The use of blockchain technology in the scientific research workflow

Bonnie Lawlor*
Chair, IUPAC Subcommittee on Publications, Radnor, PA, USA

Abstract. This paper is based upon a lightening talk that was given on February 14, 2023 at the 2023 NISO Plus conference. The presentation described the research for the creation of a white paper on the use of blockchain technology along the scientific research workflow that will be released in early 2024 by the International Union of Pure and Applied Chemistry (IUPAC).

What the three-year study found is that the technology is indeed being used in almost all the steps in the scientific research workflow - from hypothesis development through to publication - by commercial organizations as well as by non-profits and across all market sectors, even governments. For example, the U.S Department of Health and Human Services (HHS) uses blockchain technology in a pilot program, the Grant-recipient Digital Dossier (GDD), to manage their grant program more efficiently. As of July 2021, GDD had reduced the time required to complete grant assessment tasks from four-plus-hours to a fifteen-minute process.

This paper briefly summarizes the findings, discusses the pros/cons of the technology, and provides a glimpse of how the technology is impacting the future of scientific research.

Keywords: Blockchain technology, scientific research workflow, bitcoin, IUPAC, Distributed Ledger System, ARTiFACTS, Bloxberg Consortium, time stamping

1. Introduction

The development of the white paper was initiated because during the 2019 IUPAC World Chemistry Congress a representative from one of IUPAC's Member Countries raised the question "What is Blockchain Technology?" They felt that it had become a prominent buzzword, not unlike Artificial Intelligence and asked if IUPAC could provide information on how it was impacting science in general and chemistry in particular. At that time, I was already personally fascinated by the technology and had written a paper on it the prior year for IUPAC's news journal, *Chemistry International*. After the Council meeting the current IUPAC President, Javier García-Martínez, suggested that I gather a team to develop a white paper on the use of blockchain technology in scientific research and in 2020 we began a three -year journey to do just that. While I am not a technologist myself, all of the other members of the team are, and several had already been using the technology and were well-positioned within the global blockchain community. This facilitated our access to blockchain experts from around the world and we unearthed a wealth of information which was not presented at NISO Plus due to time restrictions. The white paper itself will drill down in much more detail, plus provide references to a lot of reading material for those interested in learning more. But before I provide some highlights of what we found, let me briefly describe it and its history.

---

*Tel.: +1 610 293 0207; E-mail: chescot@aol.com.

2. What is blockchain technology???

According to a report [1] from the National Institute of Standards and Technology (NIST), "Blockchains are immutable distributed digital ledger systems (i.e., without a central repository) and usually without a central authority. At its most basic level, they enable a community of users to record transactions in a ledger public to that community such that no transaction can be fraudulently changed once published". Indeed, an alternate name for blockchain is Distributed Ledger Technology (DLT), not unlike the double-entry book-keeping method that dates back to the fourteenth century. Basically, it is a technology platform that can be used to support many diverse applications. While this new technology gained popularity because of its use by those interested in cryptocurrencies, specifically Bitcoin, the technology is simply the engine "under the hood" of Bitcoin - an engine that can be used for other purposes as well. Indeed, "Blockchain is to Bitcoin, what the Internet is to email.

In fact, blockchain technology actually *predates* Bitcoin by almost twenty years. It was co-invented in 1991 by Stuart Haber and W. Scott Stornetta who both worked at Bell Communications Research (Bellcore) and who were attempting to ensure the integrity of digital records via time stamping. They believed that the ability to certify when a document was created or last modified would be essential to the resolution of conflicts over such things as intellectual property rights. Their initial efforts involved working on a cryptographically-secured chain of blocks such that no one could tamper with the timestamps of documents - hence the name "blockchain".
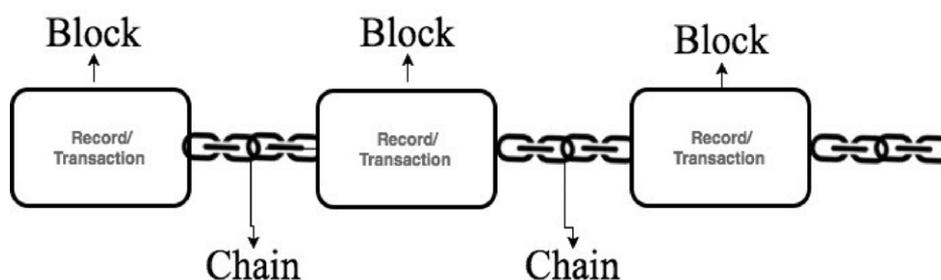
Cc8be069bda4863a638d900f827235ef (hash for "I love green tea")

The blocks record and confirm the time and sequence of transactions and each block contains a digital fingerprint or unique identifier called a "hash", timestamped batches of recent valid transactions, and the hash of the previous block. The hash of the previous block links the blocks together and prevents any block from being altered or a block being inserted between two existing blocks. Each subsequent block strengthens the verification of the previous block and hence the entire blockchain. The method renders the blockchain tamper-evident, lending to the key attribute of immutability. The technology is important for many reasons, but the primary one is that it provides incontrovertible proof-of-creation of an idea, research data, etc. Once hashed to a blockchain and time-stamped, the data can be changed, but such changes are captured and time-stamped, making fraudulent tampering visible. This makes digital goods immutable, transparent, externally-provable, decentralized, and distributed.

Haber and Stornetta left Bellcore in 1994 to co-found a spin-off company, Surety, which provided time-stamping services based upon their algorithms and it was the first company to provide commercial blockchain-based services. Things remained quiet on the blockchain front until 2008 when Satoshi Nakamoto released a White Paper entitled, "Bitcoin: a peer-to-peer electronic cash system" [2] which

proposed a system of electronic transactions that did not require a reliance upon trust - the middleman, in this case a bank, was removed. Haber and Stornetta's seminal paper, "How to Time-Stamp a Digital Document" [3], is referenced in the Bitcoin White Paper.

Over the decade since Nakamoto built upon Haber and Stornetta's work there have been a series of enhancements to the technology, primarily driven by the realization that it did *not* need to be tied to Bitcoin - it could be used for all sorts of cooperative efforts between organizations - including scientific research – and by 2017 its use in science was well on its way. Let me show you a few examples of what we found when we started research for the White Paper.

3. The Scientific research workflow

We started our journey by defining our view of the scientific research workflow - what areas did we want to examine to see how blockchain was being used and how well it was performing. We looked at five steps in the flow: Step #1: Develop Hypothesis/Define an idea; Step #2: Seek Funding; Step #3: Perform the Experiment/make observations; Step #4: Perform analysis/make insights; Step #5: Publish/share results. During our research over the past three years, we held about a dozen in-depth interviews with major global players across diverse disciplines who are successfully using blockchain technology for a variety of purposes in the scientific workflow and the white paper will provide details on these uses.

4. Step #1: Develop hypothesis/define an idea

The major use of blockchain in the first step is for the time-stamping of ideas - the purpose for which the technology was originally developed. It provides researchers with their proof-of-concept and intellectual property ownership and many organizations around the globe have been offering this service for years. All the providers of this service, just a few of which are listed below, work the same way.

- ARTiFACTS   https://artifacts.ai (basic system free to researchers)
- Bernstein   https://www.Bernstein.io (can try for free)
- Bloxberg   https://bloxberg.org (member-based global Consortium)
- Stampery   https://www.stampery.com

You upload your information so that it can be "hashed" in order to create a tamper-proof timestamp for it. A hash is a concatenation of letters and numbers that uniquely identifies your asset - similar to a human

fingerprint. The hash is uploaded to the blockchain, where it cannot be changed or deleted afterward. The time when the hash is added to the blockchain is your tamper-proof timestamp. The players listed above are based in Europe with the exception of ARTiFACTS - a U.S. company based in San Diego, CA, U.S.A. It was launched in 2018 as the first blockchain-based service for securing provenance (via time stamping) of unpublished research, making *all* work eligible for formal citation recognition and impact reporting. They have since expanded and provide services for every step in the research workflow except for funding, and individual researchers can use their basic system for free. ARTiFACTS has recently embarked on a new path with its launch of *Verify* [4]. *Verify* is a platform designed to assist in the detection of substandard and falsified (SF) drugs at all stages in the pharmaceutical supply chain. *Verify* fills an important gap in safety and regulatory practices, through its collaboration with the Lieberman Lab at the University of Notre Dame, to capture-analyze-track the data associated with chemical analysis of medicine samples. They are a group to watch.

5. Step #2: Seek funding

From what we found, blockchain is not widely-used as a funding mechanism for scientific research… *yet*. I assume the main reason is that once money comes into play, so does legislation and regulatory policies. There were some initial players that emerged in 2017–2018, but they seem to have either disappeared or have changed their business objectives. That said, there are a few examples of blockchain funding mechanisms and even stronger examples of the administrative use of blockchain to manage the funding/grant processes more efficiently and cost-effectively.

One of the new funding organizations is Molecule [5] - a collaborative platform that connects stakeholders primarily in drug development. Their main focus is allowing researchers to license their Intellectual Property to interested investors using IP-Non-fungible tokens. In August 2022 they announced that, for the first time, they would be funding drug discovery research based in the United States at the University of Washington in Seattle.

It should be noted that uses of blockchain technology on the *administrative* side of funding are beginning to emerge. The U.S. department of Health and Human Services (HHS) is the largest grant funding agency in the United States. They successfully used blockchain technology to reduce the time it took to find the best deal for purchases of equipment, clinical tools, etc. from four to five months to finding it in real time. In February 2020 it resulted in a contract that will save the government $30 million U.S. dollars over a five-year period. With that successful blockchain implementation, HHS almost immediately initiated a pilot, the Grant-recipient Digital Dossier (GDD), to manage their grant program more efficiently. As of July 2021, GDD had reduced the time required to complete grant assessment tasks from four-plus-hours to a fifteen-minute process [6].

6. Step #3: Perform the experiment/make observations

It is in this area of the scientific research workflow that the inherent features of blockchain technology are invaluable. The following information can be captured in this step and if posted to a blockchain can make the research transparent and reproducible. (1) *Who did the experiment?* The digital identity and signature of the researcher(s) can be hashed to the blockchain while maintaining the privacy of the individual(s) involved. In addition, the expertise credentials of the experimenters can be hashed as there are organizations who now use blockchain for the creation of micro-credentials. As a real-life example

an organization called Hashed Health was used to ensure that the doctors/nurses who moved from one geographic location to another during the peak of the Covid-19 pandemic had the necessary experience and expertise.

(2) *When was the experiment performed?* As I mentioned earlier, blockchain timestamping is a key feature of the technology.

(3) *What materials/methodology were used?* Details on the exact reagents/animals/people and methods that were used can be compiled and linked via a hash posted to the blockchain to ensure that the provenance is captured. This is a step that facilitates the reproducibility and transparency of scientific research because the information is immutable.

(4) *What equipment was used?* Details on the manufacturer, when the equipment was obtained, its maintenance, prior uses and users, the location of the experiment, etc. can be hashed to the blockchain. Going back to the micro-credentials mentioned above, details on the training/expertise of those using the equipment can also be added to provide a level of confidence in the results that are reported.

(5) *What data was gathered*? The raw data gathered during the experiment is a digital asset that can be stored, shared, verified, etc. and this is a common use of blockchain technology. Once the data is hashed to the chain it cannot be tampered with. The immutability of a blockchain adds a layer of trust to the data.

As of this writing there are several organizations who offer services in support of the experimental phase of the scientific workflow. ARTiFACTS, which I already mentioned is one of them. Another is the Open Science Chain. Launched in 2018, this is a consortium blockchain that is funded by the U.S National Science Foundation (Award: 1840218 [7]) and is based at the San Diego Supercomputer Center at the University of California San Diego. Its main objective is to support data sharing to enable independent verification of scientific data and to foster reuse for the advancement of science [8]. Users can search, view, and validate scientific datasets, including lineage information and create "research workflows" linking one or more entries in the OSC blockchain and other repositories such as GitHub, thereby documenting an auditable record of the data workflow process behind the research hypothesis. The project has been very successful and, in the summer of 2021, received an additional half million dollars from the NSF for expansion. The service is free to members of the academic and research community.

7. Step #4: Perform analysis/make insights

The next step in the scientific workflow process involves taking the results from the experiment performed in Step #3 and performing one or more calculations, analyses, and visualizations on that data in order to generate conclusions, make insights, and to guide the next stage of the experimental, discovery cycle. In many respects step #4 is very similar to step #3. It executes an operation or an algorithm on experimental data that was generated *in vitro* or *in vivo* in step 3. In other words, step #4 takes data and uses computational processes to generate more results, again in the form of data, which support, or not, the experimenter's initial hypothesis.

In step #4, we have to all intents and purposes, the same "actors and artefacts" as we had in step #3, namely: (i) the person who ran the analysis (the "who"); (ii) the methods they used (the "how"), which in this case comprises the software, algorithms, and mathematics they used; and (iii) the results, observations, and conclusions that were generated. There is a new Electronic Laboratory Notebook, Labii [9], that uses blockchain to secure research data, and in 2021 the National Institute of Standards and Technology (NIST) funded a research project for the development of a prototype Digital Research Notebook system called KnowLedger. The concept is to build a system where research data is captured semantically immediately

at birth and stored on the blockchain, incrementally as knowledge is added. This is the brainchild of one of the authors of the blockchain white paper, Stuart Chalk, Professor of Chemical Informatics, University of North Florida, USA.

8. Step #5: Publish/share

The final step of the research workflow is that of sharing and or publishing research results. ARTiFACTS is a mainstream player in this space, providing services for publishers as well as authors. As an example, the *Journal of the British Blockchain Association* has collaborated with ARTiFACTS on the world's first blockchain application specifically designed to enable researchers to create a permanent and immutable public record of research material in real time. Another publisher, Partners in Digital Health, had also joined ARTiFACTS to register each original research article's provenance on the blockchain for every author to create a permanent and immutable public record of their work for the scientific community. The Bloxberg Consortium mentioned earlier also supports this step as does Orvium, currently in collaboration with the European Organization for Nuclear Research (CERN) [10].

9. Other applications

Blockchain technology is also being used for related activities in research and science education, among them being those listed below:

- Data sharing within closed networks/communities.
- Near real-time research tracking (e.g., Covid-19 tracking is done via blockchain by the U.S. Department of Health and Human Services; the same can be applied to any research having global priorities).
- Authentication of data at the source for auditing, compliance, and regulatory purposes.
- Equipment management (maintenance records, training, usage history, ownership history, etc.) - can even ensure compliance with the Good Manufacturing Practices (GMP) required by regulatory agencies.
- Supply-chain tracking for both tangible/physical and digital/data assets.
- Decentralized publishing/peer review.
- Research Funding/tracking.
- Crowd-sourcing/collaborative research (both public and within companies) where blockchain helps to confirm who owns what because it can demonstrate who did what.
- Degree and qualification certification.
- Identity certification of people and objects.
- Research data re-use (including sales of such data).
- Retrieval of chemicals/pharmaceuticals at the end of their life cycles or during their life cycle for re-purposing.

While the focus of our research was blockchain technology usage in science, our journey exposed us to its usage across many industries - finance, law, environment, global trade and commerce, insurance, real estate, media and entertainment, supply chain management, etc. A recent report from the Davos 2023 World Economic Forum states that "blockchain technology offers more promises than problems and that as a technology it will continue to grow exponentially, and its use cases expand. The real-world applications of blockchain, many already in use by organizations focused on international development, offer greater utility and cost savings" [11].

10. Bloxberg Consortium

One of the major examples of how far blockchain has become utilized in the global scientific community is the Bloxberg Consortium [12] that is under the umbrella of the Germany-based Max Planck Society. In response to requests from researchers working at its member institutes for time stamping of their research, the Max Planck Digital library built their own blockchain and launched the consortium in February 2019. They began with eleven members and there are now more than fifty members worldwide despite the pandemic - it is emerging as the global blockchain for science and aims to foster collaboration across the global scientific community. Organizations who do not want to build their own platform can utilize Bloxberg's if they sign the Manifesto and abide by the governance rules. Commercial entities can also join the consortium, but they cannot serve as authority nodes. In May of 2022 Bloxberg members voted to approve what is called the "Bloxberg Association for the Advancement of Blockchain in Science" that will be founded under German legislation this year. They also agreed to introduce tokenomics to ensure sustainable funding of the infrastructure [13]. Keep your eyes on them!

11. Failed projects

During the course of the many interviews that we held with active players in this space, we found that blockchain technology is not always the perfect solution for a specific problem. Two notable examples are the Blockchain for Peer Review Project initiated by Digital Science that failed supposedly because a solution to the problem already existed in another form. The other was a project that was being conducted by a group within Elsevier's Healthcare Education Group that failed because users of the product that they were designing wanted to use it for clinical registry and one of the key attributes of blockchain - its immutability - ran afoul of the laws around "the right to forget and the right to be forgotten" which vary from country-to-country. The white paper will cover these in detail.

12. Lessons learned

Indeed, the people we interviewed discussed many of the lessons that they learned in their use of blockchain technology. Key are the following:

- Blockchain Is not the perfect solution for every problem.
- Potential users need to understand what they want to accomplish.
- Administrative use cases can really benefit from blockchain technology.
- Its use requires learning and thinking differently so that you can see how it fits in – it needs a champion!
- Blockchain applications need to be seamless and embedded in the systems that researchers already are using.
- Use of blockchain requires some learning and a new mindset

When we interviewed Dr. Naseem Naqvi, Editor-in-Chief for *the Journal of the British Blockchain Association*, he said that he believes in applying an evidence assessment framework when considering the potential use of blockchain technology and tells people to take what he calls the "PICO" approach: P - what is the problem; I - what is the new intervention; C- how does it compare to the existing intervention; and O - what are the outcomes and the evidence that it is better and that it works? This is explained in detail in his article on evidence-based blockchain [14]. Also, it should be emphasized that legal and regulatory

issues must be considered, especially when a blockchain involves tokenization. The white paper includes a section on this topic, not to supply legal advice (none of the authors are lawyers), but to make readers aware of the issues that must be considered.

13. The future

What we learned from our interviews is that usage of the technology is growing in general and is accelerating in the life and health sciences. We found that all parts of the research cycle *can* take place within a blockchain system and we expect new developments to emerge. Paper at regular intervals as the technology is enhanced and new use cases and new players emerge. In fact, a section of the white paper covers in detail the evolution of the technology. We are in the third stage of its development and technologists are already discussing potential fourth generation changes along with the impact of quantum computing.

While the scope of blockchain technology's impact is hard to predict since it's in the early stages of broad adoption, we believe that its future looks bright. It was selected as one of the Top Ten Emerging Technologies in Chemistry in 2021by IUPAC and a recent report from the European Chemical Industry Council, "Artificial Intelligence and Blockchain: Insights and Actions for the Chemical Industry", boldly states that blockchain technology holds the potential for disruption across the chemical enterprise [15]. Only the future will tell. Our goal is to update the white paper at regular intervals as the technology is enhanced and new use cases and new players emerge.

14. White paper authors

In closing, I want to recognize my fellow co-authors of the white paper who have worked tirelessly over the past three years to make the white4 paper a reality. Because of them our three year journey has been a pleasurable adventure:

Stuart Chalk, Professor of Chemical Informatics, University of North Florida

Jeremy Frey, Professor of Physical Chemistry, University of Southampton

Kazuhiro Hayashi, Director of Research Unit for Data Application, Japanese National Institute of Science and Technology Policy

David Kochalko, Co-founder, ARTiFACTS

Richard Shute, Research Consultant, Curlew Research

Mirek Sopek, Chief Technology Officer, MakoLab SA.

About the author

Bonnie Lawlor is retired from the scientific publishing industry where her prior positions include Executive Vice President, Database Publishing at the Institute for Scientific Information (now Clarivate), Senior Vice President and General Manager of ProQuest's Library Division, and Executive Director of the National Federation of Advanced Information Services (NFAIS). Bonnie is active in IUPAC serving as the Vice Chair of the U.S. National Committee for IUPAC and in other volunteer positions. She is also an ACS Fellow as well as an NFAIS Honorary Fellow. Bonnie earned a B.S. in Chemistry from Chestnut Hill College (Philadelphia), an M.S. in chemistry from St. Joseph's University (Philadelphia), and an MBA from the Wharton School, (University of Pennsylvania). E-mail: chescot@aol.com.

References

[1] D. Yaga, P. Mell, N. Roby and K. Scarfone, *NISTIR 8202 Blockchain Technology Overview* National Institute of Standards and Technology, U. S. Department of Commerce, 2018, https://nvlpubs.nist.gov/nistpubs/ir/2018/nist.ir.8202.pdf, accessed September 11, 2023.

[2] See: https://nakamotoinstitute.org/bitcoin/, accessed September 11, 2023.

[3] S. Haber and W.S. Stornetta, How to time-stamp a digital document, *The Journal of Cryptology* **3**(2) (1991), 99–111, Available at: https://link.springer.com/article/10.1007/BF00196791, accessed September 11, 2023.

[4] See: https://artifacts.ai/verify/, accessed September 11, 2023.

[5] See: https://www.molecule.to, accessed September 11, 2023.

[6] See: https://governmentciomedia.com/how-blockchain-transforming-federal-grants-management, accessed September 11, 2023.

[7] See: https://www.nsf.gov/awardsearch/showAward?AWD_ID=1840218, accessed September 19, 2021.

[8] See: https://opensciencechain.org, accessed September 11, 2023.

[9] See: https://www.labii.com., accessed September 11, 2023.

[10] See: https://orvium.io, accessed September 11, 2023.

[11] A. Nath, In Davos, Blockchain Yields More Promises Than Problems, https://www.coindesk.com/consensus-magazine/2023/01/18/in-davos-blockchain-yields-more-promises-than-problems, accessed September 11, 2023.

[12] See: https://bloxberg.org, accessed September 11, 2023.

[13] See: MPDL, Fourth Bloxberg Summit, https://www.mpdl.mpg.de/ueber/presse/812-fourth-bloxberg-summit-2022.html, accessed September 11, 2023.

[14] N. Naqvi, *The Journal of The British Blockchain Association* **3**: (2020), 1–13. doi:10.31585/jbba-3-2-(8)2020, accessed September 11, 2023.

[15] Accenture, AI & Blockchain: Chemical industry insights and actions, https://www.accenture.com/us-en/insights/chemicals/ai-blockchain-chemical-industry, accessed September 11, 2023.