Disclaimer:

All the ideas and work in this paper are my own and do not have any relation with my profession affiliations.

This is not a white paper for an ICO.

Open Science Organization

A decentralized autonomous organization for a better scientific ecosystem

Gajendra Jung Katuwal
gjk0112358@gmail.com
https://github.com/open-science-org

Aug 15, 2017

Abstract

Modern day science has been a collective effort of different researchers and research groups. The collective intellectual power gathered by the scientific ecosystem is akin to the group intelligence (the "wisdom of crowds") where a wise crowd is generally more intelligent compared to individual members in the group. The additional gain in intelligence is primarily dependent on the efficient flow of value (capital, information, etc.) among the members. The scientific community is still dependent on the rudimentary system for value flow, which is highly centralized, opaque, and redundant. Thereby it suffers from several problems: funding agencies are controlled by a small number of people, majority of scientists' time is consumed by grant application writing with small success rates, a slow publication process, institutional biases, underpaid researchers, irreproducible publications, high subscription fees for journals, and a general focus on quantity of scientific publications over quality. These problems can be solved and the overall efficiency of the scientific community can be improved by creating an open and decentralized scientific ecosystem based on blockchain technology (or its variants). In this publication, a simple Distributed Autonomous Organization (DAO) based model of open and decentralized scientific community called Open Science Organization (OSO) is presented. This publication attempts to demonstrate how blockchain technology can be utilized to create an efficient scientific ecosystem.

Acknowledgements

(in alphabetical order)

- **Abinash Koirala**, Software Engineer, Move Inc.
- **Bikrant Gautam**, Web developer, ePromos.com
- **Emin Gün Sirer**, Associate Professor, Cornell University
- **Joseph Netti**, President, Rochester Institute of Technology (RIT) Blockchain Lab
- **Robert Chen**, MD/PhD Candidate, Georgia Tech & Emory University

1 Collective science and its efficiency

In the past few centuries, science has transitioned from largely individual pursuits to collective pursuits. This transition is mainly fueled by the professionalization of science, improved communication between the scientists, and increase in the complexity of the scientific problems (Fernandez, 1998). This has led to the *present day scientific community* – similar to a group of intelligent entities interacting with each other to achieve a *collective intelligence* that is greater than the intelligence of the individual members. The collective intelligence of the group is dependent on both the contribution of individual intelligence and the system unifying them. In other words, the collective success of our scientific pursuits, in addition to the quality of individual scientists, is highly dependent on the efficiency of the ecosystem within which they interact. An ideal ecosystem would be the one which achieves the following two objectives:

1. **The Efficient flow of value** within the community to optimize the direction and magnitude of the scientific progress to align with the overall organic growth of the human race and its environment. Here, the term 'value' means anything that is useful for a successful scientific endeavor. Value can be in the form of funding, data, information, computational resources, etc.
2. **A sense of fulfilment** among the members of the scientific community.

This publication is mainly related to objective 1, the efficient flow of value or the creation of an efficient scientific ecosystem.

2 Problems with current scientific ecosystem

The current scientific ecosystem is not well-suited for the efficient flow of value and is riddled with problems such as centralized funding, cut-throat competition for funding, a closed and slow review process of publications, irreproducibility of scientific findings, etc. The problems associated with the current ecosystem are discussed briefly in the next section.

2.1 Funding

2.1.1 Funding biases and research directions

In the current scientific ecosystem, there are three sources of academic research funding: government, private companies, and philanthropists. Government funding is fundamentally susceptible to the politics and might be influenced by the biases and aspirations of the politicians who may not have a strong grasp on the current science. For example, the recently proposed budget cuts to the Environmental Protection Agency (EPA).

In addition, the authority to control the funding directions and allocation in governmental funding agencies are concentrated in a few people. Most of the funding agencies do receive reviews from the external reviewers. However, the selection of the external reviewers itself and the final decision made are highly susceptible to the biases and affiliations of the members of the decision group.

The studies funded by private funding suffer from the 'sponsorship bias' — a form of experimenter's bias where the results of studies tend to support the interests of the funding sponsors. The effect of sponsorship bias has been more prevalent in pharmaceutical drug studies, tobacco research, and chemical toxicity studies (Krimsky, 2013). Books by Ben Goldacre entitled "*Bad science: Quacks, hacks, and big pharma flacks*" (Goldacre, 2010) and *Bad pharma: how drug companies mislead doctors and harm patients* (Goldacre, 2014) are good reads to get better insights on how pharmaceutical companies can influence the results by manipulating the clinical trials.

2.1.2 Imbalance in demand and supply of funding

The inflation-adjusted academic funding as a share of GDP has been decreasing since the last decade, while the current scientific pipeline is producing an ever increasing number of PhDs looking for research grants. This imbalance between the demand and supply has created a hyper-competitive atmosphere which has been detrimental to the organic growth of scientific research. For example, NIH funding as a share of GDP has been falling since the last decade. The buying power of the funding is further diluted with the ever rising cost of biomedical research. On the other hand, the current scientific pipeline is producing far more biomedical scientists, making it very difficult to win grants; the success rate for competing RPG applications in 2016 was 19.1% (https://nexus.od.nih.gov/all/2017/02/03/fy2016-by-the-numbers/).

This hyper-competitive atmosphere resulting from the imbalance in demand and supply of funding has affected science negatively in several ways. First, the atmosphere promotes short-term gain vs long-term gain in science; it suppresses the creativity, original thinking, and risk taking and promotes safe research. In addition, it promotes quantity over quality of scientific publications and encourages the scientists to take the short-cuts and exaggerate their results due to the pressure of publishing. Moreover, scientists are spending more time writing the grants instead of conducting the scientific research due to the overly competitive funding scenario.

2.2 Review

The current academic review process is closed, incredibly slow, and suffers from several other problems such as institutional bias and survivor bias.

2.2.1 Closed review process

The current review process is closed in a sense that the rationale behind the decision to accept or reject a publication is not disclosed to the public. The reviewers and editors will have subjective biases regarding different scientists, research labs, institutions (e.g. favoring their alma maters), scientific methodologies, scientific principles, scientific findings (e.g. favoring positive replications promoting survivor bias). In addition, due to the pressure to publish (e.g., "publish or perish") created by scarce funding, the assigned reviewers often lack the time to thoroughly review the manuscripts and may even hand over the tasks to their students who may lack the expertise. Furthermore, the funding company might be secretly influencing the review committee to promote the acceptance of the publication serving its interest. For example, Richard Smith, former editor of the British Medical Journal, said —

"All journals are bought — or at least cleverly used — by the pharmaceutical industry"

as reported in The American Scholar (https://theamericanscholar.org/flacking-for-big-pharma/#.WV_uWhPyvBI). The only fool-proof way to minimize these subjective biases is to make the entire review process open so that the scientific community can act as a wise crowd to judge the quality of a publication.

2.2.2 Slow review process

The median review time — the time between submission and acceptance of a publication — is approximately 100 days and the median online publication time — the time between acceptance and online debut of a publication — is approximately 25 days (http://blog.dhimmel.com/history-of-delays/). These delays are very long and they slow down the pace of scientific research.

2.2.3 Unable to address the low replicability of scientific findings

In addition, under the current review model, there is no way to promote the replicability of the studies. Replicability of many scientific studies has been alarmingly low, thus questioning their credibility. A recent study by Baker (2016) reports that 70% of researchers have tried and failed to reproduce the experiments of other scientists, and more than 50% have failed to reproduce their own experiments.

2.3 Closed and expensive dissemination

The total scientific, technical, and medical (STM) information publishing market was estimated to be $25 billion and the annual revenues generated from English language STM journal publishing alone was estimated to be $10 billion in 2013 (Ware & Mabe, 2015). The cost of publishing can range from $2000 to $10000 per publication (Ware & Mabe, 2015). In addition, researchers and universities pay a heavy amount to subscribe to journals for accessing their publications. For example, Harvard University paid more than $16 million as subscription fees in 2012 alone (https://academia.stackexchange.com/questions/29923/reference-for-annual-journal-subscription-costs-paid-per-university) (Bergstrom et al., 2014; Phan et al., 2011). In addition, when a researcher's university does not have a subscription to a journal or if he/she is not associated with a university, he/she has to pay the subscription fee personally which is higher than bundle rates provided to universities.

The most absurd thing about the traditional publishing business model is that (https://www.nature.com/news/open-access-the-true-cost-of-science-publishing-1.12676):

the peer-review— the most critical process behind the value creation of the current scientific publishing industry—is done for free by the scientific community.

3 Open Science Organization (OSO)

The open science organization (OSO) model described below has been kept as simple as possible because the main aim of this publication is to show how a better scientific ecosystem can be created utilizing blockchain technology. I am sure that much better models which are algorithmically and socio-economically sound, will be proposed by other members of the scientific community in the near future.

OSO will be a decentralized and autonomous organization (DAO) (Deegan, 2014) based on blockchain technology (Dwork & Naor, 1993; Nakamoto, 2008; Wood, 2014) with native tokens called OSO. It will be non-profit and flexible (in a sense that its form can be changed as per the consensus of the organization members). There will be a cost for each transaction performed in the OSO network. The revenue collected from the cost of a transaction will be shared between the OSO reserve fund and the nodes participating in the transaction. The OSO reserve fund is the main account of OSO network owned by all the entities in the network.

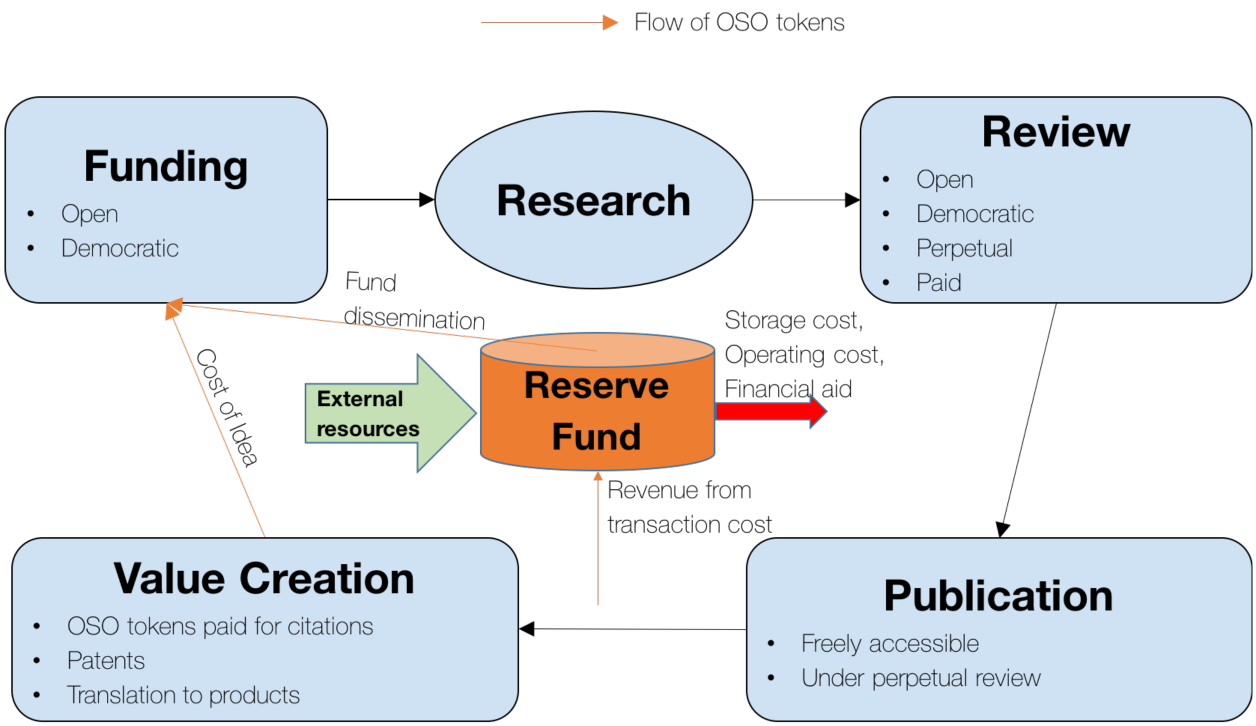

Figure 3.1 Open Science Organization (OSO) — a non-profit flexible DAO in which funding, review, and publication processes are open and democratic. The value created by scientific research will flow back to the funding step, thus creating a perpetual scientific ecosystem.

The three key characteristics of OSO ecosystem are 1.) an open democratic funding process, 2.) an open democratic perpetual review process, and 3.) the flow of value (in OSO tokens) created by the scientific results (publications, products, patents, etc.) back to the funding, thus creating *a perpetual system for scientific research*. In OSO ecosystem, all steps/procedures will be algorithmic or will be based on community voting. Even certain aspects of algorithmic steps (e.g. algorithms used for decision making) are subjected to change based upon voting. Any individual or organization who holds OSO tokens will be considered an entity (or member) in the OSO ecosystem. Each entity will have a voting power proportional to their expertise (e-value) in the corresponding scientific domain and the amount of value (OSO tokens) they are willing to invest/stake.

Let me first describe the following key concepts: Interplanetary Idea System (IPIS), expertise value, expertise function, vote value, and vote function.

3.1 Interplanetary Idea System (IPIS)

IPIS will be a protocol/model by virtue of which each unique idea can be identified and its evolution over the time can be tracked. In addition, it will give us a framework to evaluate the relationship between two different ideas. One way of modeling IPIS would be by a directed acyclic graph (DAG) that grows over time. The idea behind this model of IPIS is motivated by the Interplanetary File System (IPFS) (Baumgart & Mies, 2007) and the IOTA tangle (Popov, 2016). Each node of the DAG will be an IPIS object representing a unique idea. An IPIS object will contain all the information required to represent a unique idea. If there is a publication or data (stored as an IPFS object) related to an idea, its IPIS object will contain the cryptographic hash of the IPFS object. A very simple model of IPIS is presented in Figure 3.2.

IPIS will allow us to track the evolution of ideas over time as well as quantify the similarity of two ideas. The similarity score can be a hybrid of the market-driven score (decided by the scientific community via voting) and the algorithmic score generated from the IPIS protocol. IPIS will facilitate the efficient sharing of the ideas and prevent the community from reinventing wheels, thereby accelerating the pace of the scientific research. In addition, a good visualization of IPIS objects will help the scientific community to identify the pain points of scientific research. Similarly, it can be used to find the optimal distribution of available resources (e.g. fund allocation) required in order to maximize scientific throughput. Loosely speaking, IPIS will store the collective thought process of the whole scientific community.

TO DO:
Develop a formal mathematical model and network protocol for IPIS
https://github.com/open-science-org/IPIS

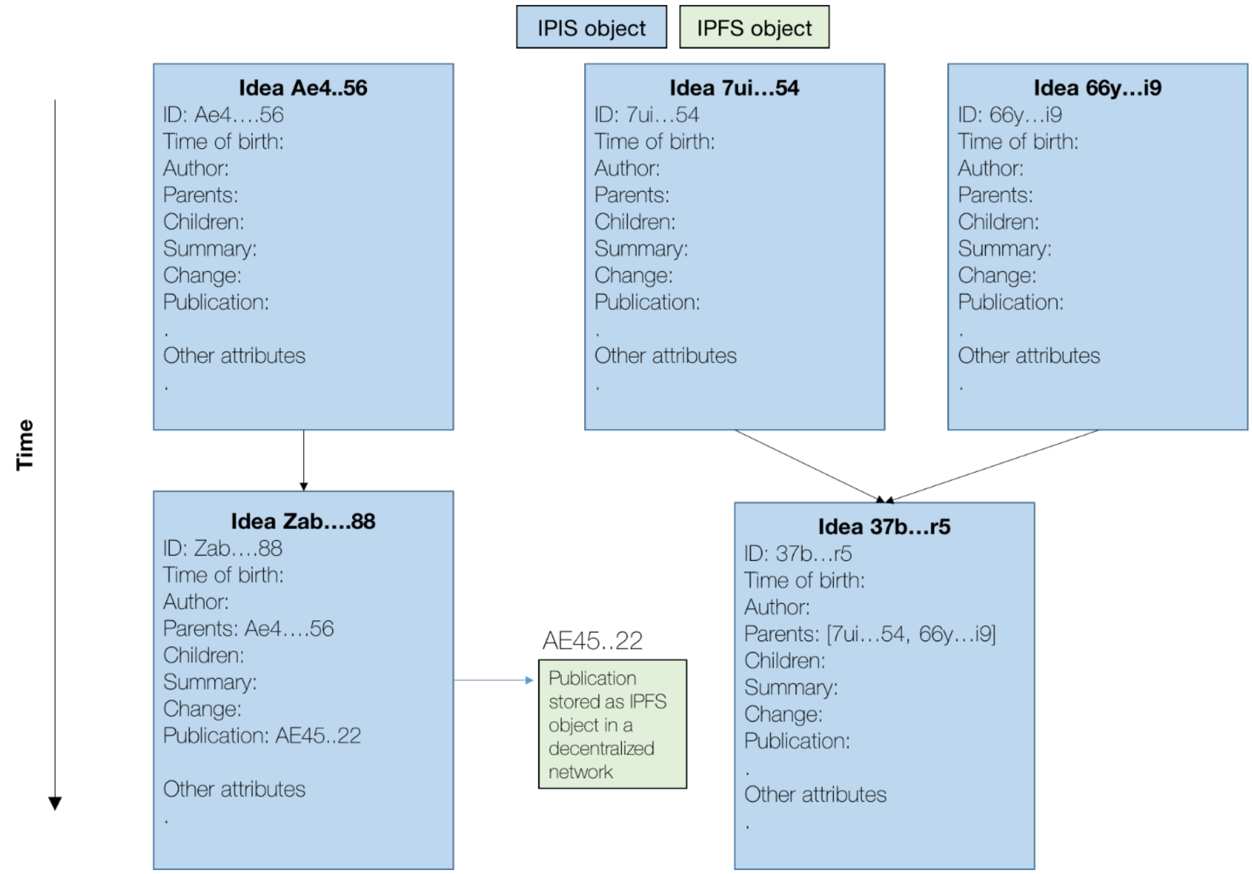

Figure 3.2 A naïve model of Interplanetary Idea System (IPIS). Idea ID is similar to the digital object identifier (DOI) of digital publications. Idea *Zab…88* is derived from idea *Ae4..56* and the strength of connection between them is determined by both IPIS protocol and market. A publication related to the idea *Zab…88*, whose cryptographic hash is *AE45..22*, is stored as an IPIS object in OSO network. If required, the storage can be outsourced to any proof-of-storage peer-to-peer network.

3.2 Quantification of expertise of an entity in a research domain

The level of expertise or $e - value$ of an entity (general individual, researcher, research group, organization, etc.) in a research domain D will be computed using e-function. The form of $e - function$ will be decided by voting, and maybe domain dependent (or a subspace of idea space defined by IPIS). The $e - value$ is defined as:

$$e - value = e - function(V, PEER)$$

Where, V is the aggregated value created by the entity related to the domain D and is defined as:

$$V = \sum_i s_i \, v_i$$

where v_i is the total value created by a scientific publication or an idea and s_i is similarity of the i[th] publication or idea with the domain D. The similarity function will be decided by voting and maybe different for different research domains.

$PEER$ is the aggregated peer rating of the entity given by other entities for the domain D. It is defined as:

$$PEER = \sum_i e_i * peer_i$$

$peer_i$ is the rating given by the ith peer and e_i is the e-value of the ith peer.

3.3 What does voting mean in the OSO model?

In OSO model, a voting by an entity for an idea/project is equivalent to how many OSO tokens the entity is willing to stake for the idea. The percentage of OSO tokens that is actually transferred is different for different scenarios such as fund allocation, fund matching, reviewing, etc. For example, during funding process, voting for a project/idea (if it is a smart contract) is equivalent to sending all the staked OSO tokens to the smart contract. That means, you are supporting the project as well as investing in it. Whereas in review/feedback stage, OSO tokens might be used only for signaling.

The value of a vote ($v - value$) from an entity will be calculated according to $v - function$ as below:

$$v - value \ = \ v - function(nOSO, e - value)$$

Where *nOSO* is the number of OSO tokens staked for the voted idea/project, $e - value$ is the expertise value of the voter related to the project. The $e - value$ will be positively correlated with the $v - value$ to ensure that the more informed votes have more weights. The $v - function$ will be task, domain, and time dependent and its form will be decided by voting itself. It will be designed to find the optimal trade-off between economic and intellectual forces.

3.4 Open & Democratic Funding

Every procedure related to the funding will be open to the public in OSO network. The funding mechanism will be democratic in nature and will be decided by voting. In OSO funding mechanism, there will be three different voting stages: 1) **Direction vote:** decides the funding distribution to different research domains and projects, 2) **Popularity vote:** decides the popularity of the submitted proposals, 3) **Matching vote:** decides the matching of the funded projects to the fund applicants. The different voting stages will have different v-functions.

3.4.1 Funding priorities/direction (Direction vote)

The decision making for fund allocation to different research domains and projects will be conducted using a hybrid voting model which will be the mix of proportional voting and first-past-the-post (FPTP) voting models. This hybrid approach ensures a healthy trade-off between “long-shot” research projects aiming for fundamental breakthroughs and hot research projects fueled by current market demand and popular interests.

3.4.2 Proposals submission (Popularity vote)

Research areas and projects can be proposed by any entity inside OSO. An entity or a group of entities can propose a project either requesting the fund to conduct it or only to share the idea so that other capable entities can conduct it. In order to control spams and ensure quality proposals, OSO tokens will have to be staked to propose an idea/project. On parallel to the proposals submission, the popularity voting will be ongoing.

3.4.3 Matching funding with ideal entities (Matching vote)

Voting will be done for each (project, proposal, entity) tuple and the entity that gets the highest number of votes will receive the allocated funding.

A simple model of a democratic funding mechanism can look like:

1) Direction vote: decision on funding priorities
    a) Fund Splitting (80% proportional, 20% FPTP) - this threshold itself can be decided by voting.
    b) Proposals submission
    c) Popularity vote (proportional)
        i) Research domains and their relative sizes will be decided by voting
            (1) This proportional voting may continue recursively within each research domain until it terminates to individual ideas/projects. For example, it may start with the research domain of biology and then branch into genetics, cellular biology, etc. and then genetics may again branch into genomics, epi-genetics, etc.
            (2) The termination of this recursion can be decided by voting itself. This will allow to incorporate the research projects sharing more than one research areas.
        ii) Popularity + Matching vote under each research domain (FPTP)
            (1) selection of top projects within a research domain. After the recursion terminates to individual projects, the top projects to be funded will be decided by FPTP voting
2) Popularity + Matching vote (FPTP)
    a) 20% of the total funding goes to the overall top projects

3.5 Open & Democratic Review Process

The review process will be open to the public. It will consist of two stages: initial review and perpetual review.

3.5.1 Initial review

At first, the quality of a publication is estimated by combining the decision from a review algorithm and a set of reviewers. The quality of publication can be multidimensional and its form can be decided by voting. One important feature of the OSO model is — *paid and quality peer review model* — in which the reviewers will be rewarded with OSO tokens for their work. This will serve two purposes. First, it will ensure quality reviews by providing economic incentives to the reviewers. Second, it will discourage spam publications. The acceptance/rejection of a publication in a journal/channel will be decided by the majority vote of the reviewers. There can be a separate journal/channel (or a hierarchical spectrum of publication channels) to publish rejected publications explaining the reasons for the rejection. This will provide examples for new authors to learn how to write good publications and reduce the positive-result bias (bias when only interesting things at a particular time appear more).

In a different model, the concept of rejection/acceptance can be completely removed and the publications can be promoted/demoted to different levels (within a channel or a research domain) solely depending upon their quality (market-driven). In addition, a publication can be shared among multiple channels.

The quality of a publication can be quantified by following main metrics (can be domain and journal/channel dependent) and will be decided by voting:

1) Reproducibility:
    a) How **reproducible** is their work?
    b) Have they **open sourced** their code?
    c) Have they **clearly described every step** in their methodology?
2) Readability:
    a) How easy is it to read and ascertain key concepts from a publication? The hope is that this will decrease publications with very convoluted writings. This will be decided by readers' feedback.
3) **Number of citations and quality of the publications citing them**. Citing publications will cost some tokens (cost of idea) and those tokens will be transferred to the authors of the cited publications. So the basic idea is that publishing your publication is like rolling out a product which has initial cost but will generate revenue and the revenue can be used for research to publish other good quality publications.
4) Importance of the publication – judged in two ways:
    a) **Incremental change vs. significant leap**? Significant leaps are more difficult to make and therefore should indicate higher quality.
    b) **Economic impact of the publication**? Economic impact refers to the extent of value a publication adds to the broader community. These metrics are subjective so we have to use multiple sources of information such as readers' feedback, feedback from the top authors of the respective field i.e. with high e-values, feedback from the industry leaders specially to quantify the economic importance of the publication, etc.

3.5.2 Perpetual review process

Starting with the initial quality/rating achieved from the initial review process, the publication will go through a *perpetual review process*. After the initial review, the quality of a publication will be adjusted according to the customers (readers/researchers) feedback and changes in the literature and technological advances. You can think of the publication as a movie or a book where the critics give the initial review and the audience/reader get a rough initial estimate of its quality. However, the rating/quality of the publication might significantly change later depending on its impact and utility, as judged by the readers. For example, a mathematical proof at one time point might look boring and not useful, and the publication might not get that much interest, but it might turn out to be a crucial piece for a significant technology after 10 years. So even if the publication might have less than 5 ratings (assuming the quality scale is 0-5), it will be a 5-star publication after 10 years. On other hand, the perpetual review process also acts as a regularizer to insure reproducibility. For example, a researcher might publish a stellar publication with ground-breaking results and the publication receives 5 stars in the initial review stage. But if nobody can reproduce its results later, it is expected that its quality would decline.

3.6 Cost of publication

There are three different costs of publication in OSO ecosystem. For the entities that can demonstrate the financial need, the OSO reserve fund will be utilized to cover the publication cost.

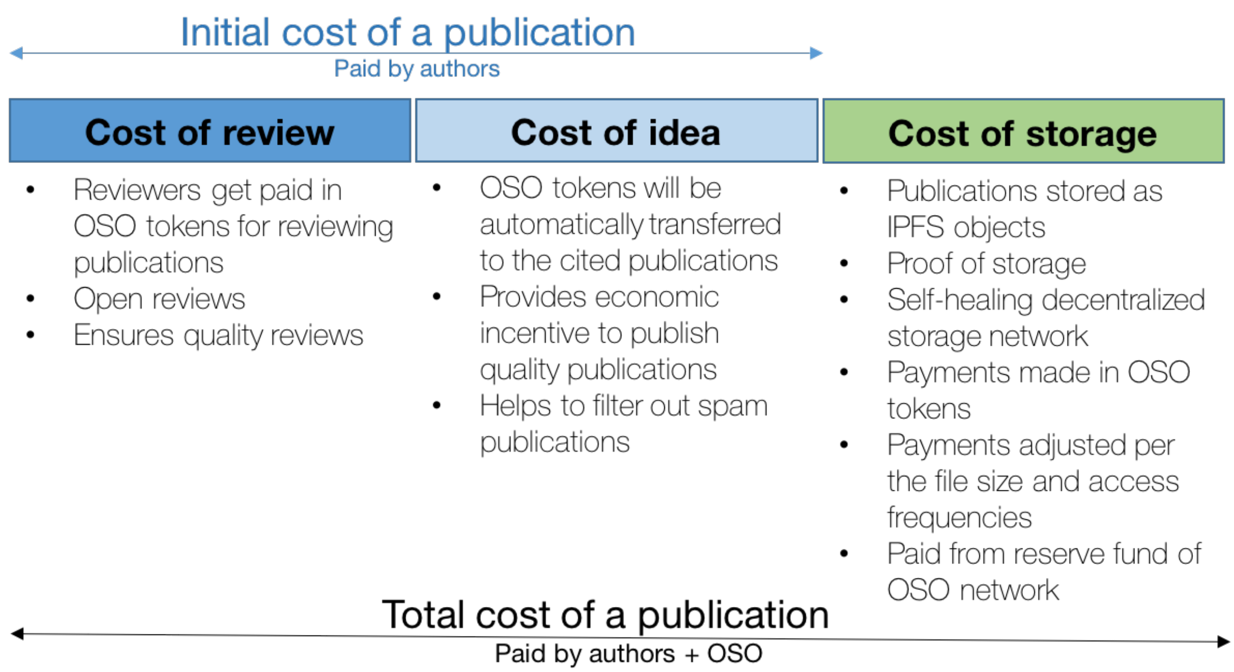

Figure 3.3 Cost of publication

3.6.1 Cost of review

Reviewers will get paid in OSO tokens for their work. See section 3.5.1 for detail.

3.6.2 Cost of idea

Whenever a publication cites previous publications, it is in essence borrowing the idea form them. In OSO network, the new publication has to pay OSO tokens to the cited publications and the cost related to it is called the *"cost of idea"*. Here, the belief is that the cost of idea concept will give an economic incentive to academics to continue producing high quality work in the ecosystem. The number of OSO tokens paid to a cited publication will be based on the quality (determined by the perpetual review process) of the publication. Thus, more OSO tokens will be transferred to the authors who publish high quality publications and the received tokens can be utilized to do high quality research which will then produce more high quality publications, hence creating a *strong positive cycle for high quality open research*. This will solve the problems (extremely competitive, 50% of the time wasted in grant writing) with the present funding scenario to some extent. In addition, it will encourage young students and researchers to produce high quality publications.

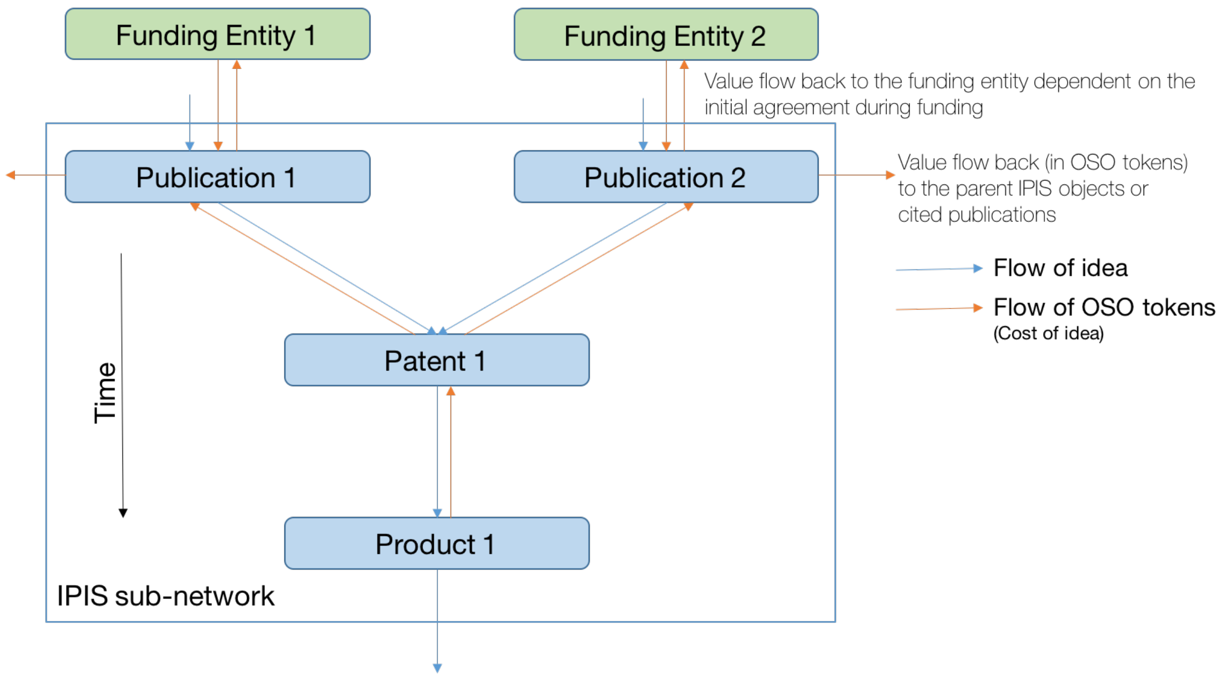

Figure 3.4 Flow of value in OSO network. Product 1 is based on patent 1 which in turn is based on publications 1 & 2, which were funded by funding entities 1 & 2 respectively.

The value created by an IPIS object flows upstream until it reaches the funding origin. Here, when product 1 creates an economic value, a portion of it is transferred to patent 1 and so on. The portion of funding that flows back from an IPIS object to another can depend on several factors such as nature of IPIS objects, nature and strength of relationship between them, any pre-agreement, etc. The model for value flow will be domain dependent and will be decided by voting. Let us assume a simple model where an IPIS object has to give 10% of the total value created to all its parents. Here, if product 1 generates \$100, patent 1 will receive 10% of \$100 = \$10 and publication 1 & 2 each will get ½*10% of \$10 = 50 cents (assuming strength of connection is same i.e. equal cost of idea). Likewise, funding entities 1 & 2 will get 10% of 50 cents = 5 cents from publication 1 & 2 respectively.

3.6.3 Cost of storage

Publications will be stored as IPFS objects in a self-healing decentralized peer-to-peer network sustained by proof-of-storage (Vorick & Champine, 2014; Baumgart & Mies, 2007). The protocol describing the storage network hosting the publications will be a subset of IPIS protocol which itself will be a subset of the OSO protocol. The storage nodes will be paid in OSO tokens from the OSO reserve fund and the cost of storage of an IPFS object will be dependent on its size and access rate. The storage can be outsourced to any proof-of-storage peer-to-peer network if required.

4 Conclusion

The collective output of the scientific community is fundamentally dependent on the environment in which the members interact with each other. Current scientific ecosystem is closed, biased, slow, and inefficient in general. An open, fair, fast, and efficient scientific ecosystem can be created using blockchain technology (or its variants). In this publication, open science organization (OSO) — a flexible decentralized autonomous organization (DAO) — was proposed. All processes in the OSO ecosystem will be open and democratic.

5 Other Ideas

- Connecting the OSO network with proof-of-work blockchains so that the scientific computation can be outsourced to these blockchains. This will transform the wasteful proof-of-work computations to useful computations (Halford, 2014).

6 Challenges

- Incorporating diverse research entities
  - cross country collaboration
  - cross organizations collaboration
- Incorporating research entities form private companies and security agencies
- Protecting business interests while making everything transparent

7 Drawbacks of OSO

- 1
- 2

8 References

Baker, M. (2016). Is there a reproducibility crisis? *Nature*. 533. p.pp. 452–454.

Baumgart, I. & Mies, S. (2007). IPFS - Content Addressed, Versioned, P2P File System. *Proceedings of the International Conference on Parallel and Distributed Systems - ICPADS*. 2 (Draft 3).

Bergstrom, T.C., Courant, P.N., McAfee, R.P. & Williams, M.A. (2014). Evaluating big deal journal bundles. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (26). p.pp. 9425–30.

Deegan, P. (2014). *Chapter 14—The Relational Matrix: The Free and Emergent Organizations of Digital Groups and Identities*.

Dwork, C. & Naor, M. (1993). Pricing via Processing or Combatting Junk Mail. In: E. F. Brickell (ed.). *Advances in Cryptology --- CRYPTO' 92: 12th Annual International Cryptology Conference Santa Barbara, California, USA August 16--20, 1992 Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 139–147.

Fernandez, J.A. (1998). The transition from an individual science to a collective one: The case of astronomy. *Scientometrics*. 42 (1). p.pp. 61–74.

Goldacre, B. (2014). *Bad pharma: how drug companies mislead doctors and harm patients*. Macmillan.

Goldacre, B. (2010). *Bad science: Quacks, hacks, and big pharma flacks*. McClelland & Stewart.

Halford, R. (2014). *Gridcoin: Crypto-Currency using Berkeley Open Infrastructure Network Computing Grid as a Proof Of Work*.

Krimsky, S. (2013). Do Financial Conflicts of Interest Bias Research? *Science, Technology, & Human Values*. 38 (4). p.pp. 566–587.

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. *Www.Bitcoin.Org*. p.p. 9.

Phan, T., Hardesty, L.C., Hug, J. & Sheckells, C.L. (2011). *Documentation for the Academic Libraries Survey (ALS) Public Use Data File*.

Popov, S. (2016). *The Tangle*.

Vorick, D. & Champine, L. (2014). *Sia: Simple Decentralized Storage*.

Ware, M. & Mabe, M. (2015). The STM Report: An overview of scientific and scholarly journal publishing. *The STM Report*.

Wood, G. (2014). Ethereum: a secure decentralised generalised transaction ledger. *Ethereum Project Yellow Paper*. p.pp. 1–32.