doi:https://doi.org/10.15407/emodel.40.06.53

УДК 004.056.53+349

H.A. Kravtsov, Cand. Sc. (Eng.), **A.V. Zupko,** Post-graduate,
G.E. Pukhov Inst. for Modeling in Energy Engineering
of National Academy of Sciences of Ukraine
(15, General Naumov Str., 03164, Kiev, Ukraine,
тel. (044) 4241063, e-mail: hryhoriy.kravtsov@gmail.com, andrey.zupko@gmail.com)

Blockchain and Science

The scientific community is actively discussing how blockchain technology can solve some specific challenges like limited access to research results, General Data Protection Regulation (GDPR) compliance, reproducibility crisis and absence of negative results that are rarely shared. In this paper authors make an attempt to address the main advantages of the blockchain technology and simulate the situation when some steps in a research lifecycle can leverage these advantages. Some examples how blockchain can streamline the whole scientific process are shown.

Keywords: blockchain, GDPR, compliance, transparency, trust, decentralization, security, fraud prevention, value exchange, micropayments, consensus.

Blockchain features. The blockchain is a special type of a database or peer-to-peer distributed ledger [1]. All discussions about blockchain application in Science are built around pros or cons of blockchain.

It seems to be one of the most expensive ledgers and in many cases, it doesn't make any sense to replace existing databases with blockchain. But in some cases blockchain may have a significant positive impact because it has several interesting features:

Transparency and Trust. It's immutable and append-only — records can only be added to that database and never removed or changed. It's updateable only via consensus or agreement on the state of the data among peers.

Decentralization. Blockchain databases are distributed among multiple computers (nodes) that store full or partial copies of that database. This removes monopoly and single point of authority and keeps trust between parties.

Security and Fraud prevention. The blockchain records are secured through cryptography. Every transaction is signed with a personal digital signature. If a record is altered, the signature will become invalid and the peer network will know right away that something has happened. Blockchain doesn't have a single point of failure and can't be changed from a computer.

Value Exchange and Micropayments. Cryptocurrencies allow transferring value between parties without banks, governments. Transaction cost is almost zero comparing to Visa/Mastercard payments (600,000 transactions for $0.01 in Stellar).

A distributed peer-to-peer network has one significant disadvantage — lack of trust. This problem was formulated back in 1982 under the title “The Problem of Byzantine Generals” [2]. The solution of problem is precisely the technology of blockchains. And, depending on how this problem is solved, blockchains are divided into different types which have been explained in [3]. Firstly, we have to explain what does consensus mean.

Consensus is defined [3] as a general agreement of a state that the blockchain is in. That means, if Alice sends $100 worth of Bitcoin to Bob, Alice will lose $100 worth of Bitcoin from her wallet, and Bob will gain $100 worth of Bitcoin in his wallet. The catch is that every clean transaction has to be recorded on the Bitcoin public ledger, and a consensus algorithm ensures no malicious transactions nor changes can be made on the blockchain itself. In accordance with [3] we know the following types of consensus algorithms: Proof-of-Work (PoW), Proof-of-Stake (PoS), Delegated Proof-of-Stake (DPoS), Proof-of-Authority (PoA).

Proof-of-Work. Most of us might heard of PoW, especially since the first public blockchain — Bitcoin, uses PoW. In this example, PoW is explained by means of Bitcoin.

PoW is conducted through miners (the people keeping the blockchain running by providing a huge amount of computing resources) competing to solve a cryptographic problem — also known as a hash puzzle. These miners help to verify every Bitcoin transaction, where it involves producing a hash-based (SHA256) PoW that is based on previous transaction blocks (read up on the Merkle Tree for more information) and forms a new branch with a new transaction block. This means that the work is rather difficult for the miners to perform but easy for the network to verify. The first miner who manages to produce the PoW will be then awarded by some Bitcoins. The amount of Bitcoin awarded decreases over time. Over the years, as the difficulty level in mining Bitcoin has increased tremendously, resulting in PoW being notorious for the amount of energy it requires to keep the blockchain running.

Proof-of-Stake. Unlike PoW where new transaction blocks are created based on computational work done by solving a complex cryptographic puzzle, PoS allows a forger (instead of a miner) to stake any amount of cryptocurrency he/she has, to be probabilistically assigned a chance to be the one validating the block. The probability based on the amount of cryptocurrency staked.

Additionally, for most PoS systems, instead of receiving a cryptocurrency reward (in the above case, the Bitcoin miners receive some Bitcoins for solving a PoW), the forgers instead of take the transaction fees as rewards.

The idea of putting coins to be ‘staked’ prevents bad actors from making fraudulent validations — upon false validation of transactions, the amount staked will be forfeited. Hence, this incentivises forgets to validate legitimately. Last year, PoS has gained attention, with Ethereum switching towards a PoS from a PoW consensus system.

Delegated Proof-of-Stake is similar to PoS in regard to staking but has a different and a more democratic system that is considered to be fair. Like PoS, token holders stake their tokens in this consensus protocol. Instead of the probabilistic algorithm in PoS, token holders within a DPoS network are able to cast votes proportional to their stake to appoint delegates to serve on a panel of witnesses — these witnesses secure the blockchain network. In DPoS, delegates do not need to have a large stake, but they must compete to gain the most votes from users.

It provides better scalability compared to PoW and PoS as there are fully dedicated nodes who are voted to power the blockchain. Block producers can be voted in or out at any time, and hence the threat of tarnishing their reputation and loss of income plays a major role against bad actors. No doubt, DPoS seems to result in a semi-centralised network, but its traded off for scalability. Like PoS, DPoS has also gained attention over the years with several projects adopting this consensus algorithm. Since it was invented by Dan Larimer, DPoS has been refined continuously, from BitShares to Steem and now in Ethereum.

Proof-of-Authority. PoA is known to bear many similarities to PoS and DPoS, where only a group of pre-selected authorities (called validators) secure the blockchain and are able to produce new blocks. New blocks on the blockchain are created only when a super majority is reached by the validators. The identities of all validators are public and verifiable by any third party — resulting in the validator’s public identity performing the role of proof of stake. As these validators identity are at stake, the threat of their identity being ruined motivates them to act in the best interest of the network. Due to the fact that PoA’s trust system is predetermined, concerns have been raised that there might be a centralised element with this consensus algorithm. However, it can be argued that semi-centralisation could actually be appropriate within private/consortium blockchains — in exchange for better scalability.

Cryptocurrencies are cryptographically secure digital money, the internal accounting unit of any community that declares its confidence in a given unit. The issuance of cryptocurrency is carried out during the course of mining. One major feature of the cryptocurrency is the anonymity of a sender and a recipient. This feature is often criticized by governments as a way to launder criminal money. However, it should be borne in mind that anonymity ends at the moment when the holder of the cryptocurrency tries to transfer it to fiat money, i.e.

money that is emulated by central banks, for example, the dollar, the euro. The ability of the cryptocurrency to be converted into fiat money makes it possible to use the cryptocurrency as a tool for financing new projects. And it is divided through the procedure for the initial placement of coins of the cryptocurrency (ICO) [4].

Scientific projects funding with ICO. Not every scientific project can be financed through the ICO procedure. Below are the expectations from the project, which can be subject to the subsequent release of the cryptocurrency:

The scientific project should have practical application with ability to monetize the results of study.

Monetization of results can be carried out within electronic commerce.

Particular attention is paid to projects with potential speculative cost.

For example, the energy market stated its interest in blockchain technologies. Grid+ platform attracted 29 million dollars [5] on the initial placement of the crypto token. The Grid + platform is designed to monitor consumption and increase savings when consuming electricity. The Power Ledger system is a distributed network for the sale / purchase of renewable energy and raised $ 17 million during the ICO [6].

Challenges in the scientific environment. Let's describe the problems in Science and how real blockchain projects can be applied to some of these problems.

The limited access to research data and results. A detailed description of the research design and full research data set is rarely available. Knowledge is controlled by centralized companies and scientific discoveries are kept behind paywalls [7].

General Data Protection Regulation (GDPR) compliance. Very few organisations comply with all GDPR requirements.

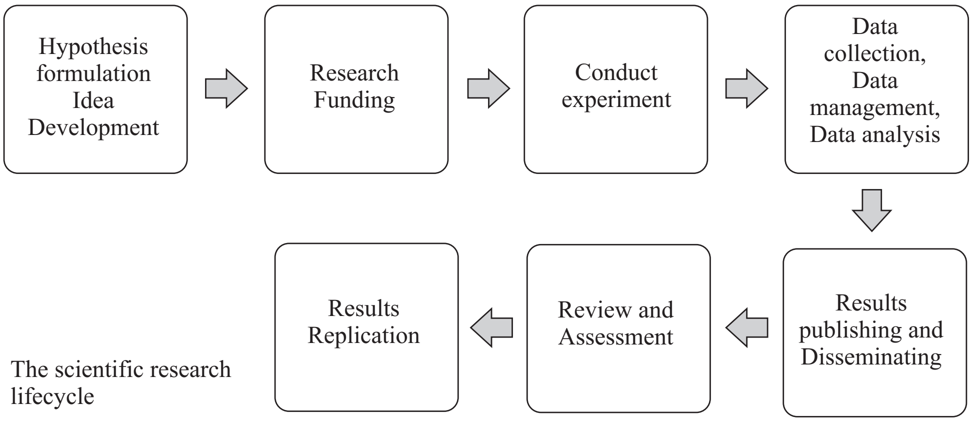
Replication and Reproducibility crisis. Results confirming established information are rarely published, increasing thus the concern over the reliability of the scientific reports.

Only successful results are published. Even though failure is a necessary part of making progress, null hypotheses and negative results are rarely shared within the scientific community.

When you have a hypothesis and funding it is hard to find collaborators and contributors worldwide especially when you need some very specific skills.

Figure shows the main steps of the scientific research lifecycle [8]. If to apply blockchain on different stages of scientific research some specific problems may be solved.

The GDPR is the biggest overhaul of the European Union (EU) data protection law in more than 20 years. It replaced the EU Data Protection Directive and aims to create unified data protection legislation covering all individuals in EU.


Hypothesis formulation Idea Development
Research Funding
Conduct experiment
Data collection, Data management, Data analysis
Results publishing and Disseminating
Review and Assessment
Results Replication

The scientific research lifecycle

Compliance is crucial due to the impacts personal data processing can have upon people's lives. GDPR revises and enhances the requirements on organisations to consider data protection and accountability, providing individuals new rights over how their data is used.

Hypothesis formulation, Idea Development. A Study design can be pre-registered to a blockchain to avoid further arbitrary alteration of study design after the experiment. It can also prevent the arbitrary suppression of research studies from being published in case the results do not meet certain expectations.

Research Funding. ScienceRoot project manages a platform that lists grants from around the world. Researchers can also present their ideas and crowdsource funding with the ICO process.

Conduct experiment. Researchers can upload to a blockchain all the data they collect during the experiment. All these data have direct attribution to its owner and leverage "proof-of-existence" mechanism. This will add trust to the final research results. A researcher can keep all the data encrypted and private till final results are published. Researchers can grant limited access to those who conduct review process before publication.

Results Replication. The Replication Foundation aims to fund scientists who want to replicate some specific research findings — an important part of scientific process. The Foundation acts of decentralized autonomous organization whose rules and financial transaction records are maintained on a blockchain. Everyone can submit proposals for replication studies and funding. The community can then vote on these proposals and if a quorum is reached in a certain time period, the funding is transferred to the proposal's author.

Find collaborators and contributors for your research. Many research studies require some involvement of experts from very different areas. And it is very hard to identify where such experts exist and if they are available for a new research project. Sometimes appropriate expert is sitting next door but you even don't know that he has appropriate skills. NaomiHire blockchains marketplace aims to solve this problem. NaomiHire has the most precise skills matching algorithm on the market built on unique mathematical theory the calculus over classification which is used by artificial intelligence [9].

Universities can upload detailed scientists profiles with all their hard and soft skills and academic experience. Research groups that look for some people with specific expertise can create a job request and fill the ideal profile of the person they are looking for. NaomiHire AI automatically matches scientists with available jobs and builds a very detailed matching report based on skills relevance and cost efficiency. Both parties can use Smart-Contracts to sign an agreement for specific work.

Science and GDPR. In May 2018 every active internet user received hundreds of emails with a request to accept new agreements because of GDPR. But some researches show that most of the organizations were no ready to be compliant with the GDPR. Many companies just updated their Privacy Policies and Terms of Use agreements. But GDPR goes far beyond these documents. Especially GDPR is critical for HealthCare and social scientific researches where user data is actively used. Below are some of the key GDPR requirements:

- All processing should be based on a legitimate purpose and customer has to be aware of what data company process and how a company use it.
- Collect only that data which is necessary, and not keep personal data once the processing is finished.
- The customer can ask to delete or transfer his personal data.

Companies to notify customers where they share information with other organizations.

Blockchain can help to follow all these rules. Customer personal data can be stored inside some permissioned Blockchain. All data is encrypted with end-to-end encryption. Only the customer has a private key to decrypt those data. The customer complete control of their personal information and can determine what is used by companies and how. In case if some company ask to provide access to specific personal information the customer can now use their own digital signature (or fingerprint) and combine that with a company's signature to unlock and release those specific data. It provides restricted access that can only exist if there is verification from both the customer and the company.

Blockchain could also control sharing data across systems and organizations. When a company needs to share data with some 3rd party partners the customer receives a request and has to formally provide his digital signature for such an action. Finally, the customer may revoke or limit access to his personal data any time.

Conclusion

Scientific landscape has some significant challenges and community is actively looking for a solution. Blockchain features like Transparency and Trust, Decentralisation, Privacy and Security may help to some of those challenges and streamline all scientific processes.

REFERENCES

1. Imran Bashir. Mastering Blockchain, Distributed ledgers, decentralization and start contracts explained.
2. Gadi Taubenfeld. Synchronization Algorithms and Concurrent Programming, Prentice Hall; 1 edition.
3. Evan Tan. Types of Consensus Protocols Used in Blockchains, available at: https:// hackernoon.com/types-of-consensus-protocols-used-in-blockchains- 6edd20951899.
4. Intial Coin Offering (ICO) Investopedia, available at: https://www.investopedia.com/terms/i/initial-coin-offering-ico.asp.
5. Grid+ Raises $29 Million as Blockchain Fever Grows, Jason Deign, greentech media, Sep. 22, 2017, available at: https://www.greentechmedia.com/articles/read/grid-raises-40-million-as-block chain-fever-grows#gs.zd2DOPg.
6. Blockchain Energy Trading Startup Power Ledger Raises $17M in Cryptocurrency CO Jeff st. John, greentech media, Sep. 06, 2017, available at: https://www.greentechmedia.com/articles/read/power-ledger-blockchain-energy-trading-startup-raises-17-cryptocurrency#gs.6G6CYYI.
7. Blockchain for science and knowledge creation, Dr.med.SönkeBartling, Benedikt Fecher, August 2016, available at: https://www.researchgate.net/publication/306107836_Blockchain_for_science_and_knowledge_creation_-_A_technical_fix_to_the_reproducibility_crisis.
8. Development of the research lifecycle model for library services, K.T.L. Vaughan, MSLS; Barrie E. Hayes, MSLS; Rachel C. Lerner, MSLS, *Journal of the Medical Library Association,* October 2013.
9. Kravtsov, H.A., Koshel, V.I., Dolgorukov, A.V. and Tsurkan, V.V. (2018), “Trainable model of the calculus over classifications”, *Elektronnoe modelirovanie*, Vol. 40, no. 3, pp. 63-76.

Received 17.09.18

Г.О. Кравцов, А.В. Зупко

БЛОКЧЕЙН І НАУКА

Наукова спільнота активно досліджує, як за допомогою технології блокчейн можна вирішити такі наукові проблеми, як обмежений доступ до результатів досліджень, відповідність загальному регулюванню захисту даних (GDPR), криза відновлюваності та відсутність негативних результатів, які рідко публікуються. Зроблено спробу показати основні переваги технології блокчейн та розглянуто ситуацію, коли ці переваги можна використати на деяких етапах життєвого циклу наукового дослідження. Наведено приклади того, як за допомогою блокчейн можна впорядкувати весь науковий процесс.

Ключові слова: блокчейн, GDPR, дотримання, прозорість, довіра, децентралізація, безпека, запобігання шахрайству, обмін цінністю, мікроплатежі, консенсус.

Г.А. Кравцов, А.В. Зупко

БЛОКЧЕЙН И НАУКА

Научное сообщество активно исследует, как с помощью технологии блокчейн можно решить такие проблемы в науке, как ограниченный доступ к результатам исследований, соответствие общему регулированию защиты данных (GDPR), кризис воспроизводимости и отсутствие негативных результатов, которые редко публикуются. Сделана попытка показать основные преимущества технологии блокчейн и рассмотрена ситуация, в которой эти преимущества могут быть использованы на некоторых этапах жизненного цикла научного исследования. Приведены примеры того, как с помощью блокчейн можно упорядочить весь научный процесс.

Ключевые слова: блокчейн, GDPR, соблюдение, прозрачность, доверие, децентрализация, безопасность, предотвращение мошенничества, обмен ценностью, микроплатежи, консенсус.

KRAVTSOV Hryhoriy Alekseevich, Cand. Sc. (Eng.), acting senior scientific worker of the G.E. Pukhov Institute for Modelling in Energy Engineering of National Academy of Sciences of Ukraine. In 2000 graduated the Pavel Nakhimov Sevastopol Naval Institute. The scientific interests — cyber security of smart- grids, cryptography, development of heterogeneous information systems.

ZUPKO Andrey Vasilevich, PhD student, G.E. Pukhov Institute for Modeling in Energy Engineering of National Academy of Sciences of Ukraine. Graduated with Masters degree in 2007 from the Taras Shevchenko National University of Kyiv. Area of scientific research — blockchain, artificial intelligence, machine learning, data science.