Information Services & Use 38 (2018) 95–98
DOI 10.3233/ISU-180003
IOS Press

The blockchain and its potential for science and academic publishing

Joris van Rossum*
Director, Special Projects, Digital Science, 1 Mark Square, London EC2A 4EG, UK

Abstract. Blockchain allows for decentralized, self-regulating data, ultimately creating a shared infrastructure where transactions are saved and stored. Scientific information in its essence is a large, dynamic body of information and data that is collaboratively created, altered, used and shared. It lends itself well to the blockchain technology because that technology has the potential to solve challenges around peer review, irreproducibility, and metrics. Other applications of blockchain technology such as cryptocurrencies and digital rights management systems also have a potential relevance for academic publishing.

Keywords: Blockchain, cryptocurrency, bitcoin, scholarly communication, publishing

1. Current challenges in scholarly communication

When researchers want to communicate their findings, they usually use different - and to a large extent disconnected - systems in their research workflow. For example, spreadsheets or lab software are used to capture the results of an experiment. When results are collected, an article is written using a local writing application or a cloud-based collaborative writing tool. The article is then submitted to a publisher through a submission system. After review and acceptance the manuscript is converted to PDF and HTML, and then is hosted on a publisher platform, from which it can be downloaded. Access to this publisher platform is often facilitated by librarians. Citations are collected in citation databases that are distributed through librarians or via freely-accessible databases.

This workflow is problematic for several reasons. First of all, research only becomes accessible at the point of publication. Everything that takes place prior to this - such as collecting and analyzing data, peer review, etc. - is not transparent. This lack of transparency leads to problems around reproducibility, i.e., the inability of researchers to reproduce experiments in order to validate the conclusions made in research papers, a practice that is a cornerstone of the scientific method. In a 2016 poll on *Nature.com*, two-thirds of respondents indicated that current level of reproducibility is a major problem, with 52% saying that there is a "significant crisis" in the ability of researchers to validate prior work.

Peer review - the process whereby research papers are evaluated by researchers working in the same field before a paper is published in a journal - is in a similar situation. The process remains opaque. There is also a lack of visibility and recognition for reviewers, with their review work remaining largely unnoticed.

*E-mail: j.vanrossum@digital-science.com.

Scientific results are primarily published in academic journals, which have a strong tendency to publish positive and novel results. Moreover, scientists themselves are more inclined to report on their successful outcomes rather than on failed experiments. A lot of research that did not lead to positive results, therefore, remains unpublished, and unknown. Moreover, as the productivity of researchers is predominantly measured in terms of journal article output, scientific efforts leading to negative results and non-research activities (e.g. reviewing articles and grant applications, sitting on scientific committees, or even micro-contributions such as participation in brainstorms, informal comments) are undervalued.

Additionally, there are challenges in research and scholarly communication that have to do with commercial interests. Research is essentially a non-commercial activity, but ironically the business of scholarly communication is one of the most lucrative industries in the world, dominated by a few large publishing giants. This causes several issues. High prices charged by commercial publishers for subscriptions challenges library budgets, and implies that not all content is made accessible to scientists at institutions. Partly as a reaction to the problems associated with the subscription model, Open Access, the model whereby payment is shifted from the reader or library to the author granting universal access to the article, has been introduced. But several decades after its introduction, still only a minority of articles are open access. Moreover, open access has introduced its own set of problems, such as the incentive of publishers to accept articles potentially leading to less rigorous quality norms, and the appearance of so-called "predatory" publishers, i.e., exploitative publishers that charge publication fees to authors without providing the editorial and publishing services that are associated with legitimate journals.

2. How could blockchain technology help scholarly communication?

2.1. Cryptocurrencies

A science blockchain could accompany the introduction of a cryptocurrency, which would add an economic layer to the blockchain. This 'bitcoin for science' could be used to make micropayments to publishers for consuming their content. It could also introduce rewards for scientific activities, such as peer review, statistical support, exchange of lab equipment, outsourcing specific research, or the hosting of data. Eventually, initial coin offerings, a form of crowdfunding using cryptocurrencies, could be used to fund entire research projects. In this way, a crypto economy could evolve in science reflecting the value merits of a number of activities.

2.2. Disseminating content & digital rights management

The role of blockchain has been researched predominantly in general (i.e. non-academic) publishing, where the move to online has led to a shift in revenue allocation from content creators and publishing companies to hosting companies, social media giants, and advertising intermediates. To some extent, this shift is caused by an inherent characteristic of the World Wide Web, namely the use of hyperlinks. Hyperlinks are one-way pointers to content, but they do not point back to the users that click on them. Hence, there is no mechanism for allowing small automatic payments (micropayments) for usage. Given this, the only choice for publishers is to open up content and base a business model on advertising, or impose unfriendly paywalls with expensive credit card payments. Micropayments could also form an alternative for the dominant business models in academic publishing (subscriptions and open access), each coming with their own challenges.

An interesting potential dimension of the blockchain is digital rights management. The coupling of usage to micropayments already makes rights management more straightforward, but digital rights can also relate to more complex aspects like reuse, permissions, and royalties that are currently intermediated through large institutions and complex products. The combination of a central database with smart contracts could bring huge advantages. Through the blockchain, ownership of content is automatically established, and the use of content and the payment of royalties are executed through smart contracts in which the rights are stored.

An additional advantage of content being disseminated via the blockchain is that usage can be accurately counted. Currently, content is downloaded and shared via different platforms (e.g. publisher platforms, ResearchGate, PubMed Central, etc.), which makes the tracking of usage difficult. This is problematic not only for publishers, but also for researchers and institutions for whom readership and usage is an important metric. A blockchain would make usage counting and reporting both accurate and simple at the same time.

2.3. Storing research data

The decentralized nature of the blockchain opens the way to create a datastore in which research activities from the entire research ecosystem can be collectively stored. Although creating this datastore is technically possible with current database technologies, the inherent need for a central gatekeeper and owner of such a datastore makes its realization highly unlikely. Such a decentralized data store would mean that whenever a researcher, for example, uploads data, performs statistical analysis, writes and submits an article or reviews a manuscript, this would be automatically tracked and recorded, making research significantly more reproducible. The risk of fraud would also be reduced. Moreover, it would also make it significantly easier to collect reliable and complete data on the performance of researchers, research, and universities, which would allow for more sophisticated as well as more reliable metrics to be built on top of that. Additionally, it will allow metrics to be based on activities that are currently not well not well recognized (e.g., peer review).

3. Will scholarly communication take place on the blockchain?

In light of its obvious advantages over the current ecosystem, it is tempting to predict that scholarly communication and other research activities will eventually take place on the blockchain. Its potential impact touches many, if not all, challenges around scholarly communication, especially those to do with trust, reproducibility, transparency, and access. However, there are also reasons to be cautious.

Science has evolved over hundreds of years, and with its history comes a significant amount of legacy in technology, systems, organization, as well culture. This legacy makes any change difficult, despite the challenges associated with the current system. The likelihood and success of a blockchain for scholarly communication would also depend on its level of implementation. For example, information stored on the blockchain could be restricted to traditional researcher roles, publications, and use of content (e.g. authorship of scientific articles, usage and citations). But it could also reward unconventional roles and affect wider aspects of the research workflow, including peer review, publication of datasets, hypotheses, etc., which would increase the level of complexity. In connection to this, an essential question is whether blockchain technology will be embraced by existing players such as the major science publishers or whether it will be successfully introduced by external parties, such as Scienceroot and Pluto. Scienceroot is an open access blockchain-based scientific ecosystem that combines all of

the functionalities required during the scientific discovery process – from funding through research to publishing (see: https://www.scienceroot.com/) and Pluto is a decentralized scholarly communication platform powered by etherium blockchain (see: https://pluto.network/). Both of these players have the potential to disrupt the current scholarly communication ecosystem and only the future will tell. The use of blockchain technology is growing and innovative applications bear watching.

About the Author

Dr. Joris van Rossum has been working for twenty years in the academic publishing industry. In 2015 he co-founded Peerwith, a marketplace for researcher services. Before that, Joris was Director of Publishing Innovation at Elsevier. Earlier roles include Head of Scirus, a vertical search engine for scientific information, and Senior Product Manager for Scopus. In his current role, Joris investigates the potential of blockchain technology for research and scholarly communication for Digital Science. Email: j.vanrossum@digital-science.com.