A Distributed Science Manifesto

Authors

- **Christopher W. Tessum**
  (iD) [0000-0002-8864-7436](#) · () [ctessum](#) · (🐦) [ctessum](#)
  Department of Civil and Environmental Engineering, University of Illinois

- **You?**

Abstract

Traditional institutions for disseminating the results of scientific results ("publication") and allocating limited resources to new scientific inquiries ("funding") have shepharded immense increases in social benefit, but are in some ways not well adapted to the modern scientific endeavor. This document lays out design constraints for a new system, specifically that it should:

1. Sort rather than filter scientific outputs
2. Provide distributed, compensated review
3. Minimize barriers to access to and participation in science
4. Distribute funding decisions
5. Preserve data ownership
6. Overcome incumbent network effects

We then specify a candidate framework that meets these design constraints. This framework—provisionally called the Distributed Science Frameowork (DSF)—acts as both a publisher and funder of science. This framework would contain a monetary credit system, where users could gain credits by providing scientific reviews, and could spend credits to publish and download articles. The cost to publish or download an article would be comparable to the compensation for providing a review, allowing users to participate in the system without having to pay money. Critically, the framework would also include a funding mechanism, where scientific funding organizations and philanthropists could contribute credits which would be distributed among authors according to the quality of their previously contributed articles. This would both provide a distributed, flexible means of funding science and a financial incentive for new users to overcome network effects inherent to established publishers.

(Citations and supporting information to be added throughout.)

Introduction

Scientific discovery is necessarily a slow process: results of individual studies must be verified, vetted, and synthesized, and decisions must be made how to allocate resources among future investigative directions before work can continue. These processes happen both at the micro scale within an individual scientist or lab group, and at the macro scale within members of a society. The document is concerned with the latter scale.

At the society scale, vetting and synthesizing scientific results and allocating resources to future represent a coordination problem: many scientific stakeholders must coordinate to allocate

constrained resources toward publishing existing studies and funding new ones.

Typically, the coordination problem is solved by centralization. Organizations and individuals act as gatekeepers to determine which studies get peer-reviewed and published and which researchers and ideas are funded to continue their work.

The current system has delivered great increases in human understanding and quality of life, and should not be discarded, but has well-documented shortcomings. An incomplete list includes the monetization and the monopolization of access to academic research output by for-profit publishing companies, and the concentration of funding allocation power in the hands of a small number of decision-makers who—like all humans—are biased may have narrow interests. The most common alternative to centralized decision-making — market-based coordination — can be effective for funding research with clear monetization potential, but is not considered effective for research where the profit potential is not clear.

The centralization of scientific funding, synthesis, and dissemination is a bottleneck for the rate scientific advancement. This document describes design constraints for a distributed system for scientific funding, dissemination, and synthesis, and then describes the components of a component distributed science framework.

Design constraints

Sort, don't filter scientific outputs

Traditional scientific journals make binary publishing decisions: articles are either accepted or rejected. This was necessary when physical, printed journals were the only method of dissemination, but is vestigial in the current electronic-delivery paradigm. This is evidenced by the rapid adoption of preprint servers (e.g., https://arxiv.org/) and journals with pre-review "discussion" versions, both of which are much less selective of which articles they will publish than conventional journals.

However, preprint and discussion publishers seem to be supplementing rather than replacing traditional journals, suggesting that the peer-review process serves a valuable function of signaling which articles are the most important and useful.

Provide distributed, compensated review

The traditional academic publishing system relies on reviews volunteered by scientist peers, where the definition of "peer" is determined by article authors and journal editors. In addition to the fact that this results in for-profit publishers benefit from unpaid labor, this system results in reviewers being hard to find, the final number of reviews being small, and the review process taking a long time. Additionally, the amount of thought put into each review can vary widely. Because reviews are not credited or compensated, ideas and attitudes in reviews are skewed toward those of individuals for whom those considerations are less important, and therefore not necessarily representative of scientific stakeholders as a whole.

Scientific review should be compensated, and it should not be limited to the social networks of the article authors and journal editors. Solutions exist for distributed review (e.g., https://prereview.org/, https://openreview.net/) and review recognition (e.g., https://www.sciencemag.org/careers/2016/02/getting-credit-peer-review), but stop short of offering monetary compensation or an equivalent incentive.

Minimize barriers to access to and participation in science

In practice, the legacy publication system is only accessible to scientists who work for institutions that can afford subscriptions to journal packages, and even then it can be difficult to download articles, especially when off-campus. For scientists and aspiring scientists not associated with such an organization, even when between jobs, and especially for those in low and low and middle income countries, access is often not possible. Open-access journals aim to solve this problem, but in practice fees are often just shifted from readers to authors, making open-access publication an option only for highly-funded researchers.

Distribute funding decisions

The current system of centralized funding decisions through public and private granting agencies and foundations, which rely on tacit knowledge accumulated within a relatively small set of individuals to predict which projects will be worthwhile and successful. However, centralized decision-making can be inefficient and can propagate biases held by in-group grant officers and reviewers against out-group applicants and ideas.

We believe a distributed decision framework could provide a complementary utility to the existing centralized systems. Distributed decision systems tend to have different biases from centralized systems, and the software systems which facilitate distributed decision making can be designed to reduce those biases. Buterin et al. (2018; https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3243656) describe a new system for distributed funding decision which will serve as the core of our framework.

Preserve data ownership

Scientific manuscripts are valuable, as are the reviews contributed by scientific peers and data that is collected regarding numbers of article downloads, social media shares and sentiment, etc. Scientific publishers should commit to sharing revenue generated using this data with the individuals the contributed the data.

Overcome incumbent network effects

Existing publishers benefit from strong network effects: scientists want to publish in established, prestigous journals to gain a wide readership and to signal the quality of their work. Any successful upstart must have a strategy from overcoming these network effects and establishing its own.

Distributed Science Framework

Here we propose a candidate framework for meeting the design constraints above, the Distributed Science Framework (DSF). The candidate framework acts both as a publisher and a funder of science. At its core is the concept of Liberal Radicalism (LR; Buterin et al., 2018), a system for the efficient allocation of grant funding. The key components are users, credits, articles, reviews, and grants.

Components

Users

DSF users are online entities that are verified to be individual humans. Verification is critical to avoid the gaming of the system, and will be accomplished by requiring each user to have an email address issued by an institution that is registered with DSF.

Credits

DSF credits are a digital currency system that can be converted to and from dollars or other global currencies. Credits are used to distribute grants and review compensation and to pay for article uploads and downloads.

Articles

Articles are scientific manuscripts, equivalent to articles in existing scientific journals. Authors pay a fee to upload a manuscript, which is published immediately. Authors must provide contribution percents for each author (adding up to 100% for all authors) upon uploading an article. Articles are only available to users; this could be enforced using DRM or blockchain technology, or using weaker methods currently favored by traditional publishers. A new user gets one article download for free, subsequent articles must be paid for using credits. The credits required to download an article will be approximately equivalent to the credits given for reviewing an article, so users not able or wishing to pay for articles with money can pay for a new article by submitting a review of their previous download.

Reviews and metareviews

Users can submit reviews of articles they have read both using quantitative rubrics and text responses, similar to existing journals. Unlike existing journals, reviewers would be compensated using credits. Users will also receive compensation for submitting metareviews (reviews of reviews).

Grants

Agencies and individuals who wish to provide financial funds for science can register grants through DSF. Funders can specify criteria that users must meet to be eligible for a grant (including topics of previous articles, location, etc.) and restrictions regarding what the grant money can be spent on. Then, grant funds will be allocated among qualified users based on the quality of their previous articles.

Article quality will be determined by an algorithm that includes number of unique users that have downloaded the article, citations, and reviews. Review scores will be weighted by results of metareviews and by the quality scores of articles authored by the reviewer.

Funds would be allocated among multiple authors of a paper according to percent contribution values submitted when the article was uploaded.

It is worth noting that this system of grant funding inherently makes decisions on the merits of past contributions rather than ideas for future work. This gives funders less control over which projects are funded than is customary with traditional funding mechanisms, but also creates opportunities for philanthropic individuals and organizations that want to "give back" but do not have the expertise to determine which projects would be useful. This system also would remove the burden associated with preparing grant proposals, and reduce the barrier to participation associated with the common practice of using funding from a previous grant to generate "preliminary results" for use in applying for a future grant.

Operational overhead and revenue model

DSF would need to be implemented and operated by an organization. Functions of the organization would include creating and maintaining the DSF software infrastructure, moderating user content, and policing for abuse of the system.

These activities could be funded by grant overhead. As machine learning and artificial intelligence technology improves over time, AI could potentially be used to synthesize information contained in the articles and reviews into revenue-generating products (for example, automatic literature reviews) that could provide operating revenue and additional compensation for authors and reviewers.

How DSF meets the design constraints

- **Sort, don't filter**: Articles are published immediately, but ranked according to reviews and metareviews. DSF will include searching and sorting algorithms the prioritize the return of high-quality articles.
- **Distributed, compensated review**: Users will be paid for reviews, and financially-constrained users can submit reviews in lieu of paying for article downloads and uploads
- **Minimize barriers to access to and participation in science**: Downloading and publishing articles can be paid for by submitting reviews
- **Distribute funding decisions**: Grants are distributed algorithmically based downloads and a distributed review system.
- **Preserve data ownership**: Users are compensated for reviews and for publishing high-quality articles.
- **Break network effects**: Early adopters be incentivised to participate by the relatively easy access to grant funding. As the number of users increases, the competition for grant funding will increase, as will the prestige associated with having a highly-rated article.

Core assumptions

Here we list some core assumptions built into the design of DSF:

1. Users can be verifiably matched to individual humans: The success of any liberal radical approach relies on successfully mapping system users to individual humans (Buterin et al, 2018). If one human can act as many users, then a distributed system is in danger of becoming a centralized one, centered around that individual. DSF will initially use email identities issued by reputable institutions to perform this verification.

2. Users can be prevented from colluding to affect outcomes: The system could also be subverted by users colluding the boost rating of an article by flooding it with positive reviews and downloads or by creating a citation network among many low quality articles. This type of collusion may not be completely avoidable, but we expect that a combination of automated and human moderation and strict user verification can minimize the damage caused by collusion.

3. It is possible to build an algorithm that ranks article quality in a way consistent with traditional considerations of article scientific rigor, novelty, etc, and that authors will trust it to rate their work: We expect that because the algorithm will rely heavily on quantitative reviews from human users, and because the review scores will be weighted in part based on traditional metrics such as the number citations accumulated by a reviewer's articles, that the system will be familiar and transparent enough to be adopted.

4. Funders will be amenable to giving up control of the projects they fund: This may be objectionable to some funders, but it also may remove barriers to participation by other funders.

5. Monetary incentives for publication will be enough to overcome network affects associated with current funding and publication systems: This seems plausible given that money is known to be a strong motivator.

References