Using Blockchain Technology in Scientometrics

Dmytro Lukianov[1], Kateryna Kolesnikova[2], Alen Mussurmanov[2], Rostyslav Lisnevskyi[2] and Valeryi Lisnevskyi[3]

[1] GIGACloud, Kyiv, Ukraine
[2] International Information Technology University, Manas St. 34/1, Almaty, 050040, Kazakhstan
[3] The National Defence University of Ukraine, Povitroflotskyi Avenue, 28, Kyiv, 03049 Ukraine

Abstract

The paper proposes to consider the possibility of using Blockchain technology to create a new generation scientometric index that can take into account both the direct and indirect influence of the authors of scientific works on the development of scientific thought as a whole. An example of a possible model for implementing the calculation of the index is given, taking into account both the logic of direct citation of the author's works and the analysis of the entire chain of past and future publications using the author's ideas. The applicability of this method for assessing the effectiveness of publication activities in the activities of both scientific organizations and individual authors is shown.

Keywords

Scientometrics, Hirsch index, impact factor, citation, blockchain, SciWallet, intellectual property, plagiarism, A. Osterwalder's value proposition

1. Statement of the problem and purpose of the study

Currently, we increasingly hear that existing methods for assessing the publication activities of scientists [1] are very subjective [2]. Moreover, due to the increasing spread of the approach to assessing scientific activity based on the calculation of formal quantitative indicators, in particular, the use of the increasingly popular "Hirsch index", when financial decision-making begins to depend on its presence and, accordingly, value, Manipulative approaches to managing such indicators are also emerging [3]. It is no secret to anyone in the scientific community that such a formal quantitative approach to assessing scientific activity carries a number of shortcomings [4], and, to some extent, even contradicts its own purpose, in particular in ensuring transparency in comparison to scientific contribution of various authors to a particular area of scientific activity [5]. At the same time, the number of ways to achieve growth in cherished indicators is only growing - from very dubious [6] to completely worthy of attention [7]. Although, of course, the very idea of creating the H-index is very positive and, in the long term, is quite capable of providing an assessment for the best of the best [8], but, nevertheless, the need for new approaches is becoming increasingly urgent [9].

2. The main material of the article

In connection with the above, the purpose of this study is to propose a model for assessing the contribution of authors of scientific works that would allow the assessment of such contributions in the long term, and would also be maximally protected from manipulations that could in the short term obtain the maximum quantitative assessment of "scientific activity" not creating significant (at least from the point of view of "interest of the scientific community") results.

Therefore, the objectives of this study are the following:
1. Define long-term "success factors" for published scientific output;
2. Conduct an analysis of modern mechanisms that facilitate taking into account such factors;
3. blockchain technology) to create a new index of accounting for the scientific contribution of authors, taking into account the previously cited long-term "success factors";
4. Determine restrictions on its use;
5. Provide appropriate recommendations for use.

Research results. To determine the long-term "success factors" of published scientific results, it is proposed to consider the following value creation model proposed by A. Osterwalder [10]:
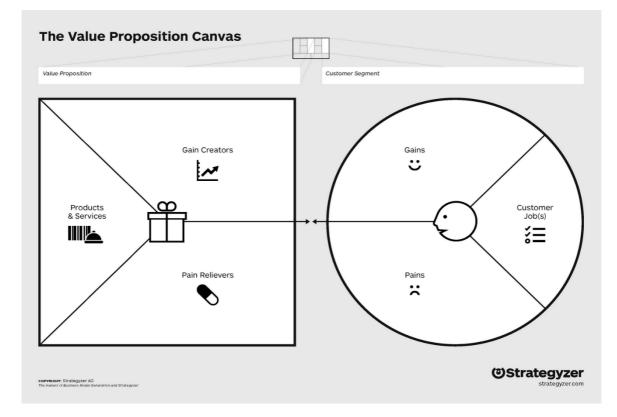

Figure 1: Osterwalder's value proposition model (based on [10])

For each of the blocks of the model, when modifying it for the purposes of this study, we will consider the following parallels:
- Customer Job (s) - A set of actions that must be performed by the author (team of authors) of a scientific work for its publication, including the process of preparing the work itself, its design, reviewing, editing, recording in scientometric databases, etc.
- Gains - The activity of the author (team of authors) that brings him the greatest "satisfaction" ("benefits", "values") from specific elements of the entire process of scientific activity, including the growth of quantitative indicators (citation index, impact factor, etc.) .
- Pains - The activity of the author (team of authors), which brings him the greatest "inconvenience" (leads to loss of "benefit", "value") from specific elements of the scientific process (or the need for its administrative, etc. support), including loss of time , the need to check the correct and timely reflection of indexes in accounting systems, etc.
- Gain creators - Supporting processes (including everything that can be associated with external support - both from the organization and other surroundings of the author), capable of maximizing "satisfaction" ("benefits", "values") from both specific elements of everything the process of scientific activity, and to obtain the final result.

- Pain relievers - Auxiliary processes (including everything that can be associated with external support - both from the organization and other surroundings of the author), capable of minimizing the "inconvenience" (leading to the loss of "benefit", "value") from specific elements of the process of scientific activity (or the need for its administrative, etc. support).
- Products & Services is a set of properties of the "new science citation index", which allows maximizing the benefits from the process of scientific activity, as well as minimizing the negative factors associated with such activity (for example, as noted in [4]).

As modern mechanisms that help take into account the above factors, it is worth noting, along with the SCOPUS project [11], such Internet resources as Researchgate [12] and Google Scholar [13], as well as the Academia project [14].

All of these above tools also offer their own vision of quantitative assessment based on the calculation of the Hirsch index, and Researchgate also offers the calculation of this index both in the classic version and without taking into account self-citations in the author's works.

In order to fully take into account, the proposed factors, the authors of this study propose to use a technology such as blockchain [15].

Currently, the idea of using this technology in scientific activities is not something extraordinary. At a minimum, it is worth noting not only the emergence of "concept" notes [16, 17, 18, 19, 20], or detailed concepts supported internationally [21], but also the emergence of projects aimed at creating products for the scientific community [22, 23, 24, 25]. Moreover, some of the projects have already found their audience and offer already developed blockchain- based services [26, 27], and the budgets being discussed can probably blow the imagination of any rector (and maybe even a minister whose "department" includes science management) in the post-Soviet space [28].

However, almost all implementation examples considered by the authors of this study have their limitations on the use of their products; moreover, no "encroachments" on the power of the H-index in the scientific community were explicitly found, despite proposals even from "scientific coins" [29-31];

To develop the proposed concept, the authors took the idea expressed by Eugene Garfield [32], the famous author of the first Science Citation Index, which he and his colleagues created in the USA back in the early 60s of the 20th century and is now known as the Science Citation Index (SCI), who is also called the "grandfather of Google" [33] that "Citations are a reward system, a bargaining chip with which we pay our colleagues." The technological part of the model is based on the ideology of blockchain technology itself , described briefly in the main conceptual "primary source" for all "electronic money" proposed for the first time for Bitcoin [34].

The continuation of Garfield's statement above - "Failure to cite the sources used in the work is a form of plagiarism" could well be rephrased as "Failure to cite the sources used in the work is a form of robbery", if this is also considered from sides of copyright and intellectual property protection. On the other hand, why not protect this "wallet" using modern technology? Why not make this "bargaining chip" way not more tangible, but at least "calculated" and provide each author - both in the present and even in the past and future with his own "wallet" and start keeping records of it ("scientific wallet") contents? And why not create a scientific index based on this content? For example, the "scientific wallet" proposed by the authors of this study is SciWallet , which could thus become a new scientific citation index.

General description of model elements:
- Scientific work (publication) = "block";
- Author citation = "chain";
- Article writing = "mining";
- Accounting for citations = replenishment of the "scientific wallet";
- "Wallet" = an index of the scientific value of the author's (and therefore the author's) work, currently the most recognized is the "Hirsch Index", supported by SCOPUS;
- "Verification body" for publications = an organization that assigns DOI to a specific publication (for example, based on the functionality of the Researchgate portal);

- "Verification body" for authors = the organization that carries out the identification of the scientific worker, for example ORCID [36].

In order to draw parallels between the capabilities of blockchain and the needs in the field of scientometric activities, let us turn to the description of the technology [15]: Blockchain is a distributed ledger of records of events in the digital world. The key component of the blockchain is the transaction log, and transactions themselves are the only way to change the state of the ledger. The key properties according to [15] are the following four parameters, each of which can be adapted for the needs of scientometric activities:

1. A transaction can be completed either to completion or not at all ("hanging" the operation in an intermediate state is unacceptable). However, entries in the transaction log can only be made with the consent of the majority of network participants.

Proposal of the authors of the article: a transaction (recognition of a "block") is considered "complete" only after publication in the appropriate publication (for example, a scientific publication included in the corresponding scientometric database), and only if the "block" is assigned an identifier (DOI), as well as the author has his own identifier (for example, ORCID).

2. An important feature of the transaction log in the blockchain is its immutability. This property means that you cannot silently remove a transaction from the log or add a new one in the middle of it.

The authors of the article propose: use this feature to verify the authenticity of the "block".

3. The property of immutability is achieved through cryptographic techniques, and not through trust in the organization or people. The two simplest cryptographic algorithms used in blockchain are hash functions and electronic signatures, which ensure the integrity of transactions and are responsible for authorization.

Suggestion of the authors of the article: use the DOI and ORCID functionality.

4. Although the blockchain is a distributed system, and each node can generate transactions, this does not mean that all participants in the blockchain network have equal rights - in almost any implementation of this technology, a distribution of roles has been introduced into validators (participants who write transactions in the journal), auditors and easy clients. Moreover, this division is true not only for private blockchains but also for public blockchains, such as Bitcoin.

The authors of the article propose: to consider the editorial offices of scientific publications (journals) as "validators", representatives of scientometric databases (for example, SCOPUS) as "auditors", and the actual authors of scientific works as "light clients".

When preparing proposals for using the properties of blockchain, the authors used the properties, as can be seen from the text, primarily described in relation to the "first generation" of the commercial implementation of this technology - the creation of "bitcoin". On the other hand, the system can be more complex, for example, when a larger number of information sources are interconnected (in our case - information about co-authors, keywords, data about the journal in which the publication was made, etc.). In this case, as noted in [37], more complex functions can be used to combine hashes obtained from processing data blocks (article structure) and obtaining the resulting hash (integral publication score):
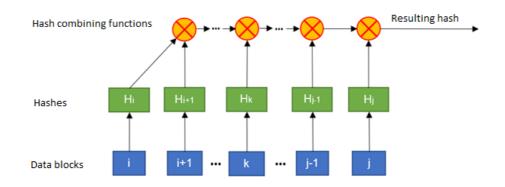

Figure 2: Operation of a binary function (according to [37])

Moreover, if we take into account, as noted in [38], that there are two types of chain: Public Blockchain - an open, expandable database. This type of blockchain is used in the Bitcoin cryptocurrency. Each participant can write and read data, as well as a Private or private blockchain, which has restrictions on writing/reading data, in which case priority nodes can be established. A subtype of Private Blockchain is an exclusive blockchain. In such a chain, a group of individuals is established to process transactions. In the context of the presented research, this allows us to build both open (and, accordingly, free) models of this kind of "wallet", and closed (and, accordingly, paid) models. For example, "Public SciWallet " could be considered for use in Google Scholar (or similar open platforms), and "Private SciWallet " could be used in scientometric databases such as SCOPUS.

As a very minimal model of a "block" that could be built into blockchain technology, consider the following structure:

1. Job title;
2. Information about the author(s), including the ORCID of the author(s);
3. Publication DOI;
4. Publication date information;
5. Information about the sources used, including the DOI of such sources.

The logic of the work is as follows - any publication that is a member of the SciWallet system, in case of citation, "receives" X points, which are distributed both among co-authors (if any) and between publications (and subsequently between the authors of such publications) that were used in the preparation of the work itself.

It is possible to use a system of "weights" for distribution both among co-authors and among sources, with each source being considered as another "co-author", regardless of the number of authors of such a source. To determine the number of points itself, a system of increasing/decreasing coefficients can be applied depending on the "quality" of the publisher in which the work referencing the publication calculated in the model was published. In turn, the "points" obtained in this way are distributed among the cited works and their authors.

For example, if X = 100, one author and nine sources used, in the simplest version of the model, which can simply treat the contribution of all participants as "equal", then 1/10, i.e. 10 points are "transferred" to the account of a specific source publication, which, in turn, are distributed in a similar way within the co-authors and sources of the cited publication.

In this case, it is possible to provide a "score" in the form of that same "bargaining chip" for those publications that were not directly indicated by the authors of the publication calculated in the model due to the fact that such works served as sources for those publications that were directly used by the authors of the work. sources of the publication in question. In such a model, the number of "points" for each publication, as well as for each author along the entire chain, can be taken into account. In this way, each "block" can include the information necessary to be included in the "publication blockchain ".

Moreover, to "start" such a system, it is possible to provide for the provision of an initial "credit" to each first-time registered author so that he "repays" his "debt" to those authors whose work he has already used, regardless of the possible replenishment of his own SciWallet. To ensure a more reliable assessment, an additional rule can be introduced - for example, in such a system the author may not have a "displayed" public "index-wallet" until he makes the "first contribution" to "predecessors" - works cited in the amount of a fixed volume corresponding to the value of this primary "loan", and only after that receive a public index in the form of the "balance" that has developed at that moment. With this approach, however, it is also possible to obtain a negative balance for the "scientific contribution" of such an author.

Despite the apparent simplicity of this approach, this model, according to the authors, can provide a truly objective quantitative indicator of the "scientific contribution" of the author, and will effectively motivate authors to be responsible both for the content of the works proposed for publication and for the choice of the publisher, and also to promote the results of their research in the professional community.

An obvious obstacle to the implementation of the proposed model is the incomplete coverage of both authors and publications themselves with digital identifiers.

Of course, further improvement of the proposed model will require serious consideration of the best practices that have been developed in world practice, in particular, as was done in the dissertation work [35], although as of 6 years ago. On the other hand, it is already possible, taking into account the possibility of obtaining data from existing scientometric databases, to conduct a pilot project to create a prototype of such a model based on real data.

3. Conclusions

As can be seen from the results of the study, we are on the verge of big changes. Moreover, these changes are inevitable, if only because too many initiatives in this direction have already started. As noted in [15]: "It is safe to say that new projects on the blockchain will be based on its main advantages - openness, security, safety. Therefore, blockchain will be a good help for any services where users are concerned about the problems of fraud and the safety of personal data."

As another evidence of the applicability of blockchain technology for the implementation of the proposed model, we present the following figure:
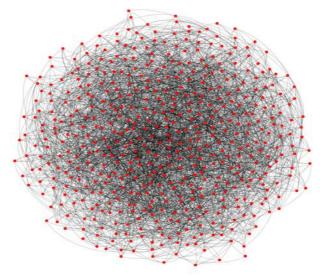

Figure 3: Visual display of a decentralized Bitcoin server (based on [39])

This visualization, showing the connections between blocks for an example of a decentralized Bitcoin server, can also be built for a decentralized server that stores information about the connections between the works of various authors and "citation chains" in the scientific community, incl. using data that can be obtained from already existing scientometric databases, for example, by modifying the approach to visualizing global connections between researchers. As other parameters of such scientometric models, one can consider not only geographic location, but also analyze interdisciplinary connections:

And if in the world of cryptocurrency, one can already observe an ever-increasing variety of all kinds of electronic " coins ", then, quite possibly, those projects that have already started in the field of application of blockchain technology in scientific activity, also, with their further development, they will come to the idea of \u200b\u200b"digitizing" that very "small coin" that Eugene Garfield spoke about and, albeit by solving the "inverse problem", assessing the real contribution of each of the researchers to the total world knowledge, perhaps forming registries of both "creditors" and "debtors" in the digital world.

4. References

[1]   Interpretation of three important scientific concepts: impact factor, citation index, h-index [Electronic resource] URL: https://wos- scopus.com/traktovka-treh-vazhnyh-nauchnyh-ponjatij/.

[2]   Tsyganov, A.V. Brief description of scientometric indicators based on citation rate [Electronic resource].
      URL: https://academcity.org/sites/default/files/docs/attachments/cyganov.pdf
[3]   How to "increase" the Hirsch index [Electronic resource] URL: https://luchecon.livejournal.com/903188.html.
[4]   Polyanin, A.D. Disadvantages of citation and Hirsch indices. Maximum citation indices [Electronic resource] URL: http://www.nbuv.gov.ua/node/1365.
[5]   Gainetdinov, R. "The competition critically lacked transparency" [Electronic resource] URL: https://www.gazeta.ru/science/2010/11/02_a_3434703.shtml.
[6]   Chasing deer [Electronic resource] URL: http://scientific.ics.org.ru/gid-park/pogonya-za-olenyami1.
[7]   K. Kolesnikova, O. Mezentseva and O. Savielieva. (2019). Modeling of Decision-Making Strategies In Management of Steelmaking Processes. *2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT)*, Kyiv, Ukraine, pp. 455-460. doi: 10.1109/ATIT49449.2019.9030524.
[8]   Sednev, Yu. Citation, rating (first thousand with H>100 Hirsch) - analogue of the human central nervous system [Electronic resource] URL: http://e-cambridge.info/ratings-citation-thousands-over-100s.
[9]   Shpachinsky, I.L. Blockchain as a means capable of overcoming corruption: world experience and prospects for technology development in Ukraine [Electronic resource] /I.L. Shpachinsky, A.V. Kushchev
      URL: http://molodyvcheny.in.ua/files/journal/2017/12/11.pdf.
[10] The Value Proposition Canvas [Electronic resource]
      URL: https://strategyzer.com/canvas/value-proposition-canvas.
[11] Scopus [Electronic resource] URL: https://www.scopus.com/.
[12] Researchgate [Electronic resource] URL: http://researchgate.com/.
[13] Google Academia [Electronic resource] URL: https://scholar.google.com.ua.
[14] Academia [Electronic resource] URL: https://www.academia.edu/.
[15] "What do they eat it with": what is blockchain [Electronic resource] URL: https://habr.com/company/bitfury/blog/326340/ – Access date: 06/25/2018.
[16] Could Bitcoin technology help science? [Electronic resource] URL: https://www.nature.com/articles/d41586-017-08589-4 – Access date: 06/25/2018.
[17] Zlenko, O.A. Study of the possibilities of applying blockchain technology to the process of confirming authority of a scientific publication //O.A. Zlenko, L.A. Makushkina/ [Electronic resource] URL: https://www.scienceforum.ru/2017/2248/26939.
[18] Novikova, O. Blockchain technology for tracking scientific research [Electronic resource] URL: https://freedman.club/tehnologiya-blockchain-dlya-otslejivaniya-nauchnih-issledovanii/.
[19] K. Kolesnikova, A. Mukhamediyeva, K. Alpysbayev and O. Kolesnikov. (2023). Development of a Cognitive Model of the Interaction of Knowledge Carriers in the Educational Environment. *2023 IEEE International Conference on Smart Information Systems and Technologies (SIST)*, Astana, Kazakhstan, pp. 322-326. doi: 10.1109/SIST58284.2023.10223564.
[20] How will blockchain change the fundamental scientific field? [Electronic resource] URL: https://decenter.org/blockchain/572-blockchain-in-science-rus.
[21] Grech, A. Blockchain in Education /Alexander Grech, Anthony F. Camiller [Electronic resource] URL:
      http://publications.jrc.ec.europa.eu/repository/bitstream/JRC108255/jrc108255_blockchain_in_education(1).pdf.
[22] Everything is fair. Belarusians have started a blockchain revolution in science [Electronic resource] URL: https://dev.by/lenta/main/nauka-po-chestnomu.
[23] Blockchain will become the basis of the copyright management system [Electronic resource] URL: https://indicator.ru/news/2017/12/13/blokchejn-upravlenie-avtorskimi-pravami/.

[24] Is it time for a "blockchain revolution" for science? [Electronic resource] URL: https://scientificcoin.com/blog1_ru.

[25] Kolesnikova, K., Lukianov, D., Gogunskii, V., Iakovenko, V., Oborska, G., Negri, A., Koliada, A., Dmitrenko, K., Olekh, T., & Bespanskaya-Paulenka, K. (2017). Communication management in social networks for the actualization of publications in the world scientific community on the example of the network researchgate. Eastern-European Journal of Enterprise Technologies, 4(3 (88), 27–35. https://doi.org/10.15587/1729-4061.2017.108589.

[26] IPChain Features [Electronic resource] URL: https://ipchain.ru/possibilities/.

[27] Globex SCI is the world's first blockchain platform for the exchange of scientific publications! ICO review. [Electronic resource] URL: https://goldvoice.club/@duosre/globex-sci-pervaya-v-mire-blokchein-platforma-po-obmenu-nauchnymi-publikaciyami-ico-obzor/.

[28] Startup Incubator Science Could Raise Up to $100 Million in Planned Token Sale [Electronic resource] URL: https://www.coindesk.com/startup-incubator-science-plans-100-million-token-sale/.

[29] SMART CONTRACTS FOR SCIENTIFIC DEVELOPMENTS [Electronic resource] URL: https://scientificcoin.com/blog4_ru.

[30] Fuel the engine of scientific progress - buy the Globex SCI token [Electronic resource] URL: https://globexsci.io/ru/.

[31] Science Blockchain releases fourth quarter NAV statement [Electronic resource] URL: https://medium.com/@scienceinc/science-blockchain-releases-fourth-quarter-nav-statement-1638fa8a99b0.

[32] Eugene Garfield 1925 - 2017: Visionary information scientist [Electronic resource] URL: https://www.cwts.nl/news?article=n-q2y2c4&title=eugene-garfield-1925--2017-visionary-information-scientist.

[33] Eugene Garfield: Librarian* & Grandfather of Google [Electronic resource] URL: https://blog.lib.uiowa.edu/hardinmd/2010/07/12/eugene-garfield-librarian-grandfather-of-google/.

[34] Nakamoto S. Bitcoin: A Peer-to-Peer Electronic Cash System [Electronic resource] / Satoshi Nakamoto. – 2008./ / Mode Access : https://bitcoin.org/bitcoin.pdf.

[35] Riviera, E. Mapping scientific literature. Structuring scientific communities through Scientometrics [Electronic resource] URL: https://boa.unimib.it/retrieve/handle/10281/40095/59466/Phd_unimib_732752.pdf.

[36] Connecting Research and Researchers [Electronic resource] // Access mode: https://orcid.org/.

[37] How it works: Merkle trees in the Bitcoin network [Electronic resource] URL: https://habr.com/company/bitfury/blog/346398/.

[38] What is blockchain in simple words [Electronic resource] URL: https://prostocoin.com/blog/blockchain-guide.

[39] What is blockchain? Let's tell you in simple words [Electronic resource] URL: https://coinspot.io/beginners/chto-takoe-blokchejn-rasskazhem-prostymi-slovami/.

[40] Beauchesne, Olivier H. Map of scientific collaboration between researchers [Electronic resource] URL:http://olihb.com/2011/01/23/map-of-scientific-collaboration-between-researchers/.