DOI:10.1145/3770921

Can a token-based peer-review system turn "lazy" reviewers into "eager" ones?

BY MATTEO FRANCIA, ENRICO GALLINUCCI, AND MATTEO GOLFARELLI

From Volunteerism to Duty: Reforming Peer Review with Tokens

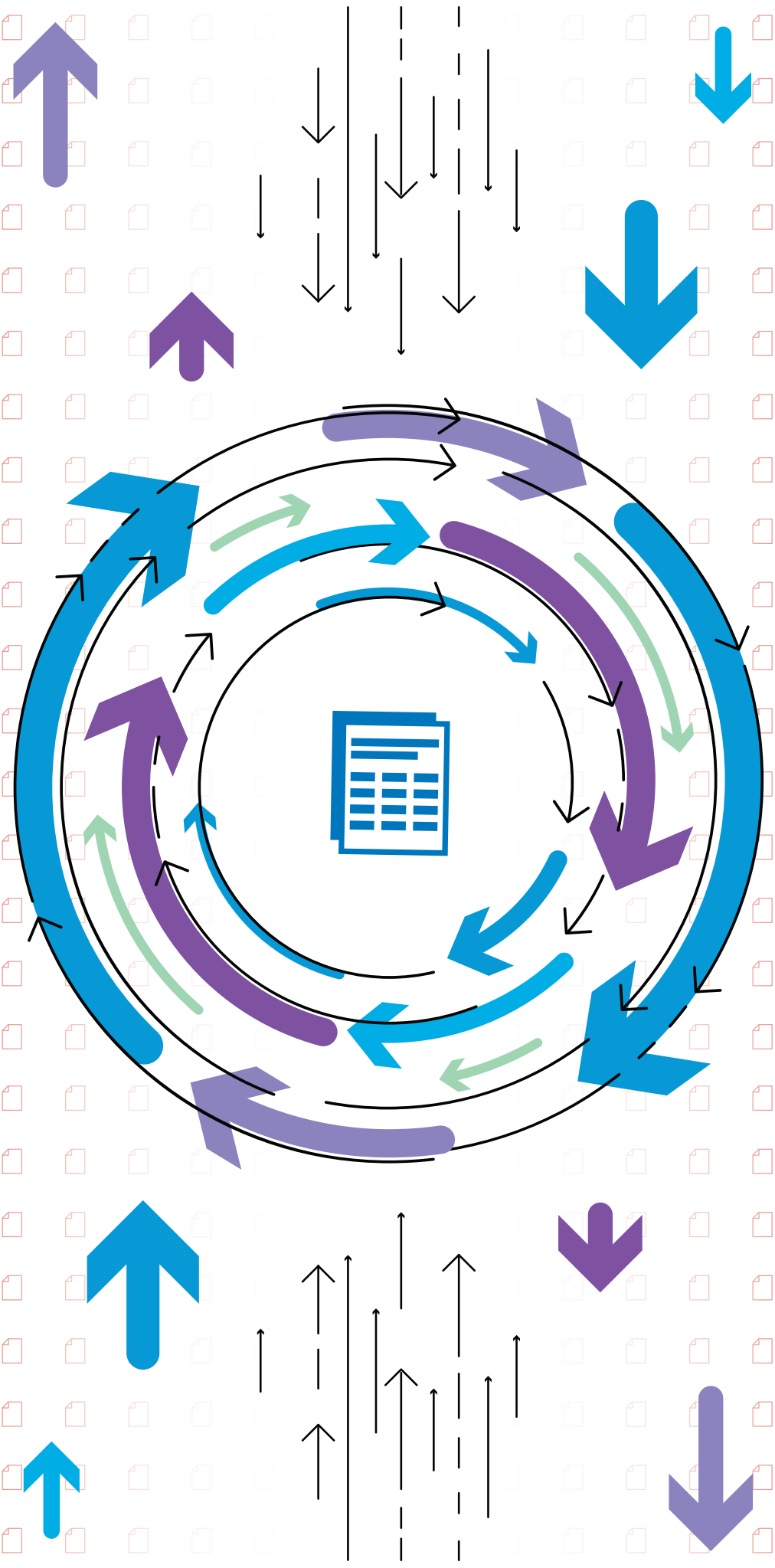
REVIEWING IS A time-honored practice. Peer review, in its modern form, began to spread in the 18th and 19th centuries, becoming standard in the 20th century. Although today software supports reviewers in handling the process and performing automatic controls (e.g., plagiarism detection), the review process has remained fundamentally unchanged for centuries. But over the past decade, there has been a progressive lengthening of the review time for scientific articles.

This problem was felt acutely during the COVID-19 pandemic, when the rapid circulation of pandemic-related findings was particularly critical.[1] Today, it is estimated that the average review time for journals in computer science is about three months, exceeding one year in many cases. The result: a major slowdown in innovation and the circulation of ideas, as well as harm to authors, who see their scientific productivity reduced. And in a fast-moving world, the risk of obsolescence of ideas also increases. In addition, the frantic search for reviewers and the risk of delayed reviews greatly increase the workload of associate editors. A recent survey within the data-management community showed how the problems of high reviewer load and low review engagement are ongoing, highly relevant issues that cause significant concern to the community.[7]

The problem arises from an exponential increase in submissions that is not counterbalanced by the availability of reviewers. This phenomenon seems to be related to emerging geographic areas; for example, for China, India, and Iran, the ratio between the number of submitted papers and reviews is more than 3 to 1.[9] Why the disparity? One answer is that while authoring is considered a central part of researchers' work on which careers are often built, reviewing is done on a voluntary basis and

» key insights

- **The peer-review system is facing a critical imbalance: While paper submissions continue to grow, reviewer availability remains stagnant, resulting in longer decision times and increased editorial workloads.**
- **Voluntary and reward-based initiatives have proven insufficient in resolving this issue.**
- **We need to transform peer review from a voluntary activity into a reciprocal obligation.**
- **Although the basic idea is simple, making it work and accepted by the academic community raises many challenges.**



is bound only by an ethical covenant between researchers. The problem is more severe in scientific journals than at international conferences. At the latter, the program committee and the author community overlap much more, so there is a common interest in completing reviews by the conference deadline.[2]

An orthogonal problem is how to ensure the quality of reviews—and, in turn, the quality of the paper, which partially depends on the work of the researchers who read and evaluate it, requesting additions and explanations, or reject it. Currently, there is no mechanism for certifying the quality of reviewers,[8] and the final judgment is left to the editor's assessment.

Previous attempts to curb the problem have been based on the logic of rewarding and reducing the review effort:

▸ ReviewerCredits[4] is a start-up that offers a reward (even monetary) to researchers who rely on the platform. In exchange for their willingness to review texts, researchers receive virtual credits and a certification index that measures their activity so that not only what they publish, but also what they are asked to evaluate, counts in their resume. The platform also offers a service to publishers: Journals that subscribe to ReviewerCredits in search of reviewers have guaranteed credibility and transparency, so it becomes easier for reviewers to assess the quality of journals and for the latter to find accredited reviewers. The ReviewCredit business model is based on fees paid by publishers.

▸ Peerage of Science[6] is a collaborative peer-review platform designed to improve the quality and efficiency of the scientific review process. Authors submit their manuscripts, which are then evaluated by volunteer reviewers based on quality and scientific rigor. The review process is open and interactive, allowing reviewers to comment on each other's assessments to refine the overall evaluation. Once the peer review is completed, authors can submit the revised manuscript to one of the partner journals, increasing the chances of acceptance thanks to an already established assessment. The system encourages reviewer participation through a scoring mechanism based on the quality of their evaluations.

Despite the attempts made so far, it is apparent that voluntary and reward-based approaches have failed to prove their effectiveness.

▸ Publons[9] was a platform that allowed researchers to track, verify, and gain recognition for their peer-review and academic editing work. Users could record the reviews they had conducted for scientific journals, building a public profile that showcased their contributions to the scientific community. Publons collaborated with publishers to verify reviews without disclosing their content, ensuring the confidentiality of the process. In addition, it provided statistics and useful metrics to monitor the impact of one's reviewing activity. In 2019, Publons was integrated into Web of Science, becoming part of the academic reviewer recognition system managed by Clarivate.

The need for an adequate reward system for reviewers is widely recognized.[5,11,12] However, in the increasingly fast-paced and competitive world of research, there is no certainty that awards (even if prestigious) will attract a sufficient number of reviewers. As emphasized by our survey, none of those interviewed consider these approaches truly useful.[3] In fact, recent studies reveal that awards might actually have a (counterintuitive) discouraging effect,[13] as they make awarded reviewers aware that their efforts are greater than the average, while the awards received are perceived as inadequate. Despite the attempts made so far, it is apparent that voluntary and reward-based approaches have failed to prove their effectiveness.

For these reasons, we propose a compulsory approach, based on tokens, that relies on the following elementary mechanism:

▸ Each group of authors, in order to submit a paper, must provide a number of tokens equal to the number of reviews required by the journal.

▸ For each review made, the reviewer receives one token from the journal.

The authors must also serve as reviewers, and each submission generates an equal number of reviews; this ensures that the system is in balance. Furthermore, as we argue in the next sections, the adoption of a

token-based approach brings with it a series of positive consequences. We emphasize that the token-based approach is designed to support—not replace—the voluntary-based review that we believe is a natural task for a researcher. Apart from the ethical motivation, reviewing is essential for keeping up to date and finding new research ideas. And if the reviews are distributed evenly, reviewers will be able to choose only those that are most interesting to them, which may increase rather than reduce the spirit of volunteerism.

Finally, the token-based mechanism is designed as a security system to solve the *ethical* problem of non-collaborative and lazy researchers. The use of rules and laws to correct unethical behavior is often necessary when civic education alone proves insufficient. Examples range from smoking bans in public spaces to strict penalties for plagiarism in academia.

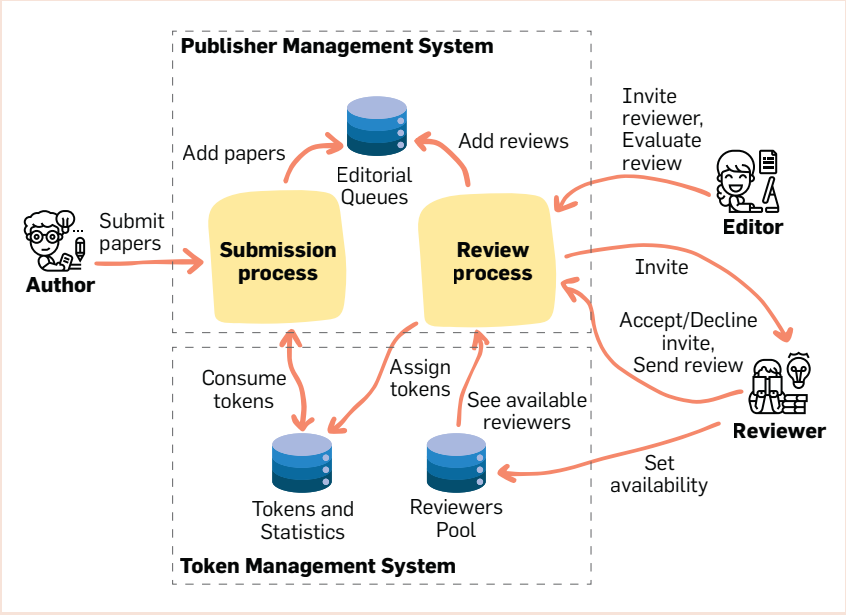
In the remainder of this article, we discuss our proposal in detail by using statistics and simulations to show how the problems we have identified are addressed by the system. Statistics are calculated by integrating DBLP data (with 2.85 million authors and more than 3 million papers published since 1936, excluding publications on arXiv) with Scimago information (with almost 2,000 journals) and statistics manually retrieved from the official websites of more than 700 journals.

System Architecture and Outcomes

The functional system architecture is shown in Figure 1. The Token Management System (TMS) interacts with the Publisher Management Systems (PMS) (e.g., ACM Editorial Manager or IEEE Author Portal) to implement the following process.

Authors submitting a paper must have a number of tokens equal to the number of reviews required by the journal (typically two or three). The PMS contacts the TMS, which verifies the token availability and subtracts them from the authors' availability. Submitted papers are assigned to editors, who start the review process. Reviewers register on the TMS for specific subjects and set their availability to accept reviews. Given the paper's subject, the editors invite available reviewers, though the reviewer still has the option of accepting or declining the invitation. When the reviewer submits a review to the PMS, the editor verifies the review quality and communicates to the TMS the assignment of a token to the author. In the case of a late or low-quality review, the editor may decide to assign only a fraction of a token.

Figure 1. Functional system architecture and data flows.



To prevent fraud and rule exchanges, tokens can be implemented as a digital entity, representing a fungible unit of value. Tokens can be transferred through cryptographic protocols and governed by rules running in the TMS. Authors and reviewers must be identified using a universal identifier such as ORCID. Note that the mechanism ensures authors' and reviewers' privacy, since all information linking an author or a reviewer to a paper remains in the PMS. The TMS also enables the calculation of statistics on the activity and quality of reviewers, which can be useful to both editors and reviewers themselves to certify their work.

Implementation. A simulator for the submission-review process has been implemented in Python, using the agent-based modeling library Mesa.[a] An in-depth description of the simulation model and the source code are available at the Business Intelligence Group website,[3] enabling interested readers to run the simulator with different parameters.

The configuration of the simulation discussed here is based on the statistics for 2024, extracted from DBLP, Scimago, and journal websites (e.g., number of authors per paper, journal acceptance rate, average review time). The model simulates a community of 135,972 groups of around four authors. Each group of authors submits an average of nine papers per year and at the average acceptance rate of journals (19%). Each group of authors also acts as a reviewer. Different groups of authors give widely disparate availabilities for voluntary reviews, from four to 34 reviews per year (19 on average). We assume that eager authors have a lower review time (10 days on average, with SD of 2) than volunteer reviewers (81 days on average, with SD of 37). The limit to voluntary reviews can be exceeded only if the reviewer/author needs tokens.

The simulation was run for 2,190 iterations, with each iteration representing a day (totaling six years). The first three years simulate the peer-review system in current conditions (submissions are free and reviews are

a https://mesa.readthedocs.io/latest/

purely voluntary); then, the token-based mechanism is activated (at Day 0) and the simulation runs for three more years. The simulator models the TMS and the PMSs as a single central entity (where submitted papers are collected and sent for reviewing) and groups of authors as separate autonomous agents. Each agent has a status (either *lazy* or *eager*, effective only when the token-based system is active) and two queues: the *submission queue* (papers ready to be submitted to a journal) and the *reviewing queue* (papers to be reviewed); papers in the reviewing queue are associated with a reviewing date, chosen by the agent. In every iteration, each agent performs the following actions:

- Add a new paper to the submission queue (with probability 2.3%, so as to match the estimated total amount of papers submitted annually) and pick a journal to submit to (with 73% probability of picking a journal with two reviews needed, three otherwise).
- Review papers in the reviewing queue if the reviewing date corresponds to the current iteration. If the token-based system is active, the agent earns one token.
- Submit papers in the submission queue. If the token-based system is active, submission is possible only with enough tokens available (which are then spent). In the absence of enough tokens, the agent can decide to anticipate the reviewing date of papers in the reviewing queue (only in the amount needed to get the required tokens); if the reviewing queue is empty or insufficient to grant the needed tokens, the agent becomes *eager*.

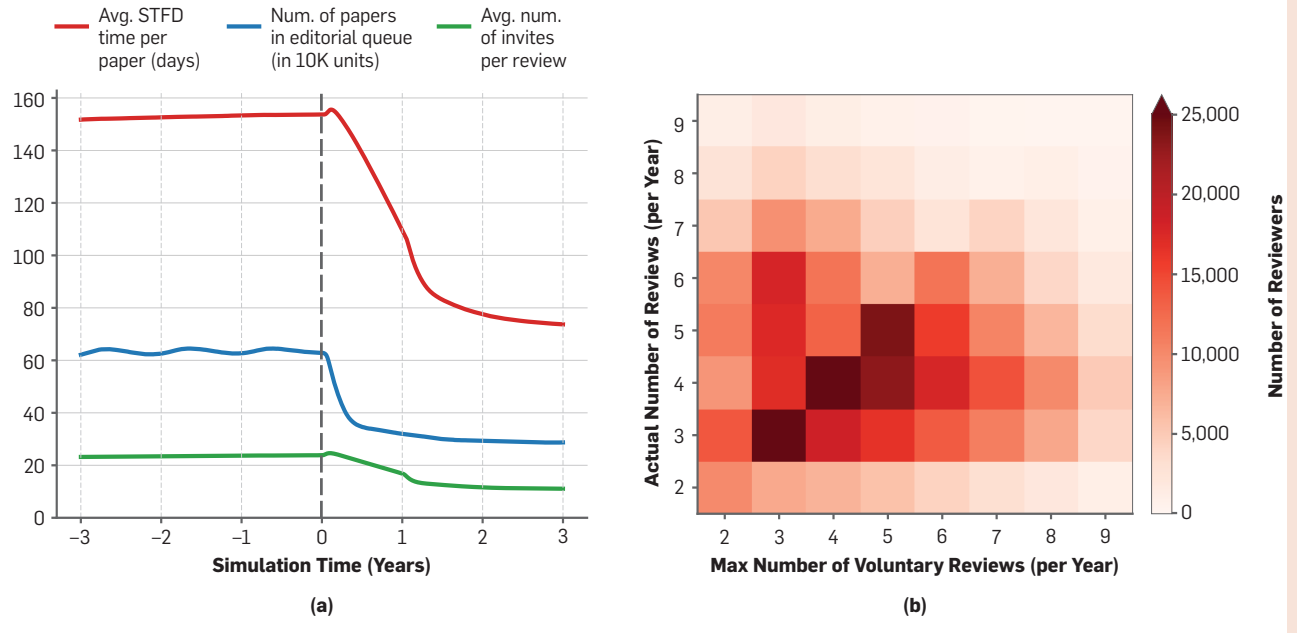
Then, the central entity representing the TMS and the PMSs carries out the following operations for every submitted paper:

- If the paper has been submitted over the last four iterations, it is kept on hold.
- An invite is sent to a random agent for every review needed by the paper; if the token-based system is active, priority is given to eager agents. Agents answer immediately based on their status: A lazy agent accepts with 20% probability (based on survey outcomes) and only if the yearly availability for voluntary reviews has not been met over the past 365 iterations; an eager agent always accepts and possibly goes back to a lazy status (if its reviewing queue satisfies its needs for tokens).
- If all needed reviews have been assigned, the paper is moved to a waiting queue (from which it is removed upon collection of all reviews).

Results. The simulation outcomes (averaged over five runs) are reported in Figure 2, and include the following:

- The activation of the token-based mechanism at Day 0 results in a strong reduction of the *editorial queue* due to authors who are *eager* to review papers to earn the tokens needed for submitting their own papers (blue line in Figure 2a).
- Given the availability of eager reviewers, the number of invitations done by associate editors to assign reviews is also halved (green line in Figure 2a). This results in a further reduction in the submission-to-first-decision (STFD) time, since each submission takes four days in our simulation (red line in Figure 2a).
- The matrix in Figure 2b shows the final distribution of the researchers based on the maximum number of *voluntary reviews* they initially foresaw (*x*-axis) and the number of reviews *actually* done (*y*-axis); for instance,

Figure 2. (a) Results of a six-year simulation of the submission-review process; the token-based mechanism has been activated on Day 0. (b) Heatmap of reviewers, distributed by the actual number of reviews done vs. the theoretical maximum number of reviews they were willing to do (values taken at the end of the simulation).



the top-left cell contains researchers who volunteered for only two reviews but actually did nine, because they needed many tokens to make multiple submissions. Remarkably, cells above the diagonal of the heatmap are attributable to the token-based approach, highlighting lazy researchers who were unwilling to meaningfully contribute to the peer-review process, despite adding to its burden with their own submissions.

Challenge Analysis

In the following sections, we list risks and opportunities related to the system and how they can be addressed.

Token exchange and circulation. Authors publish in different journals and potentially on different subjects; thus, a token could be earned by reviewing a paper from a certain publisher/journal/area and spent on a submission to a different journal/publisher/area. This criterion provides maximum flexibility for the authors to submit a paper and ensures a rapid return on the reviewing effort. Nonetheless, it opens the door to imbalances and opportunistic behaviors: An author/reviewer might try to acquire tokens in journals with a low level of complexity. This would make the number of reviews insufficient for journals with a high level of complexity. The same problem could arise for narrow research areas. To ensure a stronger balance, barriers can be placed on token usability by requiring that a token be spent in the same area, or for the same level of complexity, as the journal from which it was acquired. Managing barriers between areas requires flexibility, and a margin of discretion may be given to the associate editor in evaluating the token's usability. A softer approach could be to assign greater value to reviews based on the journal complexity. The level of complexity of a journal can be determined by metrics such as its impact factor or the quartile it belongs to.

Enforcing review quality and timeliness. The token-based model can also be used to reward the punctuality and quality of reviews. In fact, it is possible to refine the basic mechanism by imposing a rule that a token is granted only if the review is delivered on time and respects a predefined level of quality. Late or low-quality reviews can progressively generate a fraction of tokens. For example, the value of the token can be reduced by a fraction (e.g., 0.1) for each week of delay, and even set to negative for undelivered reviews.

While current voluntary reviews necessarily require longer review times, it is conceivable that they would shorten significantly in a token-based approach.

The quality of a review can be assessed by the editor or by peer evaluation. It is unlikely that standardized, uniform criteria will be adopted across all disciplines, since the concept of quality can vary depending on the field and type of journal. For example, Sizo et al.[10] have surveyed more than 100 metrics. Defining these metrics in detail is beyond the scope of this article, but it is important to underline that our approach allows for managing them. We also need to consider that punitive rewards are not well received by the community.[7]

While current voluntary reviews necessarily require longer review times, it is conceivable that they would shorten significantly in a token-based approach, both because it is in the reviewers' interest to complete the review as soon as possible to have the token available, and because a delay may reduce the value of the token.

Token wear and reissuance. The availability of reviews to be assigned is a necessary condition for not disrupting the flow of submissions. Under a constant flow of submissions, the "liquidity" of the system is guaranteed: Each submission consumes the same number of tokens previously produced through reviews. However, there are some scenarios in which the number of tokens consumed is greater than the number of tokens produced. Over time, this could become problematic, as authors may be unable to submit work due to having no way to earn the required tokens, given the lack of available reviews. This may arise from late reviews gaining only a fraction of tokens, authors quitting or retiring from research and leaving their tokens unused, or volunteer reviewers acquiring more tokens than they actually consume. The problem can be easily overcome by monitoring the quantity of tokens/reviews available and by implementing

ad-hoc reissuance policies when necessary—that is, granting more tokens than are consumed by a submission, equivalent to increasing the number of reviews available. Editors can also exploit this mechanism to support the reviewing process; for instance, reviews of complex papers (or papers on less-attractive subjects) could be incentivized by issuing a higher number of tokens per review. Ensuring a sufficient level of liquidity is important to prevent further unethical behaviors, such as fake coauthorship, where the only contribution made by a coauthor is tokens. We believe this risk becomes real only in the event of a severe token shortage.

Cold start. A submission can take place only after the necessary tokens have been acquired, that is, only after the author has delivered the necessary reviews. This is a problem at system start-up because requiring reviews in advance would impose a delay on all submissions. It must also be considered that, at start-up, the system inherits the previous imbalance; that is, there are a large number of active submissions that lack reviews. Solutions to this problem include:

- Each publisher can initially provide tokens to authors in a fixed number, or based on reviews made in the previous year.
- In the start-up phase, reviews can be delivered within the paper submission period. This avoids delaying submission and ties publication to the performance of the review activity.

In our simulations, we granted three tokens to all the authors when the token-based approach was activated at Day 0. This ensured a smooth transition, since publishing was not stalled. We emphasize that the *three-token welcome bonus* mitigates the start-up phase, but the introduction of a reciprocity solution requires researchers to change their perspective: In order to publish, they must also review. Once this mental shift has been made, variations in productivity during one's career or due to seasonality will not create problems: If one is more or less productive, one will also have to do more or fewer reviews.

Rookie authors. Due to limited experience and skills, new authors may not have the opportunity to access the journal review process and accumulate tokens. This is only a problem in the case where the article is single-authored or all authors are rookies. Statistics indicate this occurrence is very rare—only 2.3% of papers in the DBLP repository in the past 10 years. In other cases, the remaining authors can provide the necessary tokens. To support submissions from rookie authors, it is necessary to provide waivers that allow submission and generate an honor debt (i.e., a negative number of tokens for the author) that must be repaid by performing reviews in subsequent years. Alternatively, a non-coauthor researcher can pay the necessary tokens, playing the role of mentor to the rookie.

Authors are required to change their mindset by reorganizing their way of working, anticipating the need for tokens by acquiring them in advance.

Complexities in transition. The main complexities in deploying the system described here are organizational and social in nature. From a technological point of view, the functional architecture proposed in Figure 1 is not particularly complex and is similar to many cloud-based applications. To ensure the reliability of the entire system, it is important that it can be managed by an independent, nonprofit entity (as in the case of ORCID). Potential points of attention include the definition of the communication interfaces between the TMS and the various PMSs, as well as the evaluation of how to make the TMS redundant in order to avoid it becoming a single point of failure.

It is much more complex to guide the transition and encourage adoption by journal publishers and the scientific community. We draw attention to the following focus points:

- *Selecting the experimentation area:* It is wise to initially apply the approach to a limited disciplinary area, both to check for any side effects when the system is actually in use and to simplify the choice of the alternatives and thresholds discussed in the previous step. For example, reasoning on a single discipline will ease the selection of the review-quality standard. As already mentioned, some effects of our proposal are difficult to assess in advance and require a field-testing phase. In principle, it would be possible to apply the system to a single journal or a single publisher rather than a complete disciplinary area,

but it is difficult to forecast whether authors would be attracted by the reduced review time or discouraged by the compulsory review task.

▸ *Setting the process details and thresholds:* The approach proposed so far still requires choosing the best solution among the alternatives for mitigating the various risks and defining the detailed parameters (e.g., quality metrics and thresholds, token decay rules). At this stage, in order to overcome aversion to change and mistrust, it is essential to involve publishers and associate editors, since the "best choice" is ultimately the one that secures the broadest consensus.

Authors are required to change their mindset by reorganizing their way of working, anticipating the need for tokens by acquiring them in advance. Only researchers with a negative review-submission balance will experience an extra workload, while virtuous researchers and associate editors should see their effort reduced after an initial adjustment phase.

Discussion

The key point of our proposal is shifting from a voluntary or reward-based approach to a compulsory model for peer review. The concept of peer review should no longer be understood solely as "among peers of similar expertise" but also as a system based on "reciprocity." It is important to note that the proposed system does not prevent researchers from voluntarily reviewing papers without necessarily earning tokens.

The in-depth discussion, statistical analysis, and simulations presented in this study demonstrate that the proposed system effectively balances the review-to-submission ratio. It allows for the evaluation and appreciation of review quality, and significantly reduces review times, ultimately shortening the publication process.

To encourage discussion, we close with the results of a survey submitted to a subset of chief and associate editors of computer science journals, chosen by varying subject areas, publishers, and quality of the journals. We collected 30 responses out of 100 invitations.[3] The main outcomes include the following:

▸ The vast majority of responses confirm the existence of the problem and the need to find a solution. The seriousness of the problem of delay in the review process in peer-reviewed journals due to the lack of available reviewers was rated 4 or 5 (on a scale from 1 to 5) by 90% of respondents, and was noted as worsening over the past few years by 97% of respondents. Their experience with the length of the submission-to-first-decision process is very diverse (from two to 30 months), with an average of 8.4 months (SD 6.4); notably, this is worse than the official metrics published by journals, which typically indicate three months on average.

▸ The most recurring concern (45%) is the risk of a decrease in review quality (as some reviewers would be *forced* to do them to earn tokens). We emphasize that our approach foresees a mechanism for rewarding review quality through the possibility of providing portions of tokens if the review does not meet quality standards.

▸ Twenty-four percent of respondents indicated that junior researchers may be disadvantaged by the system (mainly due to a lack of experience in reviewing) and recommend allowing some flexibility in the token mechanism. However, DBLP statistics indicate that junior researchers most often collaborate with seniors, who have the skills and experience to review other papers and earn the needed tokens.

▸ Only 17% are familiar with Publons and related initiatives; among them, none of the respondents perceive them as being able to tackle the problem.

We hope this article sparks a broader discussion within the scientific community, leading to a shift away from the current review-submission imbalance, which poses a significant barrier to scientific production and, consequently, to the rapid dissemination of innovations.

Acknowledgments

We would like to thank our colleagues who encouraged us to pursue this research, convinced of its relevance and potential impact, and the associate editors who participated in the survey by providing their experience and opinions.

References

1. Benjamens, S., Pol, R.A., de Meijer, V.E., and Haring, M.P.D. Peer review during demanding times: Maintain rigorous standards. *Scientometrics* 126, 7 (2021), 6115–6117.
2. Bhowmick, S.S. and Hose, K. Data-driven PC-chair-in-the-loop formation of program committees: An EDBT 2023 experience. *ACM SIGMOD Record* 53, 2 (2024), 68–74.
3. Business Intelligence Group. Token-based system for peer-review (2025); http://big.csr.unibo.it/token
4. Fund, S. and Dyke, G Recognition and reward in peer review: The ReviewerCredits vision. *Minerva Cardiol Angiol.* 72, 3 (Jun. 2024), 220–224.
5. Gasparyan, A.Y., Gerasimov, A.N., Voronov, A.A., and Kitas, G.D. Rewarding peer reviewers: Maintaining the integrity of science communication. *J. of Korean Medical Science* 30, 4 (2015), 360–364.
6. Hettyey, A. et al. Peerage of science: Will it work? *Trends in Ecology & Evolution* 27, 4 (2012), 189–190.
7. Meliou, A. et al. Peer-reviewing processes and incentives: Data management community survey results. *ACM SIGMOD Record* 52, 4 (2024), 41–46.
8. Proctor, D.M., Abraham, R., and Righi, S.E. From novice to expert: Preparing your peer review. *mBio* 16, 6 (2025).
9. Publons Clarivate Analytics. Global state of peer review 2018. (2018).
10. Sizo, A., Lino, A., Rocha, A., and Reis, L.P. Defining quality in peer review reports: A scoping review. *Knowledge and Information Systems* (2025), 1–48.
11. Teixeira da Silva, J.A. and Al-Khatib, A. The Clarivate™ analytics acquisition of Publons—Aan evolution or commodification of peer review? *Research Ethics* 15, 3–4 (2019), 1–11.
12. Van Noorden, R. The scientists who get credit for peer review. *Nature* (2014).
13. Yu, H., Liang, Y., and Xie, Y. Can peer review accolade awards motivate reviewers? A large-scale quasi-natural experiment. *Humanities and Social Sciences Communications* 11, 1 (2024), 1–13.

Matteo Francia is an assistant professor at the University of Bologna, Cesena, Italy.

Enrico Gallinucci is a tenure-track researcher at the University of Bologna, Cesena, Italy.

Matteo Golfarelli (matteo.golfarelli@unibo.it) is a full professor at the University of Bologna, Cesena, Italy.