International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

Blockchain For Intellectual Property Management In Biological Data: A Review

Harish Sharma
Research Scholar, Department of Allied & Health Science, North East Christian University
Dr. Prakash Mathew
Professor, Department of Allied & Health Science, North East Christian University

Abstract

High-throughput technology continues to rapidly advance the life sciences. As a result of these advances, large amounts of information have been generated as a result of this development. Examples include genomics (the complete set of DNA), transcriptomics (the a study of RNA), proteomics (the complete set of proteins) and other omics datasets. These datasets provide the basis for scientific discovery, bio-medical innovation and agriculture and biotechnology. These developments have resulted in the proliferation of large amounts of data for open sharing and collaboration between researchers, and increased the pace of research and innovation. However, with these advancements, major issues have been raised around the Intellectual Property Rights, including, but not limited to, ownership of datasets, authorship, provenance, misuse (unauthorised usage), and enforcement.

Intellectual Property Rights have been governed by the existing traditions of intellectual property systems, including patent system (exclusive rights to inventors), copyright (reproduction and distribution rights), trade secret (business operational knowledge) and databases rights (databases rights including the creators and users). These traditional systems were developed for static, centralised means of innovation, and have proven to be increasingly inadequate for managing biological data in a dynamic, distributed and digital environment.

Blockchain technology is one of the most promising technologies that may be used to address these IP challenges through its ability to provide decentralized, immutable, transparent, cryptographically secure and automatable smart contracts. This review analyzes the current literature on biological data management, intellectual property frameworks, and blockchain technology, to evaluate how blockchain technology can be utilized within the biological data ecosystem to create a secure and ethical approach to managing Intellectual Property. The current review synthesizes research in the areas of: Data Provenance Tracking; Ownership Verification; Secure Data Sharing; Licensing; Consent Management; and Benefit Sharing through Blockchain Systems. This study further highlights the legal uncertainties; limitations and areas for additional research on blockchain technologies within Bioinformatics.

Keywords : Blockchain Technology; Biological Data; Bioinformatics; Intellectual Property Rights; Data Provenance; Smart Contracts; Ethical Data Governance

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

1. Introduction

The advances made in high-throughput experimental and computational technology over the last few decades have transformed nearly every aspect of life science. Technologies such as next-generation sequencing, microarrays, mass spectrometry, and multi-omic platforms allow researchers to produce enormous amounts of biological data, quickly and accurately (Stephens et al. 2015). As a result, biological research has become largely driven by data, collaboration, and computational modeling, thus providing researchers with new opportunities as well as complications regarding research governance.

Biological data are no longer merely considered research output; they have become an increasingly significant component of the scientific community's intellectual assets with the potential to offer economic, academic, and societal benefits to those who create them. Genomic and transcriptomic datasets are particularly important to the discovery of new drugs, the development of personalized medicine, the enhancement of agricultural productivity, and the emergence of the field of systems biology (Edgar et al. 2002; Marx 2013). With public resources such as GenBank, GEO, ArrayExpress, and EMBL-EBI, there has been a growing number of accessible datasets, which allow for greater transparency and reproducibility in the scientific process. However, as more biological data is shared between researchers, there have been increasing concerns about ownership and attribution, unauthorized reuse of data, and the use of shared data for commercial gain.

Current legal frameworks for intellectual property rights (IPR) are unable to help researchers with such problems. Many raw biological materials do not meet the requirements for obtaining patents or copyrights, while trade secrets prevent researchers from sharing their work openly (Reichman et al., 2016). In addition, biological data sets are often created through international partnerships between multiple institutions, making it difficult to assign proper ownership, authorship, and inventorship to any individual or organization (Shabani & Marelli, 2019). Finally, centralized data systems create additional problems because they create single points of failure, a lack of transparency, and increased risk for unauthorized access and/or manipulation of data.

In light of these challenges, there has been growing interest in using blockchain technology to solve some of these data management and intellectual property issues. Blockchain is a type of distributed ledger technology that allows for digital record keeping without relying on one central entity (Nakamoto, 2008; Zheng et al, 2018). The basic attributes of blockchain—its immutability, ability to time stamp transactions, reach consensus among multiple users, and its inclusion of smart contracts—open up many new options for managing ownership, origin, licensing, and consent in complex, decentralized data networks.

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

Currently, a large body of literature provides support for using blockchain to facilitate the management of intellectual property related to biological data. The technology allows users to register data sets in a secure manner, provide verifiable evidence of the creation of the data set, track where the data is used, and automatically enforce licensing and benefit-sharing agreements via a smart contract mechanism (Yli-Huumo et al., 2016; Tapscott & Tapscott, 2016).

To address concerns regarding privacy and scalability, an emerging architecture has been developed that combines the blockchain with off-chain data storage. This architecture operates by storing the metadata and cryptographic hash of sensitive biological data on the blockchain, while keeping sensitive biological information off-chain (Zyskind et al., 2015).

Despite the interest in using blockchain to manage intellectual property issues for biological data, this topic continues to be an active research area. Most of the current work is in a conceptual format, with very little work synthesising interdisciplinary literature related to bioinformatics, intellectual property law and blockchain technology. This review is intended to fill this knowledge gap by conducting a critical review of the existing literature related to biological data governance, the challenges associated with biological data and intellectual property, and the potential benefits of blockchain-based solutions. This review will provide an overview of the current state of research and identify the major challenges and possible future developments in the use of blockchain technology for managing intellectual property related to biological data.

2. Biological Data and Contemporary Data Sharing Ecosystems

The advancement of experimental and computational technologies has fundamentally reshaped the generation, storage, and utilization of biological data. Modern biological research produces vast and diverse datasets that capture information at molecular, cellular, organismal, and population levels. These data form the backbone of bioinformatics, systems biology, precision medicine, and agricultural biotechnology, making effective data governance a critical requirement for scientific and societal progress.

2.1 Types and Characteristics of Biological Data

Biological datasets encompass an extremely diverse range of digital formats; genomic sequences, transcriptomic expression profiles, proteomic datasets, and metabolomic datasets, epigenetic markers, phenotypic observations, and clinical records. The emergence of next generation sequencing (NGS) and mass spectrometry (MS) platforms has created a dramatic increase in the scale of biological datasets, in many cases exceeding the capacity for management by traditional data management systems (Marx, 2013; Stephens et al., 2015).

Biological datasets are generally characterized as "big data", meaning that they have high volume, high velocity, high variety, and high complexity. Furthermore, because of the nature of biological datasets being noisy, heterogeneous, and context-dependent, considerable curation and annotation must be performed to make the datasets usable and reproducible (Afendi et al., 2013).

As a result, not only do these characteristics create technical challenges for managing biological datasets, they also present additional legal and ethical governance challenges, particularly regarding ownership and intellectual property rights.

2.2 Role of Bioinformatics Repositories and Databases

Publicly available curated biological datasets have been created using GenBank and other biological data repositories, including the Gene Expression Omnibus (GEO), ArrayExpress and EMBO-EBI, which allow researchers to provide a collaborative framework for conducting research. Preserving and reusing curated biological datasets promotes increased transparency and reproducibility of research, thereby strengthening the integrity of science (Edgar et al., 2002).

Although the availability of publicly accessible curated biological datasets allows an open sharing of data with the scientific community, most studies that have been published to date have indicated that there are limited means available to track how the curated biological datasets are being used after they are deposited into the repositories; enforce proper attribution for the use of data; and effectively manage licensing conditions related to the deposits that have been made into the repositories. After the files have been deposited into the repositories and are made available to users, the control over the subsequent use of these files by a user has been significantly relinquished and as such poses an increased risk for the unauthorized use of or commercial exploitation of the files by third parties who are unaffiliated with the original depositor (Reichman et al., 2016). When using data that have been deposited into the repositories, researchers have to manage their intellectual property and therefore lack any means of providing a proper attribution with respect to the deposits.

2.3 Data Sharing Mandates and Collaborative Research

Sharing data is now an essential aspect of contemporary research. Many funding groups, journals, and international collaborations now require researchers to share their data so the general public can access results generated from publically funded research efforts. The open science movement supports this idea by promoting the development of reproducible research through the provision of shared materials, and reducing instances of duplicated efforts plus accelerating the development of new ideas through improved communication, collaboration, and sharing of data (see Piwowar et al. 2007).

Despite this, the literature documents that many researchers are unwilling to share their high-value datasets with others. Researchers have expressed fear over misuse of the data, loss of competitive advantage, and lack of proper citation and recognition as reasons for being reluctant to share (see Tenopir et al. 2011). The ability to share datasets or access to shared datasets remains limited due to the existence of different countries with disparate legal systems and policies that affect how intellectual property agreements can be enforced.

2.4 Intellectual Property Challenges in Data Sharing Environments

Traditional Intellectual Property protection methods can be at odds with sharing environments on Biological Databases, because they are collaborative and open, and can be re-used, combined, and/or re-analysed many times, preventing clear definitions for ownership or establishing enforcement of use restrictions. Centralised models of Data Governance depend upon the belief in, and reliance upon, trust towards the manager of the repository and more importantly, the rules and regulations of the university or institution itself, as well as being the main avenue to resolving issues of mis-usage and disputes (Shabani & Marelli, 2019).

There is growing consensus in literature that transparency, traceability, and accountability are necessary for successfully managing Intellectual Property in a Biological Databases Ecosystem without compromising the ideals of openness and collaboration. Furthermore, the need for these attributes within Biological Databases Ecosystem is motivating the investigation of using Blockchain as a potential mechanism for managing Biological Databases and the associated Intellectual Property.

3. Intellectual Property Rights Frameworks in Biological Data

The protection of innovation through Intellectual Property Rights (IPR) fosters the promotion and growth of research as well as allowing scientific results to be marketed. For biological and Bioinformatics based research, IPR frameworks are created with a focus on providing a fair balance between the rights of data generators, institutions, funding bodies and society. However, the nature of Biological Data presents challenges with regard to applying traditional Intellectual Property Rights systems to it. In addition to the extensive literature demonstrating the shortcomings of traditional intellectual property mechanisms when applied to Biological Data in a collaborative, data, driven research environment.

3.1 Patents and Biological Innovations

In relation to Biological Research, one of the ways Inventors protect their Invention is via Patent law. Patents may protect Invention that arise from Biological Research, including inventions relating to novel biological processes, genetically engineered organisms, Diagnostic Methods & Treatment uses that derive from Biological Data. However, the existing literature has acknowledged that Raw Biological Data, such as DNA sequences, & Expression Profiles, are generally ineligible for patent protection because they are regarded as Discoveries (rather than Inventions) (Bently & Sherman, 2014), as well as being expensive, time-consuming & Geographically Limited, making them unsuitable for protecting an ever-expanding database of Biological Data that continues to evolve, & be updated. Additionally, the current Studies highlight the ethical and legal issues that exist with regards to Gene Patents & Natural Occurring Biological Material, which make the idea of a Patent system an even bigger challenge in Biological Data-driven research. (Rai & Boyle, 2007).

3.2 Copyright and Database Rights

Copyright protects the original expression in various forms, such as: (1) software; (2) curated databases; and (2) analytical representations of biological data via copyright laws. As per copyright law, copyright is not a protection for the underlying factual data represented by the creators of the databases, even when copyright protects the structure, selection, and arrangement of databases (Reichman et al., 2016). As a result of this distinction, data that have been deposited in public repositories remain vulnerable to unrestricted reuse once someone's intellectual property has been illegally removed (using unauthorized access to the data) and deposited into a public repository.

Sui-generis laws or regulations, relating to database rights, may provide additional protection in certain jurisdictions for large investment (and/or) large amounts of investment) made into the creation of a database, and therefore, may offer a greater degree of protection for those making a significant investment in creating a database. However, the literature indicates that such laws or regulations are geographically limited to only certain jurisdictions, meaning they do not apply to the entire world. Furthermore, database rights do not address issues related to provenance tracking or attributions involved in downstream ("data") reuse.

3.3 Trade Secrets and Confidential Data

Trade secrets are frequently used to protect proprietary biological data in industrial and commercial research settings. This form of protection relies on maintaining confidentiality and restricting access to authorized parties. While effective in closed environments, trade secret protection is fundamentally incompatible with open science principles and collaborative research models that emphasize data sharing and transparency (Rai & Boyle, 2007).

The literature also notes that reliance on trade secrets can slow scientific progress by limiting data accessibility and reproducibility. In academic research, excessive use of confidentiality measures may conflict with funding agency requirements and ethical expectations of openness.

3.4 Ownership, Authorship, and Inventorship Challenges

Enhanced scientific research and development; however, trade secret protections may inhibit scientific advancement due to the lack of access to data in most cases, restricts real-time data availability or verifiability by researchers of all types.

Non-technological the characteristics of collaborations can further complicate the determination of the specific nature of the relationship between authorship and ownership of the underlying data and information. The ambiguity often results in disagreements or conflicts between the various parties involved in creating the data and their rights to use, re-use, or exploit the data commercially. As long as there remain no viable, reliable means for accurately and efficiently tracking how data was created and contributed or determining the identity of each contributor, a significant lack of confidence in the authorship and ownership of data will persist.

3.5 Enforcement and Governance Limitations

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

Enforcing intellectual property rights for biological data continues to be challenging because traditional methods are largely reactionary requiring companies to rely upon litigation that are fundamentally expensive, slow and ineffective at dealing with the misuse of data across borders. The inability to effectively enforce biological data intellectual property rights is further complicated by the lack of the ability of centralized data governance models to monitor how biological data is used after the data is distributed (Reichman et al. 2016).

Current developments in the literature on these issues are increasingly highlighting the need for additional technological solutions that are capable of real-time tracking, automatic enforcement and transparent governance. As such, there continues to be an increase in interest about using blockchain technology to create a more effective and efficient infrastructure for managing intellectual property rights related to biological data ecosystems.

4. Blockchain Technology for Intellectual Property and Data Governance

Nearly every day new and innovative digital infrastructures are emerging as a result of blockchain technology. The improvements that have been made to the way biological data is governed and managed will impact the future of data ownership by society and individuals. Blockchain technology, which was originally conceived as the technological backbone of cryptocurrencies, is now being explored in many different sectors and as such has become more than just a financial technology. Blockchain is now also being considered as a solution for the protection of digital assets, and as an effective tool for managing digital assets, especially in decentralized and globally collaborative environments.

4.1 Fundamental Concepts of Blockchain Technology

Blockchain is a method of recording information that allows for multiple computers to securely verify that something has taken place and no one has changed that information after it has been recorded. It consists of "blocks," which are timestamps, cryptographic hash values of the previous blocks, and detailed records of all transactions that occurred during that period of time. Each time a transaction occurs on the blockchain, it must be verified by all computers before it can be added to the chain (Nakamoto, 2008). All transactions on the blockchain have received approval from everyone involved in the transaction, which means there are no "middle men" involved. Therefore, when a transaction is placed on the blockchain, all parties know it is valid and secure.

Because of the decentralized nature of blockchain technology, it increases both accountability and transparency, which allows people to work in a more trustworthy fashion without having to rely on a single company as an intermediary to verify whether or not the transaction has happened. This characteristic makes it an excellent tool for establishing trust between people engaging in these types of transactions (Zheng et al., 2018).

4.2 Types of Blockchain Networks

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

Blockchain networks can be broadly classified into public, private, and consortium (permissioned) blockchains. Public blockchains allow open participation and offer high transparency but may face scalability and privacy challenges. Private blockchains are controlled by a single organization and provide greater efficiency and privacy but sacrifice decentralization. Consortium blockchains represent a hybrid model in which a group of trusted institutions collectively govern the network, balancing transparency, performance, and access control (Yli-Huumo et al., 2016).

The literature suggests that consortium blockchains are particularly well suited for research and data-sharing environments involving universities, research institutes, funding agencies, and industry partners.

4.3 Smart Contracts and Automated Governance

Public, private, and consortium (permissioned) blockchain networks fall into three primary categories. Publicly available blockchain networks provide open access to anyone interested in joining, while also providing significant technical capabilities for public transparency purposes. However, they are also subject to issues related to decentralization (the ability of all participants to have equal rights to participate) along with usability and confidentiality limitations due to the size and scope of these networks.

In contrast, private blockchain networks only allow users who have been invited or approved to join. Many of these private networks operate much like public-facing blockchains, with high degrees of technical capability for increasing user efficiency and maintaining heightened confidentiality, but they do trade away some elements of decentralized structure.

4.4 Blockchain and Intellectual Property Management

Consortium blockchain networks can be seen as a hybrid form, with multiple people and organizations serving as administrators and governing bodies.The following author suggests that consortium blockchain models may be ideally suited to those environments, which will naturally include research data sharing, and that universities, funding agencies, and private-sector partners can all benefit from the availability of blockchain technology today.Blockchain-based IP registries have been proposed as complementary systems to traditional legal frameworks, offering technological enforcement and transparency while preserving the role of existing laws and institutions. Such systems are particularly relevant for biological data, where ownership is often contested and reuse is widespread.

4.5 Privacy, Security, and Regulatory Considerations

While creating many advantages through increased security and transparency, Blockchain technology is recognized by many researchers as having several challenges; Privacy, Scalability, and Regulatory Compliance. It is generally considered that storing sensitive clinical/biological information (i.e., any data that may have legal, regulatory, or safety implications) directly on a

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

Blockchain is not recommended because once data has been recorded on the Blockchain, it cannot be deleted and this clash with current Data Protection Policies. Therefore, most scholars recommend using Hybrid Blockchain Solutions that utilize both on and off Chain technologies to store metadata, hashes, and Access Logs while still providing secure storage of off Chain via encrypted storage or Cloud-based security solutions.

The scholarly community is also developing formalized data management systems that meet both Data Protection and Privacy laws, and ethical standards, using the Blockchain as a foundation for these systems.

5. Blockchain Applications for Intellectual Property Management in Biological Data

Blockchain technology has recently been gaining traction in how we manage biological data, especially with the difficulties surrounding data sharing, intellectual property and compliance to ethical concerns that many researchers have experienced. Biological data is created and donated by many different types of people; researchers, universities/colleges, funding sources, repositories (that archive all biological material), companies that use such materials (including pharmaceutical and biotech industries).

5.1 Provenance Tracking and Ownership Registration

The scientific literature has widely acknowledged that provenance tracking is an absolute necessity for the proper management of intellectual property rights in the context of biological data. By leveraging the capabilities of the Blockchain, researchers can now provide hard evidence of their authorship/provenance claims for a given biological dataset without revealing the actual data by registering a unique hash on the Blockchain at the time of dataset creation/submission and cryptographically hashing each event related to that dataset (e.g., creation, ownership, and modification).

The literature shows that because of its decentralization, transparency and lack of immutability, blockchain technology appears to be an excellent choice for resolving the issues surrounding ownership, provenance, access control and enforceability of data in biological data ecosystems.

5.2 Secure Data Sharing and Access Control

Using a blockchain, data owners have the ability to create permission-based systems that dictate when and how biological data can be accessed and reused. Smart contracts provide the mechanism for enabling easy changes to access permissions, depending on a set of rules. These systems are particularly important in environments where a sizeable number of datasets may contain sensitive biological data, such as through genomic or clinical repositories.

Numerous studies have reported on the value of hybrid on-chain and off-chain architectures for storing sensitive biological data; this includes securely storing the data off-chain while only keeping access logs, cryptographic references, and authorizations on the blockchain. Hybrid architecture balances transparency, security, scalability, and compliance with regulatory requirements.

5.3 Licensing, Monetization, and Benefit-Sharing

Smart contracts simplify the licensing and benefit-sharing processes associated with biological data. Blockchain-based licensing provides creators of biological data with the ability to define how the biological data can be used (for academic purposes only) as well as the distribution method for royalties. After being encoded within smart contracts, the terms established through these methods can be enforced automatically, eliminating the need for administrative review and reducing the potential for disputes.

The literature indicates that automated benefit sharing approaches to economic compensation related to public funding of research and/or collaborative research projects involving multiple institutions is helpful in these situations as it helps promote transparency and therefore will provide equitable distribution of royalties, as well as incentivize the generation of high-quality data while also upholding ethical principles.

5.4 Consent Management and Ethical Governance

In the area of biological data research involving human subjects, consent management is an area of concern. Most traditional consent management approaches are static, and therefore are difficult to maintain over time, creating ethical and legal considerations. Blockchain-based consent management systems are capable of providing dynamic consent management capabilities, allowing for consent preferences to be recorded, changed, and enforced in a transparent and auditable way.

Researchers have noted that the creation of immutable consent records increases the level of accountability for researchers and contributors to biological data and allows for improved compliance with ethical guidelines and governmental regulations. In addition, researchers can utilize smart contracts to ensure that all access to and usage of data is limited to only those purposes consented to by the contributors, enhancing the level of trust between the researchers and the contributors.

5.5 Interoperability with Bioinformatics Infrastructure

To be effective, blockchain designed for use with Intellectual Property (IP) must have smooth integration across "Bioinformatics" tools (aka bioinformatics databases, applications, etc.), workflows & repositories. Interoperability is recognized by researchers as a major barrier to integrating blockchain & the current bioinformatic environment, with a strong need for standardise data Metadata schemas, APIs and Governance frameworks.

6. Challenges, Limitations, and Research Gaps

'A consortium model' has been presented as an architectural structure for allowing multiple repositories and institutions to integrate with each other while creating a basis for shared trust

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

and governance. The consortium model now neatly aligns with the collaborative methods used in biological research and allows for IP management across multiple institutions. Despite the need for blockchain in biological IP and increasing interest, many technical, legal, ethical and practical barriers remain to its widespread adoption, the record of Immutable, Transparent and Decentralized properties, along with the outstanding areas of Blockchain that need to be addressed, hinder its full integration into ecosystems around biological data.

6.1 Technical Challenges and Scalability Issues

One of the most frequently cited limitations of blockchain technology is scalability. Public blockchains, in particular, face performance constraints related to transaction throughput, latency, and storage capacity. Biological data ecosystems generate a high volume of transactions, including data access requests, licensing events, and consent updates, which may exceed the capabilities of current blockchain infrastructures.

Additionally, blockchain systems are not designed to store large datasets efficiently. As a result, most proposed frameworks rely on hybrid on-chain/off-chain architectures. While this approach mitigates storage limitations, it introduces additional complexity in ensuring secure and reliable integration between blockchain networks and off-chain repositories. The literature highlights the need for robust interoperability mechanisms and standardized interfaces to support seamless data exchange.

6.2 Privacy, Security, and Data Protection Concerns

The concept of a 'Consortium Model' has emerged over the last several years as an Architectural Framework that will help facilitate the connections between multiple repositories and institutions while establishing a foundation of Collaborative Trust and Governance for Collaborative Biomedical Research. As such, 'Consortium Model' aligns with the current Collaborative approaches employed in Biological Research and facilitates Intellectual Property management among several institutions.

As interest in utilizing blockchain technology for Biological Intellectual Property continues to rise; there exist many significant Technical, Legal, Ethical and Practical challenges that have prevented the widespread use of such technology. Furthermore, while the Unique Properties of Blockchain include their Immutable, Transparent and Decentralised nature; many Grey Areas of Blockchain remain to be Addressed before Blockchain Technology can be fully integrated into the broader Ecosystems surrounding Biological Data.

6.3 Legal and Regulatory Uncertainty

Many legal jurisdictions have not determined how they will treat smart contracts or blockchain records as evidence. Though blockchain technology may prove that the data was created and

International Journal of Engineering, Science and Humanities
An international peer reviewed, refereed, open-access journal
Impact Factor 7.9 www.ijesh.com ISSN: 2250-3552

owned by the proper party, the acceptance of blockchain as a legitimate form of evidence varies widely among legal systems.

The literature suggests that an Intellectual Property system based on blockchain technology does not supersede the existing legal systems or frameworks but rather supports them.

The issues related to cross-border collaborations in research simply amplify the challenges of complying with local regulations for IP laws and ethical guidelines, as they vary greatly from country to country. Without internationally accepted standards of law for the governance of data on blockchain technology, developing a widespread blockchain-based electronic marketplace will not be successful.

6.4 Governance and Standardization Challenges

Successful use of blockchain to manage IP requires well-defined governance including role definition, responsibility assignment and decision making processes in regard to blockchain IP management. Governance has been highlighted in the literature as an under-explored, yet extremely important, area of research when considering the implementation of blockchain within an academic research setting. The questions of who owns the blockchain, what rules govern a consensus, how disputes are settled, and how blockchain users will be accountable to each other remain largely unanswered.

Additionally, standardization is another major hurdle for the success of blockchain technology. Unless there is agreement among participants on metadata used within a blockchain, smart contract design and interoperability between multiple blockchain initiatives, blockchain-based systems will experience fragmentation and/or reduced useability. The literature calls for collaborative efforts between all stakeholders - researchers, organisations, government and technology developers - to develop widely accepted standards and establish best practice guidelines.

6.5 Limited Empirical Validation and Adoption

Research on blockchains and Intellectual Property (IP) has gone largely unevaluated from a research standpoint because there is not a lot of real-life evidence or test results for what works best regarding blockchain strategies for protecting biologically-derived data. There have been many papers published regarding concepts related to developing prototype systems; however, most of these papers do not contain extensive testing to determine if the proposed solutions actually perform as intended.

Additionally, as with most developing technologies, barriers to adoption exist, including insufficient technical knowledge, high infrastructure costs related to implementing blockchain-based solutions at a large scale and the reluctance of individuals and businesses to implement new technologies. Therefore, applied research needs to take place to fill the gaps that separate blockchain solutions from traditional storage and data management practices.

6.6 Identified Research Gaps

Based on the reviewed literature, several gaps remain:

- Absence of comprehensive, empirically validated blockchain-based frameworks for managing intellectual property in biological data
- Limited integration of legal, ethical, and technical perspectives
- Insufficient focus on interoperability with existing bioinformatics infrastructures
- Lack of standardized governance and policy models
- Minimal evaluation of user trust, acceptance, and long-term sustainability

These gaps highlight the necessity of systematic research aimed at designing, implementing, and evaluating blockchain-enabled intellectual property management frameworks tailored specifically to biological data ecosystems.

7. Conclusion and Future Research Directions

The rapid growth of biological information and bioinformatic signals has changed life sciences research to a more relational and data-intensive form of working. The increase in open access data-sharing along with the numerous opportunities for large-scaled collaborative approaches has progressed science at a faster pace but increased challenges concerning individual entitlements to intellectual property with regard to data rights and related issues such as the origins, traceability, attribution, ethical guidelines for data use, and enforcement of property rights for data.

The literature analysed in this review indicates that many existing forms of intellectual property do not adequately address the changing dynamics of biological data ecosystems, including the global, digital and fluid facets of such ecosystems.

Blockchain technology represents a potential pathway for addressing the obstacles related to intellectual property in the management of biological data by virtue of blockchain's main characteristic properties, which are: Decentralization, Immutability, Transparency, and Cryptographic Security.

The advantages of blockchain to biological data management systems include the development of innovative ways of tracking provenance; registering ownership; securely sharing information; automating the process of licensing; managing consent; and transparently distributing the benefits of sharing data. Hybrid on-chain and off-chain architecture combined with smart contracts are well-established and documented in the literature as ways to optimally balance security, privacy, scalability, and compliance with the law.

While the use of blockchain technology for managing IP for biological data has the potential to do so much more for people than just being a tool to help manage and facilitate the transfer of ownership of IP, there are many technical hurdles to overcome. The ability to scale, communicate with other systems, and protect user information are some of the biggest technical roadblocks to be addressed in order for blockchain to be successfully utilized for IP management in biological databases.

There is also a lack of clarity around the legal and regulatory landscape for blockchain records and smart contracts. This creates an obstacle for researchers working collaboratively across borders. There are also issues with respect to governance and standardization that are preventing blockchain technology from being implemented for IP management in biological databases on a large scale.

Going forward, more research is needed to develop standardized and consortia-based blockchain solutions that address the many layers of legal, ethical, and technical perspectives involved in the application of blockchain to IP management. Empirical studies using biological datasets will be critical to proving that blockchain technology can be applied successfully and, therefore, demonstrate the practicality of the proposed solutions.

In addition, as the blockchain industry continues to evolve, it is important to focus on establishing systems that will allow for integration with existing bioinformatics infrastructures as well as develop protocols that will meet the legal and regulatory expectations of the International IP and Data Protection community. Finally, more research needs to be conducted to understand how trust, acceptance, and sustainability will affect the ability of individuals and organizations to adopt blockchain technology for IP management.

It should be noted that while blockchain represents a tremendous opportunity to improve IP management within the biological data ecosystem through increased confidence and transparency, the realisation of this opportunity requires appropriate framework design, interdisciplinary cooperation and evidence-based evaluation to develop effective systems. Therefore, this review provides a solid starting point for future research as well as supporting the development of blockchain-supported intellectual property management (IPM) systems which will facilitate equitable data sharing among researchers, enhance the rights of creators and initiate scientific progress.

References

1. Afendi, F. M., Ono, N., Nakamura, Y., Nakamura, K., Darusman, L. K., Kibinge, N., … Kanaya, S. (2013). Data mining methods for omics and knowledge of biological systems. *Journal of Molecular Biology*, 425(21), 4053–4069. https://doi.org/10.1016/j.jmb.2013.07.010
2. Bently, L., & Sherman, B. (2014). *Intellectual property law* (4th ed.). Oxford University Press.
3. Edgar, R., Domrachev, M., & Lash, A. E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Research*, 30(1), 207–210. https://doi.org/10.1093/nar/30.1.207

4. Marx, V. (2013). Biology: The big challenges of big data. *Nature*, 498(7453), 255–260. https://doi.org/10.1038/498255a
5. Nakamoto, S. (2008). *Bitcoin: A peer-to-peer electronic cash system*. https://bitcoin.org/bitcoin.pdf
6. Piwowar, H. A., Day, R. S., & Fridsma, D. B. (2007). Sharing detailed research data is associated with increased citation rate. *PLoS ONE*, 2(3), e308. https://doi.org/10.1371/journal.pone.0000308
7. Rai, A. K., & Boyle, J. (2007). Synthetic biology: Caught between property rights, the public domain, and the commons. *PLoS Biology*, 5(3), e58. https://doi.org/10.1371/journal.pbio.0050058
8. Reichman, J. H., Okediji, R. L., & Thomas, J. R. (2016). When copyright law and science collide: Empowering digitally integrated research methods on a global scale. *Minnesota Law Review*, 96(4), 1362–1480.
9. Shabani, M., & Marelli, L. (2019). Re-identifiability of genomic data and the GDPR. *EMBO Reports*, 20(6), e48316. https://doi.org/10.15252/embr.201948316
10. Stephens, Z. D., Lee, S. Y., Faghri, F., Campbell, R. H., Zhai, C., Efron, M. J., … Robinson, G. E. (2015). Big data: Astronomical or genomical? *PLoS Biology*, 13(7), e1002195. https://doi.org/10.1371/journal.pbio.1002195
11. Tapscott, D., & Tapscott, A. (2016). *Blockchain revolution: How the technology behind bitcoin is changing money, business, and the world*. Penguin Random House.
12. Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., … Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS ONE*, 6(6), e21101. https://doi.org/10.1371/journal.pone.0021101
13. Yli-Huumo, J., Ko, D., Choi, S., Park, S., & Smolander, K. (2016). Where is current research on blockchain technology?—A systematic review. *PLoS ONE*, 11(10), e0163477. https://doi.org/10.1371/journal.pone.0163477
14. Zheng, Z., Xie, S., Dai, H. N., Chen, X., & Wang, H. (2018). Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services*, 14(4), 352–375. https://doi.org/10.1504/IJWGS.2018.09564
15. Zyskind, G., Nathan, O., & Pentland, A. (2015). Decentralizing privacy: Using blockchain to protect personal data. In *Proceedings of the IEEE Security and Privacy Workshops* (pp. 180–184). IEEE. https://doi.org/10.1109/SPW.2015.27