

Decentralised Academic Publishing

Distributed Library Foundation

www.dislib.org

Abstract

Academic publishing is surprisingly big business. The market consists of a few reputable publishers and is predominantly fuelled by public money, where contributors, editors and readers are the same community paying the publishers just to keep the reference what is good and what is bad. The paper proposes a decentralised model of publishing where communities of scientists can be easily formed to create journals, but compared to open access there would be monetary rewards for the people contributing to keep the Journal alive (Editors, Reviewers) or towards sponsorships for conferences. The proposed blockchain technology to achieve this is Ethereum Project in combination with p2p sharing system such as IPFS. Reasoning behind this is argued and economical models to achieve it are proposed.

1 Introduction

Scientific publishing is a global, profitable business, generating revenue of at least \$30 billion annually (\$25 billion journals + \$5 billion books), as reported in 2015. [1]. To put things in perspective, global box office for film industry reported revenues of \$38.3 billion in 2016 [2] and global music industry generated revenues of \$15 billion in 2015 [3], putting top scientists' revenue generation potential somewhere between Adelle and "The Martian"¹. Scientists are furthermore extremely profitable, as shown in 2010 Elsevier's scientific publishing arm profits of 36%. [4]. The revenue is generated over approximately 2.5 million articles, putting the average revenue per scientific article to approx. 1000\$.

To understand this apparently not-so-obvious, but extremely profitable business, one has to understand the motives which drive this industry, mainly fuelled by public grants and donations. The Guardian[4] article gives the historical overview of how this industry came to be, and one can debate if the industry has made this drive to publish or it was vice-versa, however this is out of the scope of this article.

Looking at the state of the science today, dissemination and acceptance of one's scientific work in reputable journals is the most commonly used objective metric to determine a person's or institution's scientific success. In vast majority of the cases, academic success and reputation is dictated by the quantity (i.e. sheer number) and quality (i.e. number of citations, H-Factor etc.) of published academic work. Furthermore, it is expected that the scholar maintains an active role in the scientific community, by promoting overall research in one or more fields, doing

¹Hyperbole?

peer-reviews for these reputable publishers, and being an active member in one or more communities promoted by the same publishers. Finally, in order to do the science work, scientist needs to have the state of the art, and needs the work published by his/hers peers.

These three drives (write-review-read) close the full publishing loop. What is left for a publisher is to establish a name and it becomes a reference to which science community aligns. Publisher does this by maintaining good relationship and rewarding influential scientists, and once established, it is not uncommon for publisher to start charging:

1. Publishing fees: selling the right to the scientist to publish the peer-reviewed article
2. Access fees, paid by libraries: selling the right to the scientist to read the peer-reviewed article
3. Membership fees: selling the right to the scientist to be the members of the community which reviews and decides which article to publish

In today's digital age, where vast majority of the articles are sold in their electronic version, it seems in core, that the only tangible expense for the reputable publisher are server/storage fees for pdf articles, approximately 3-4MB in size, and to maintain a small community to proof-read and edit the articles, without checking the scientific content, as this is done for free by peer-review. With yearly revenue reaching \$30 billion, it seems a bit unbalanced.

Scientific community publishes and reviews the articles which are almost 100% read by the same scientific community, and yet has to pay for each step considerable amount of money to the third person just to keep the reference and consensus what is good and what is bad. It is hard to get out, as this money is then fuelling the same community to agree on the consensus (however with considerable profits). This is a textbook example where blockchain can be of benefit both from the perspective of keeping consensus between parties (who wrote what), and by introducing monetary rewards to the ones keeping the system afloat, by editing and doing peer review, and therefore making consensus and reference for quality. This could arguably be done with far lower overhead, where the excess money could be instead used to fund the research itself, or to reduce the research cost overall.

The rest of the paper discusses the plan, how this median step between pay-for-everything and open-access-voluntary-review can be achieved both from rational and technical perspective.

2 Journals are Scientific Communities

Let's define scientific journal as a core: scientific community, reviewing and choosing content to be published. The main asset for a scientific journal is journal's reputation of hosting high quality publications. This reputation is built by involvement of relevant and influential scientists in the journal's field, and by doing the good peer-review, which if done properly is a significant effort and responsible work, as it defines and distinguishes correct and important scientific work from incomplete data conclusions, speculation and in the worst case plagiarism. Once the positive loop gets formed, there is a race to publish with this high tier journal and high tier journal have the luxury of choosing only the best articles. This is normal, expected and in core, not too bad as it rewards the community organising this.

However, huge pressure built by academia to publish, with just a few high tier (and very wealthy) publishers where these communities can be organised, makes it extremely hard for a new community to form on its own. There are attempts, but they are weak and few. Open-access fully free journals are a very good and intuitive solution, but practice shows that they often fail either in attracting top scientists to promote (no funds to fuel nice conferences for example) or to provide good peer review, and this leaves the free articles to be in average sub-par to the ones published with high tier publishers, and the positive loop described earlier becomes negative. Exceptions, with extremely strong open access communities exist, but the bulk of the high quality articles still go through the few top publishers.

Scientist has absolutely no problem to publish open access, and paradoxically will even choose to pay to have the work hosted at the publisher for open access instead of publishing it for free somewhere. The only service paid is the name and consensus that this article will be properly reviewed before publishing.

Attempts to organise hosting of copyrighted content freely are rightly illegal [5] and will not be discussed. There is also a grey-zone, where authors host the papers written by them and published in a copyrighted journal, just to keep it open-access. The author of this article is not aware on any legal action regarding this, although it is more-less common practice.

It is in any case clear that the science community is completely ok with the fact that they should pay to have their work peer reviewed and published, and expect no profit from it. Therefore, the main point to attack the status-quo should be to funnel this money directly into the communities organising peer review of these articles. This funds can then directly be used to promote the same journal and possibly to fund the peer-review and editing itself. With enough incentive and level field where new communities can be quickly formed, there is a probability that big piece of scientific community would adapt this, which should in theory force the classical publishing to reduce the overhead and adapt to the model, or to be sidelined. Finally, a level field could move the whole system to the economically logical one, where readers pay, and authors get rewarded instead of the one where author pays people to read the article.

The key is transparency of the journal's finances with additional simplicity of setting up a financially backed scientific community (journal), and the competition with the "free market" would drive the quality upwards. Margin to work with is, as discussed, \$1000 per article, in average.

3 Levelling the Field

If there is a strong enough economic incentive for a scientist to contribute to, promote and be a part of a specific scientific community, the scientist might choose to participate in this community, over a more established one, even if the reputation is currently lower. If, at the inception of the new model, scientific community sets a level field in publishing, and incentives are high enough, over time this community will grow and quality of the publications will reach certain level, where it could put major pressure on established publishers to adapt the costs.

The goal would be, instead of just adopting the idea that scientific publishing should be unbounded, we persuade scientists to do the same process as they would with major publishers, but instead of using private entity as a reference and financial service provider, blockchain technologies enable the community itself to be the reference, and distribute the costs between scientists publishing, reviewing and reading the articles. If the field is set where it is completely

transparent where the money is spent and becomes easy for a scientist to create or change the community, over time this would create enough pressure to keep only the high quality and active communities as the valid references, leaving it impossible for the community organisers to use only the name and the legacy as the reference, if they fail to keep the strong scientific community alive.

A community centric journal review process and a platform where communities can be formed are proposed, where each community would have its publication journal and articles can be submitted and obtained, with the following rules:

1. Article is the record on the blockchain. It has a list of authors - their unique blockchain addresses, and unchangeable link where the article can be obtained.
2. Journal is defined as a distributed application, running on blockchain. It has a list of owners, editors and reviewers as unique addresses belonging to the people involved, as well as the rules of publishing.
3. Review process is kept separate, and once the article is "publication ready", editors of the journal can mark it to be published
4. In order to attach an article to the list of published articles, the author can then pay certain amount of crypto-currency to the journal's address, as defined by the journal's rules. This value can also be zero, for open access free publishing.
5. Journal keeps the list of published articles as the list of addresses pointing to the blockchain records of these articles.
6. The funds belonging to the journal can only be split between owners, editors and reviewers, and all the transactions are publicly available.

This should be a changeable process, and the rules proposed are the starting point. Later, journal might choose to completely remove the publication fee, and only charge the people to download articles (i.e. by keeping the list of decryption keys to the articles), or to have institutional fees (i.e. libraries keep the decryption keys to the articles). Main point, which should be kept is the full transparency of the money involved, and this is anyway dictated by the blockchain. Arguing point is that this transparency would include the financial aspect in the journal's reputation (which is as mentioned the main asset of a journal), forcing a more sensible usage of the funds obtained, and keeping the full amount of the funds within the scientific community, to be used on improvement of the services or research itself (conferences, standardisations, grants, development funds etc.).

4 Blockchain Technologies

In order to achieve this, one needs a distributed ledger where research funds can be converted to tokenised currency and vice versa. This ledger should also have the possibility to create distributed applications to be journals, and to keep the records of publications. Finally, an unchangeable distributed storage is needed to keep the data itself (papers). One can attack this problem in many ways, and the paper will note a few:

1. Creating the distributed ledger from zero, and organising the community to host it.
2. Utilising existing technologies, networks, and crypto-currency for funding.
3. Utilising existing technologies, networks, and creating a new crypto-currency for funding.

Custom ledger is the solution with the least overhead, and it would be the preferred one technologically. C. Janze [6] presents a solution for the custom blockchain (APOLLO) with currency to be used specifically for scientific publishing. It has built in functionality suited to perform and reward peer-review, making it very similar to the proposed idea and a potential candidate for the technology. The optimal solution might however run into difficulties, as it would require formation of the new community to actively develop the blockchain technology in order to make it user friendly to the whole scientific community (including non-technical part), as well as hosting by the scientific communities, libraries and institutions, with incentives to do so. Furthermore, for any implementation of the custom blockchain, one has to note that the changes to the core functionality down the road are extremely difficult, if not plain impossible, and therefore this would need to be very flexible ledger to be universal to many applications.

A possibility could be to use the "private" blockchain with proven code (i.e. Ethereum), which could solve the flexibility problem, however the acceptance and participation would still be the issue.

Utilising existing technologies could be the quicker way to the market and maybe a more stable blockchain backbone to host the journals. Ethereum [7] is the active project, providing the possibility to write smart contracts which will act as distributed journals, article records and database records. To host the articles, one just needs an unchangeable record database, and ipfs [8] could be the solution for that.

If one chooses the Ethereum project as the backbone, there is an open question if it is better to utilise the Ether (this is the crypto-currency fuelling the blockchain) as the currency to purchase/publish the papers, or it would be beneficial to have a specific currency with which the communities and journals would be funded. Using Ether comes with lower overhead, and could be a viable solution, however there needs to be a high incentive for early adopters to start organising communities and journals where they could benefit from the acceptance and development of the whole ecosystem. If the system is made, where early adopters have higher incentive to promote the whole ecosystem compared to the single journal they host, there is a reasonable expectation that they would do so, which will make higher pressure on the established publishers to accept the new currency, financing the early adopters. Therefore the big players would be forced to adhere to the consensus, compared to becoming just another player in the blockchain, with enough funds to change the rules in its favour. Creating a new crypto-currency which should be utilised solely for academic publications allows the early adopter scientists and journals to obtain the larger amounts of it with lower costs and higher risks, and they would have high interest in growing the ecosystem itself, compared to growing the single (or few) journals they edit, effectively increasing the competition.

M. Swan [9] proposes exactly this, in form of Journalcoin which *"could be issued as the token system of the publishing microeconomy to reward contributors, reviewers, editors, commentators, forum participants, advisors, staff, consultants, and indirect service providers involved in scientific publishing. This could help improve the quality and responsiveness of peer reviews, as reviews are published publicly, and reviewers are rewarded for their contribution."*

Spearpoint [10] and Swist/Magee [11] also discuss this issue in scientific publishing with the same conclusion regarding the specialised currency to be used for scientific publishing.

5 Foundation

If this path is chosen, problem lies in the potential centralisation and extreme influence by the currency creator to the whole ecosystem. Although this will be the case in the beginning, the author believes that if entity is formed with clear final goal to self-destruct after the system becomes accepted and decentralised, one can limit the centralisation of influence. It is proposed:

1. To generate a public token and open science foundation. This token will be later used as a currency in scientific publishing.
2. $n\%$ of the total tokens are sold initially to finance the development of the technology, and to promote the idea (conferences, science funding). This would be open to everybody believing the project will have success.
3. Journals are formed by scientists working in similar fields. These early adopter journals are rewarded with certain amount of tokens, depending on perceived impact, to boost early participation in publication and review process.
4. As number of communities and members grow, the early adopter help from the foundation is depleted. Existing communities are free to choose the funding method and members as they wish.
5. As number of scientists and institutions publishing rises, the demand on the token becomes higher, and early investors, communities and foundation members are rewarded.
6. Foundation slowly reduces the number of tokens in the initial account to zero, funding the development, decentralisation and even science projects with a clause to support the ecosystem by publishing, until it is either dismantled if the whole system becomes completely decentralised, or continues to provide technical support as needed in much smaller scope.

In essence we propose 3 stages of progress: First, anybody supporting the idea or believing that it will raise its value invests money in the foundation. Then, foundation uses the acquired funds to develop the technology to allow decentralised publishing, and to promote the early communities to form and to publish. Once the adoption becomes global and decentralised, early investors and adopters can financially benefit from it, and the foundation would either continue to manage the open source infrastructure, without any influence on the publishing or science itself, or be disintegrated completely if there is no need for it.

6 Technology

In core, to publish an article one needs an unchangeable source of data (pdf's and supporting information), with unchangeable proof that specific person uploaded the article. To host a journal one needs set of pointers to the articles journal published (or endorsed). To provide the

financial layer one needs a method to transfer money between people interacting with journal. This is essentially what a publisher does today, and creating the same system creates a new publisher.

The trick is to be completely decentralised by doing this, as it would remove the central authority as the reference which chooses to profit from it. And the good thing is that the technology to do so exists today. By combining Ethereum project [7] and Ipfs project [8], one makes a publisher capable of managing the records and providing finance layer, where all the finance would be organised around the community hosting the journal, and more importantly would be completely transparent.

The person publishing the article would upload the data on the ipfs in form of the article itself and supporting meta-data which would help the searchability and traceability of the article. The meta-data does not need to be firmly defined at the beginning, but it could contain the title, authors, authors' public keys, abstract, keywords, and links to the citations, if they are using the same technology, to name a few. Once the links to the uploaded data are obtained, these are unchangeable.

Second step is to prove that the author (or authors) of the paper are the actual people holding the public key mentioned in meta-data. For this, a record in ethereum blockchain is formed, signed by one or more authors, which confirms the authorship of the data.

This, on its own is enough for open-access publishing. However, if one wants the peer-review stamp on the article, a second contract on ethereum blockchain is needed - Journal contract. This contract would contain links to the articles it wishes to host as journal publications. Also it would need to contain the list of public addresses belonging to the people: who own the journal (Editors in chief), to the ones allowed to choose the articles to be published (Editors), and the ones doing scientific scrutiny of the articles (Peer-reviewers). This journal contract could have the finance mechanism, where certain amount of tokens need to be transferred to the contract once someone attaches the published article or someone requests the published article. As discussed, it is proposed that in the beginning all the articles are open access, but there is a fee to publish.

All that is left is to organise the peer review, and this will not be discussed in this article. Possible solution is to keep the today's system where potential authors submit the papers to the editors for the review without making the public record of it. After the peer review is done, authors then make the records public, and editor approves this record to be published in the journal. Finally, the authors transfer certain amount of tokens to the journal contract with the article address and it becomes an official journal publication. Journals should have the freedom to organise the peer review independently, be it public or private.

Transparency of the finance is an added benefit, as all the funds contained in a journal contract would be publicly available for everyone to see, along with all the transactions going in or out from this journal. And although the journal owners or editors should in theory be allowed to utilise this tokens as they wish, they would need to do so with a strict public scrutiny. This would, as mentioned, introduce the financial aspect in the journal's reputation (the one thing painfully missing today). Although there should be rewards (pay) for Editors and Reviewers for their contribution, keeping publicly the list of them and writing in the contract how much funds they are allowed to withdraw, would keep the needed transparency along with the clear information how public funds are spent. Finally, as for any other publisher, there should be some sponsorship funds available, for conferences and similar events. The Editor in Chief would be directly responsible for this, keeping them too at the reasonable level by the consent

of the people involved.

7 Distributed Library - dislib.org

This paper is hosted on ipfs, noted as the public record and published in the example "Journal of Decentralised Publishing". Sources for the smart contracts can be found on the github, as well as the basic implementation of the html/js to browse the contracts for journal and article, if one has an ethereum web3 interface.

Initially plan is to form the library of distributed publishing www.dislib.org, listing all the early journals created by the scientific community, and a simple method to start own journals (which could also be hosted on IPFS). All of the code is plain HTML and Java Script, requiring no database or special security, as the content is published directly on peer to peer network and indexed in the Ethereum blockchain via smart contracts. Metamask [12] and Infura [13] are initially used as public nodes to interact with Ethereum or IPFS in a user-frendly manner.

- Journal address:
0xfeedbeeffeedbeeffeedbeeffeedbeeffeedbeef (TBA)
- Science token:
0xfeedbeeffeedbeeffeedbeeffeedbeeffeedbeef (TBA)
- Github page:
github.com/TBA

References

- [1] The STM Report: An overview of scientific and scholarly journal publishing, International Association of Scientific, Technical and Medical Publishers
http://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf.
- [2] Global box office revenue from 2016 to 2020 (in billion U.S. dollars)
<https://www.statista.com/statistics/259987/global-box-office-revenue/>.
- [3] IFPI: Global Music Report 2016: State of the Industry
<http://www.ifpi.org/downloads/GMR2016.pdf>.
- [4] Is the staggeringly profitable business of scientific publishing bad for science?
<https://www.theguardian.com/science/2017/jun/27/profitable-business-scientific-publishing-bad-for-science>.
- [5] US court grants Elsevier millions in damages from Sci-Hub
<https://www.nature.com/news/us-court-grants-elsevier-millions-in-damages-from-sci-hub-1.22196>.
- [6] Christian Janze. Design of a decentralized peer-to-peer reviewing and publishing market. *Proceedings of the 25th European Conference on Information Systems (ECIS)*, pages 1713–1725, 2017. http://aisel.aisnet.org/ecis2017_rp/110/.

- [7] Ethereum project
<https://www.ethereum.org/>.
- [8] IPFS is the Distributed Web
<https://ipfs.io/>.
- [9] Melanie Swan. *Blockchain: Blueprint for a new economy*. O'Reilly Media, Inc, 2015.
<http://w2.blockchain-tec.net/blockchain/blockchain-by-melanie-swan.pdf>.
- [10] Michael Spearpoint. A proposed currency system for academic peer review payments using the blockchain technology. *MDPI Publications*, 5(3):19, 2017.
<http://www.mdpi.com/2304-6775/5/3/19/pdf>.
- [11] Teresa Swist and Liam Magee. Academic publishing and its digital binds: Beyond the paywall towards ethical executions of code. *Culture Unbound: Journal of Current Cultural Research*, 9(3):240–259, 2018.
<http://www.cultureunbound.ep.liu.se/v9/a17/cu17v9a17.pdf>.
- [12] MetaMask is a bridge that allows you to visit the distributed web of tomorrow in your browser today.
<https://metamask.io/>.
- [13] Infura is a secure, reliable, and scalable access to Ethereum and IPFS.
<https://infura.io/>.