (publication time). Finally, to check whether similar trends occur with journals from different publishers we compare the distribution of acceptance times (a metric that is common to all publishers) between different publishers and observe that independently of the publisher the overall process takes far too long for more than half of the articles.

5.1 First revision times of articles in JRT

For the time being, we have crawled $136,064$ papers from 374 journals of six different publishers. The publication years of the crawled articles range between 2012 and 2019. On one hand, we would like to highlight that the number of articles keeps growing since we continue to crawl data for more journals and more years, and on the other hand, some articles are disregarded because they do not report any values (i.e., the missing values problem). In Figure 3, we show the distribution of the first revision times of all the articles by the year of publication, and observe that the average medians in the last eight years for all the articles is 6.45 months. Surprisingly, there are articles that took more than four years for their first revision, but we removed them to make the figure more readable. Finally, out

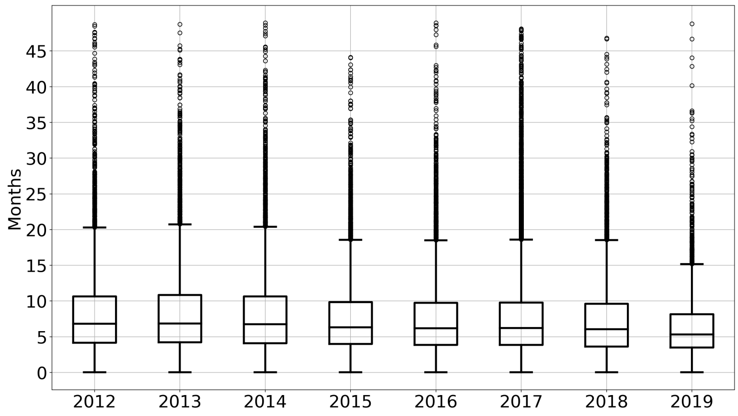


Fig. 3: Distributions of first revision times for all the papers in JRT

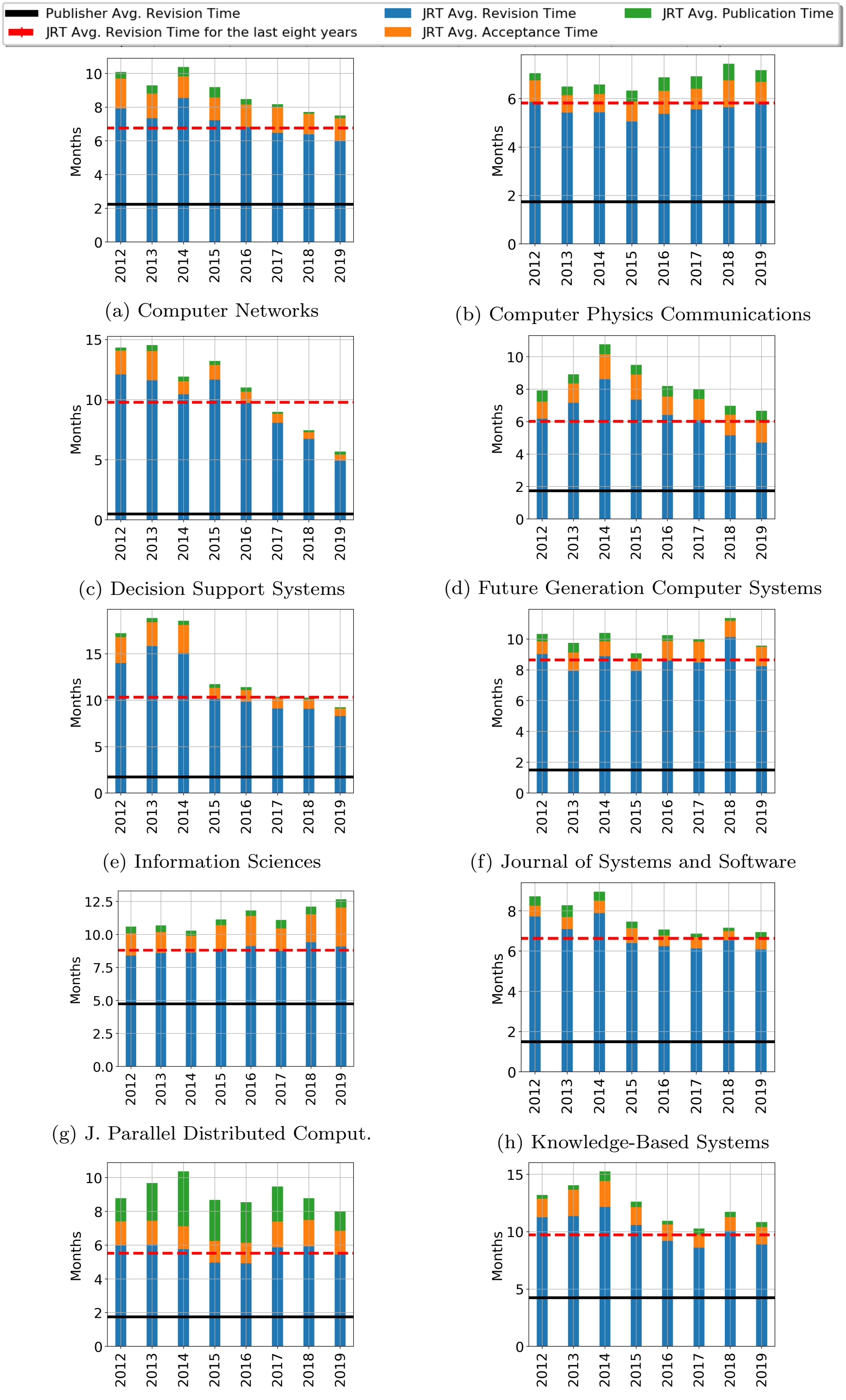
of the total number of articles only 16% took less than three months for the first revision.

5.2 First revision times of ten selected journals

Aggregated metrics provided by different journals (see Table 1) are very relevant for authors, since they give an idea of the duration of the review process. However, not much information is provided with regard to the way these metrics are computed. To this end, there is no information about the year(s) of measurement, or the articles used for the calculation, or whether rejected articles are considered or not (we contacted both Elsevier and Springer, and only Elsevier confirmed that rejected articles are included in their calculations). Hence, to validate the results provided by journals, we compare those aggregates reported on their websites (accessed on September 2020), against raw values obtained by JRT by crawling individual papers.

Figure 4 shows the results for ten journals that were selected based on i) the quality (i.e., belong to the top quartiles according to Journal Citation Reports — they are expected to more easily recruit reviewers, and are more likely to have larger editorial staffs which leads to faster process submissions [5]), and ii) the field of study (e.g., representatives from different fields within the discipline of CS such as, Computer Systems, Theoretical Computer Science, Computer Applications, and Software Engineering — to observe whether differences exist between the fields. Finally, the journals belong to the same publisher, because only Elsevier provides all the data and in the right format (see Table 1).

In Figure 4, on the one hand, using JRT we calculate among others the *average first revision time*, *average acceptance time*, and *average publication time*, which are represented using a cumulative shaded bar per year (2012-2019). On the other hand, we plot the publisher's *average first revision time* with a black horizontal line, because this is the only measure reported by them (since they do not indicate whether it is calculated with respect to a given year, it is represented as a single



(a) Computer Networks
(b) Computer Physics Communications
(c) Decision Support Systems
(d) Future Generation Computer Systems
(e) Information Sciences
(f) Journal of Systems and Software
(g) J. Parallel Distributed Comput.
(h) Knowledge-Based Systems
(i) Neurocomputing
(j) Theoretical Computer Science

Fig. 4: Values computed by JRT against those reported by publishers, for ten randomly selected journals

value). Thus, to facilitate the comparison, we also plot with a dashed red line the JRT *average first revision time* for those eight years.

Firstly, we can see that obviously *first revision time* is clearly underestimating the overall duration of the process (which includes other revision rounds and is actually 22% longer in the average for the selected journals), so journals should be using a more accurate indicator to avoid misleading authors. Even more, we can observe that the values reported for that metric deviate highly from the ones obtained through JRT (the average deviation for the ten selected journals is 527%, and the maximum deviation reaches up to 1,975%). This may be due to the fact that in Elsevier they consider rejected articles too, but we contend that to make their data reliable, publishers/journals must provide precise information on how these values are calculated.

5.2.1 First revision times of articles for ten selected journals

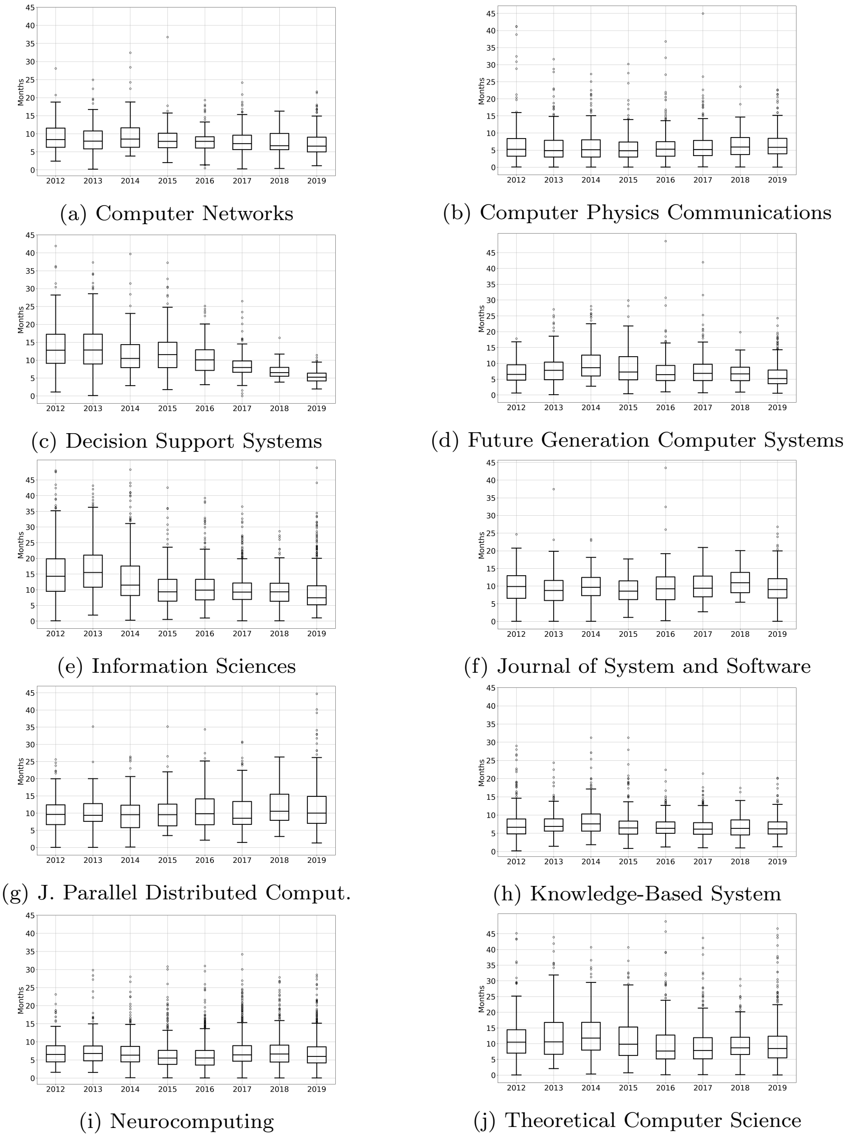
In the previous section, we reported on the aggregated average first revision times of journals. Here instead, we drill down and study the distribution of the average response times of individual articles, which are shown in terms of box plots in Figure 5. Observing the figures, we can see that (i) in most of these journals, there are a lot of outliers in the upper side (105 on average), meaning that for many articles the review process took very long, (ii) the average median for the ten randomly selected journals is 6.8 months, and (iii) the median of the duration of the first revision is not improving in the recent years (except for Decision Support Systems), meaning that the problem of long revision times is still not catching the required attention.

Furthermore, observing both Figure 4 and 5, we can realize that for instance for Computer Physics Communication, Decision Support Systems, Future Generation Computer Systems, Knowledge-Based System, and Neurocomputing, the average first revision time reported by the publisher is less than two months, yet the values calculated by JRT indicate that 75% (above the first quartile) of the articles of these journals took more than 7.1, 11.2, 8.2, 9.5, and 8.4 months (on average for the years considered), respectively. Considering the fact that these journals include the review times of rejected articles too (that are typically much lower), it means that the latter have a huge impact on deflating the average values.

5.2.2 Acceptance times per publisher

In the previous sections we closely analyzed 10 different journals, but they all belonged to a single publisher. In order to check whether journals from different publishers follow similar trends, in this section, we analyze the data per publisher. Hence in Figure 6, we plot a box plot for each publisher using all the journals and all the papers we have collected for them[8]. In Figure 6, we can observe an overall similar trend among the publishers. In the figure, we plot the *acceptance time* in terms of months; this is the only metric that is common to all the publishers. First, observe that for ACM and IEEE the *acceptance time* tends to be higher than the rest, but we believe this is the effect of having less data points. JRT is

[8] The full list of the journals for each publisher can be found here: `http://dtim.essi.upc.edu/jrt/journals.php`



(a) Computer Networks (b) Computer Physics Communications

(c) Decision Support Systems (d) Future Generation Computer Systems

(e) Information Sciences (f) Journal of System and Software

(g) J. Parallel Distributed Comput. (h) Knowledge-Based System

(i) Neurocomputing (j) Theoretical Computer Science

Fig. 5: Distributions of first revision times of articles per journal, for ten randomly selected journals

not able to crawl the same number of papers (homogeneously) for each publisher, since, as mentioned above, some publishers embed the required information inside the actual documents and this makes the problem harder, where articles need to be downloaded manually. In the future, we plan to apply a crowd sourcing effort to get more data for these publishers. For the rest of the publishers, we observe three main patterns, (i) Oxford University Press journals tend to have a lower acceptance time, with a median (4.94) slightly below the rest, (ii) Elsevier,

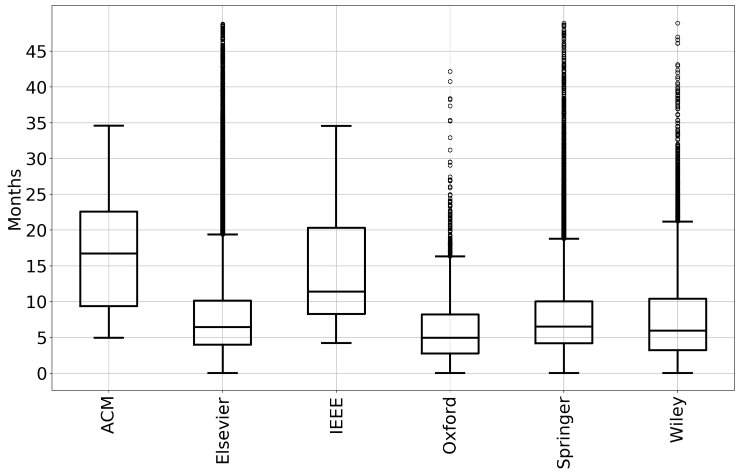


Fig. 6: Distributions of acceptance times for all the papers per publisher

Wiley, and Springer for which we crawl more data, have similar distributions with respective medians of 6.44, 5.94, and 6.5, and (iii) all of the publishers contain a lot of outliers in the upper part, hence a lot of papers that take way too long to be published. In conclusion, there is no significant difference between publishers and the review time remains an open issue that requires the attention of all the parties involved.

6 Conclusions

Based on our study of the peer review duration in journals, we claim that the process takes much longer than needed (as also supported from existing literature), but it specially takes much longer than reported by publishers. To this end, we proposed a framework, and based on that developed a tool (JRT), that extracts the 'hidden' information for accepted articles. Our results showed that, (i) half of all the articles in JRT took roughly more than six months for their first revision, and a quarter of the articles (third quartile) took more than 10 months, (ii) there is no overall evidence neither for improvement, nor for worsening in terms of the review duration for the journals of CS. Furthermore, the values computed by JRT confirm the previous studies in that CS journals tend to have long revision times (e.g., the average response time for CS journals reported in [13] is 5.5 months and the average first revision time computed by JRT is 6.36 months), (iii) for the ten selected journals, there is a huge gap between the average first revision times reported by the publishers and the average first revision times computed by JRT (on average the deviation is greater than 500%, and this we suspect is the effect of considering the review times of rejected articles too). Hence, the values reported by the publishers tend to underestimate the overall review time. Finally, (iv) comparing the acceptance times for different publishers, we observed that, independently of the publisher, the process takes far too long for more than half of the articles.

In conclusion, due to the challenges faced during the collection of the data, we advocate that critical information about the peer review process (e.g., the dates when the article was received for review, when it was accepted, etc.), instead of being buried inside HTML and PDF documents, should be provided in a programmatically consumable way (e.g., through APIs), or machine readable format (i.e., semantically annotated), so that one can easily get the data to analyse the problems in the peer review process.

References

1. Bruce Alberts, Brooks Hanson, and Katrina L. Kelner. Reviewing Peer Review. Science, 321(5885):15–15, 2008.
2. Marcel Ausloos, Olgica Nedic, and Aleksandar Dekanski. Seasonal entropy, diversity and inequality measures of submitted and accepted papers distributions in peer-reviewed journals. Entropy, 21(6):564–575, 2019.
3. Ofer Azar. Rejections and the Importance of First Response Times (Or: How Many Rejections Do Others Receive?). General Economics and Teaching 0309002, University Library of Munich, Germany, 2003.
4. Ever J. Barbero. Journal paper requirement for PhD graduation. Journal of Engineering Education, 2(2), 2008.
5. Bo-Christer Bjork and David Solomon. The publishing delay in scholarly peer-reviewed journals. Journal of Informetrics, 7(4):914 – 923, 2013.
6. Declan Buttler. The Dark Side of Publishing. Nature, 495(7442):433–335, 2013.
7. Cinzia Daraio, Maurizio Lenzerini, Claudio Leporelli, Henk F. Moed, Paolo Naggar, Andrea Bonaccorsi, and Alessandro Bartolucci. Data integration for research and innovation policy: an ontology-based data management approach. Scientometrics, 106(2):857–871, 2016.
8. Cinzia Daraio, Maurizio Lenzerini, Claudio Leporelli, Paolo Naggar, Andrea Bonaccorsi, and Alessandro Bartolucci. The advantages of an Ontology-Based Data Management approach: openness, interoperability and data quality. Scientometrics, 108(1):441–455, 2016.
9. Suhail A. R. Doi, Sherry A. Salzman-Scott, and Adedayo A. Onitilo. Validation of the CoRE Questionnaire for a Medical Journal Peer Review. Accountability in Research, 23(1):47–52, 2016.
10. Ivana Drvenica, Giangiacomo Bravo, Lucija Vejmelka, Aleksandar Dekanski, and Olgica Nedic. Peer review of reviewers: The authors perspective. Publications, 7(1), 2018.
11. Glenn Ellison. Evolving Standards for Academic Publishing: A qr Theory. Journal of Political Economy, 110(5):994–1034, 2002.
12. Glenn Ellison. Is Peer Review in Decline? Economic Inquiry, 49(3), 2007.
13. Janine Huisman and Jeroen Smits. Duration and quality of the peer review process: the author's perspective. Scientometrics, 113(1):633–650, 2017.
14. Charles G. Jennings. Quality and value: The true purpose of peer review? `http://blogs.nature.com/peer-to-peer/2006/06/quality\_and\_value\_the\_true\_pur.html`. Accessed: 2020-09-20.
15. Peter Kareiva, Michelle Marvier, Sabrina West, and Joy Hornisher. Slow-moving journals hinder conservation efforts. Nature, 420:15, 2002.

16. Arie Y. Lewin. The Peer-review Process: The Good, the Bad, the Ugly, and the Extraordinary. Management and Organization Review, 10(2):167–173, 2014.
17. J. Lotriet Cornelius. Reviewing the review process: Identifying sources of delay. The Australasian medical journal, 5:26–9, 2012.
18. Maciej J. Mrowinski, Agata Fronczak, Piotr Fronczak, Olgica Nedic, and Marcel Ausloos. Review time in peer review: quantitative analysis and modelling of editorial workflows. Scientometrics, 107(1):271–286, 2016.
19. Maciej J. Mrowinski, Agata Fronczak, Piotr Fronczak, Olgica Nedic, and Aleksandar Dekanski. The hurdles of academic publishing from the perspective of journal editors: a case study. Scientometrics, 125(1):115–133, 2020.
20. Susi Peacock. The PhD by publication. International Journal of Doctoral Studies, 12:123–135, 2017.
21. Douglas P. Peters and Stephen J. Ceci. Peer-review practices of psychological journals: The fate of published articles, submitted again. Behavioral and Brain Sciences, 5(2):187–195, 1982.
22. David Resnik, Christina Gutierrez-Ford, and Shyamal Peddada. Perceptions of Ethical Problems with Scientific Journal Peer Review: An Exploratory Study. Science and engineering ethics, 14(3):305–210, 2008.
23. Lisa Robins and Peter Kanowski. PhD by Publication: A Students Perspective. Journal of Research Practice, 4(2), 2008.
24. A. J. Smith. The task of the referee. Computer, 23(4):65–71, April 1990.
25. David J. Solomon and Bo-Christer Bjrk. Publication fees in open access publishing: Sources of funding and factors influencing choice of journal. Journal of the American Society for Information Science and Technology, 63(1):98–107, 2012.
26. Leanne Tite and Schroter Sara. Why do peer reviewers decline to review? A survey. Journal of Epidemiology and Community Health, 61(1):9–12, 2007.
27. Mark Ware and Michael Mabe. The STM Report: An overview of scientific and scholarly journal publishing. 03 2015.