On the Use of Blockchain Technology to Improve the Reproducibility of Preclinical Research Experiments

Eduardo C. Oliveira[1] [a], Rafael Z. Frantz[1] [b], Carlos Molina-Jiménez[2] [c], Thiago Heck[1] [d], Sandro Sawicki[1] [e] and Fabricia Roos-Frantz[1] [f]

[1] *More Innovation Space, Unijuí University, Rua do Comércio, 3000, Universitário, Ijuí/RS, 98700-000, Brazil*

[2] *Department of Computer Science and Technology, University of Cambridge. CB2 1TN, Cambridge, U.K.*

eduardo.oliveira@sou.unijui.edu.br, {rzfrantz, thiago.heck, sawicki, frfrantz}@unijui.edu.br, carlos.molina@cl.cam.ac.uk

Keywords: Blockchain, Distributed Ledger, Data Analysis, Decentralised Technology, Experiments Irreproducibility, Experiments Reproducibility, DApps, DStorage, IPFS, Smart Contracts.

Abstract: Preclinical research is crucial for the advancement of life sciences. The use of experimental animal models in basic health sciences historically helped humanity to understand the pathological mechanisms of diseases and to develop therapeutic strategies, medicines and vaccines. Progress in this direction depends, to a large degree, on experimentation. Therefore, it is highly desirable that research experiments conducted on preclinical research are reproducible. Regrettably, a large number of experiments are not reproducible. Factors leading to irreproducible research on preclinical studies fall into four major categories: Biological reagents and reference materials, study design, data analysis and reporting and laboratory protocols. The data analysis and reporting category concentrates 25.5% of the total factors. Is estimated that $7.19 billions of the total research budget is funding irreproducible experiments. It is widely acknowledged that sharing experimental data between different institutions and cooperative researchers worldwide helps in experiment reproducibility which results in science and technology acceleration and innovation. Data sharing involves several data operations: The researcher needs to collect the data, protect it to prevent accidental and malicious deletion and corruption and make it available to colleagues, possibly, to the general public. The execution of these operations is cumbersome and error prone unless appropriate technology is used. This paper suggests and explores the use of blockchain to improve the reproducibility of experiments.A blockchain is a decentralised database that offers several properties that can be used advantageously in the collection, storage and sharing of experimental data, for instance, it prevents deletion.

1 INTRODUCTION

Preclinical research is a key aspect for the advancement of life sciences. The use of experimental animal models in the basic health sciences historically helped the humanity to find pathological mechanisms of different diseases as well has promoted the development of many therapeutic strategies, medicines and vaccines.

A crucial aspect of preclinical investigation is the integrity of the research which strongly encourages the achievement of the 3 R's: Refinement, Reduce and Replace on the use of animal models as much as possible.

In this way, ethically and scientifically, the quality and trustworthiness on the results is measured by the methodological accuracy of the experiments, thus research data must be reported with transparency and submitted to the scrutiny of regulatory bodies, scientific community, publishers, reviewers and readers.

It is paramount that experiments conducted on preclinical research be reproducible. Sharing experimental data between institutions and cooperative researchers worldwide may turn science and technology innovation in the health field faster and efficient to solve different demands of the population.

However, a large number of experiments are not reproducible. According to Freedman et al (2015) a staggering rate ranging from 51% to 89% is esti-

[a] https://orcid.org/0000-0001-5283-7732
[b] https://orcid.org/0000-0003-3740-7560
[c] https://orcid.org/0000-0002-3617-8287
[d] https://orcid.org/0000-0002-1242-5423
[e] https://orcid.org/0000-0002-7960-0775
[f] https://orcid.org/0000-0001-9514-6560

Oliveira, E., Frantz, R., Molina-Jiménez, C., Heck, T., Sawicki, S. and Roos-Frantz, F.
On the Use of Blockchain Technology to Improve the Reproducibility of Preclinical Research Experiments.
In *Proceedings of the 25th International Conference on Enterprise Information Systems (ICEIS 2023) - Volume 1*, pages 173-179
ISBN: 978-989-758-648-4; ISSN: 2184-4992

mated, a scenario that represents a big challenge for the field (Freedman et al., 2015).

Factors leading to irreproducible research on preclinical studies fall into four major categories: The first category (C1) represents biological reagents and reference materials, the second category (C2) is related to study design, the third category (C3) to data analysis and reporting and the fourth category (C4) to laboratory protocols.

Biological reagents and reference materials concentrates 36.1% of the failures in reproducing experiments, usually related to issues with non validated or contaminated reagents and inadequacies in biological materials handling.

Study design presents 27.6% of the factors leading to irreproducibility, it is largely associated to inconsistencies or lack of methods for sampling, randomisation and blinding.

Data analysis and reporting concentrates 25.5% of the total factors and is linked to poor statistical methods, unclear criteria for missing data, data inclusion and exclusion and outliers handling. Topics such as absence of disclosure of full results and primary data also constitutes factors in this category.

Laboratory protocols accumulates 10.8% of the overall causes and correlates to lab operations not adhering to standards and best practices when conducting preclinical experiments.

This article focuses on the Data Analysis and Reporting category of factors leading to irreproducible animal research.

The following statistics illustrate the economic impact of the reproducibility problem: The annual spend on preclinical research in the United States alone is estimated to be $56.4 billions, applying a irreproducibility rate of 50% means that $28.2 billions of the total budget is funding researches that are not reproducible, $7.19 billions only on the Data Analysis and Reporting category.

Normally, critical data is generated, compiled, analysed and reported at each stage of an experiment, in many cases this data in manually written in paper notebooks, or digitally recorded in spreadsheets, word processors kept locally on the researcher computer or shared by email or a cloud based storage.

To develop his or her activities a researcher needs to preserve data collected and avoid data deletion and tampering, get access to experiment data anytime, anywhere, track changes in data, openly publish data and manage data life cycle.

The blockchain technology is a digital ledger where transactions are stored in a chronologically linked sequence of blocks (hence the name blockchain), blocks are stored and processed at nodes on a decentralised network, every node on the network can receive transactions, all the participant nodes should agree about the final state of every transaction, accept or reject it, a process called consensus, once accepted a transaction is permanently recorded on the digital ledger.

Blocks are data structures that stores a specific amount of transaction records, minting is the process of creating and adding new blocks to the blockchain.

A public or open blockchain operates over the Internet fabric with geographically dispersed nodes, a blockchain protocol implements a code based operation with no central authority, nodes are free to enter or leave the network. A incentive mechanism rewards participant nodes and provides the proper balance of activity on the network to protect it from malicious attackers.

Bitcoin is the first large scale implementation of blockchain to gain mainstream adoption and along with Ethereum represents the bedrock for several use cases supporting the worldwide flow of financial assets and fostering new applications and business models in property rights, healthcare, logistics, manufacturing, energy and transportation among others initiatives (Nakamoto, 2009).

The Blockchains feature universal principles that can address the needs of a researcher regarding the handling of data generated from experiments: Immutability, resilience, traceability, decentralisation and programmability. There is a set of technologies with built-in functionalities that relates directly with each of the universal principles, connecting them to the needs of a preclinical researcher.

Blockchain networks are now widespread present and adoption is attracting financial and intellectual investments in large scale.

Applying this computational strategy in the sharing of data between researchers might help science and technology development and, at the same time, contribute to the improvement of animal research considering the 3R's international goals. In this paper we explore the use of blockchain to improve the reproducibility of experiments.

Herein, we present the study organised as follows: Section 2, discusses the related works approaching the preclinical research irreproducibility crisis; Section 3 provides an overview on blockchain technologies that supports our proposal; Section 4, introduces our proposal based on blockchain to improve the reproducibility of preclinical research experiments; and, finally, Section 5 concludes this paper.

2 RELATED WORK

To address the reproducibility problem, several approaches have been proposed in the literature. To achieve a higher level of reproducibility in preclinical research Landis et al. (2012) recommend the adoption of a core set of reporting standards for experimental group randomisation, blinding processes for investigators and animal care takers, sample size estimation, statistical models and *a priori* criteria for stopping data collection, inclusion, exclusion and removal of data, handling of outliers and missing data (Landis et al., 2012).

Vasilevsky et al. (2013) acknowledge the important contribution of guidelines for reporting of *in vivo* experiments such as the ARRIVE guideline (Percie du Sert et al., 2020) that was proposed to improve research reproducibility. The authors point the lack of proper unique identification of material resources such antibodies and model organisms as a critical irreproducibility factor yet not properly addressed by the current guidelines. They propose to increase the identifiability through a better tracking of research resources using electronic lab notebooks, management software and resource sharing repositories that can store and report the unique identities for each resource through the entire experiment cycle (Vasilevsky et al., 2013).

The study of Collins et al. (2014) explore actions the US National Institute of Health (NIH) is planning to enhance reproducibility in preclinical research, such as training on research reproducibility, transparency and study design, implementation of checklists to assert adequate experimental design on the evaluation of grant applications, a big data initiative, the Data Discovery Index (DDI), for sharing of unpublished primary research data sets and an online open forum for comments on published articles, PubMed Commons, Manolagas et al. (2014) subscribe on these recommendations for the bone research field (Collins and Tabak, 2014) (Manolagas and Kronenberg, 2014).

To increase the value of research Ioannidis et al. (2014) recommends the public availability of research protocols, original data and statistical analysis scripts of experiments, additionally public reviews and periodic comparisons between study protocols and results might provide valuable feedback for investigators and journals (Ioannidis et al., 2014).

Elaborating on the importance of the reproducibility as a foundational component for advancements in preclinical research Begley et al. (2015) points that due to the biological variability of the subject systems is not possible to achieve precisely replication of experiments, but is still vital that the major findings of the original experiment could be confronted with the results of further iterations while accepting a reasonable uncertainty degree, states that addressing the challenge of the reproducibility crisis requires a multidisciplinary approach by the stakeholders and advocates for more rigour in study design, training, statistical methods and primary data sharing (Begley and Ioannidis, 2015).

Table 1: Categories approached by the literature.

	Categories			
Authors	C1	C2	C3	C4
Landis et al., 2012		✓	✓	
Vasilevsky et al., 2013			✓	✓
Collins and Tabak, 2014		✓	✓	
Ioannidis et al., 2014		✓	✓	
Manolagas and Kronenberg, 2014		✓	✓	✓
Freedman et al., 2015	✓	✓	✓	✓
Begley and Ioannidis, 2015		✓	✓	✓
Curtis et al., 2015		✓	✓	
Our Proposal			✓	

Providing detailed guidance for experimental design and data handling Curtis et al. (2015) proposes the eradication of undesirable and unnecessary sources of errors to promote experimental reproducibility (Curtis et al., 2015).

3 BACKGROUND

In order to build a disruptive model is important to put in perspective the core tenets of decentralisation in the context of application architecture.

Decentralised applications, commonly referred as DApps, are applications running on top of decentralised peer to peer networks (Raval, 2016). As a regular application, a DApp requires a front end interface for user interaction, a middleware for business logic processing and a persistence layer for read and write transactions, the difference is that DApps relies on a underlying decentralised infrastructure to get access to such resources.

Regarding computing and storage services, decentralised systems consume resources provided by nodes distributed through a network, the free flow of nodes entering and exiting the network is coordinated by algorithms, at the operations level, this distribution provides resources availability and fault tolerance.

A peculiar dynamics is at play on decentralised networks, constituent parts are not obliged to trust each other to collaborate, there is no central entity to arbitrate the network behaviour, therefore to promote and sustain decentralisation some strategies were developed, for instance, incentive mechanisms are used to entice the addition of new nodes and discourage the participation of rogue actors. For data integrity, consensus mechanisms are implemented to achieve a single version of the truth for every transaction among all participant nodes. To achieve a high degree of transparency and trust, usually the code supporting this type of governing model is open sourced and backed by a active community of developers.

At the application level, a DApp implements the business logic to execute processes through a smart contract, a special program that runs on top of a blockchain for the execution of well defined instructions, like agreements between parts in a traditional analogical contract (Szabo, 1994). Smart contracts presents important features: Automation, predictability, publicity and privacy.

Automation means that once a smart contract is deployed and invoked, it will be executed if all logical conditions are met, no human intervention is required.

Predictability relates to the precise outcomes generated by the execution of a smart contract, there is no room for bias or misinterpretation of the contract, unlike their traditional counterparts.

Publicity is grounded in the intrinsic nature of the public blockchain, the smart contract and associated transactions can be tracked and audited seamlessly, furthermore the terms of a smart contract are openly available for scrutiny.

Privacy is achieved through the pseudonymous condition of users at a public blockchain, transactions executed through a smart contract are linked to a blockchain address and not to a person's identity.

A blockchain network delivers processing power and a minimal storage space for code and state of a smart contract, usually a decentralised storage (dStorage) is used in conjunction with the blockchain to provide off chain storage capacity.

The Interplanetary File System (IPFS) is a protocol designed to provide decentralised storage services for application deployments and secure distribution of large data volumes, featuring version control and a location independent name space, files and other content types are accessed through a content identifier - CID (Benet, 2022). Based on a peer to peer network without central authority, it is not required that nodes trust each other to connect and transfer content.

Blokchain is the central technology for immutable transactions recording in the context of data handling in our proposal of a infrastructure for handling experiments data. It is a innovative approach that emerged from the convergence of several well established methods from the Mathematics and Computer Science bodies of knowledge such as cryptography hashes, asymmetrical encryption and peer to peer networks.

For a broader comprehension of the subject key definitions are provided (Daniel Hellwig, 2020):

- **Digital Ledger Technology (DLT):** Is a data repository that chronologically stores transactions in sequence, it is a digital version of the ledger book used to record property rights or financial records.
- **Blockchain:** A subset of the DLTs where transactions are recorded. The blockchain is a database containing all transactions. A copy of the blockchain is stored in every node of the network for availability and validation purposes. Bitcoin and Ethereum are the most prevalent cases of blockchains currently in use.
- **Block:** Chronologically linked data structures. Each block consolidates transactions that will be appended to the blockchain as a unit, the blockchain protocol specifies the number of transactions or the block size. Only valid blocks are appended to the blockchain, validation is obtained through a consensus mechanism.
- **Time Stamping:** Every transaction and every block are time stamped, allowing to trace back the proper order of transaction events since the inception of the blockchain.
- **Hash:** Hash is a mathematical function that receives an arbitrary size input and generates a fixed size unique output. Hashes are heavily deployed in blockchain operations for data integrity verification and block linkage. The hash algorithm SHA256 generates a 64 bytes output and its used for hash operations in Bitcoin (NIST, 2015).
- **Merkle Tree:** Is a tree like hierarchical data structure with a unique root, each level of the tree stores hashes computed from hashes of the previous level. In the context of blockchains, Merkle Trees holds the hashes of every transaction in the current block plus the root hash of the previous block providing a linkage of every block in the chain for integrity verification, preventing a malicious attacker to tamper data.
- **Consensus:** Every node on the network can accept transactions, but each transaction on a blockchain is only committed if the block is approved through a validation process, the network nodes must agree about the the block to be committed, this is settled trough a consensus mechanism, Proof of Work (PoS) and Proof of Stake (PoW) are the most prevalent protocols. On Proof of Work (PoW), all the participant nodes competes to win a compute intensive mathematical challenge, the winer node mints the block and appends it to the blockchain. On Proof of Stake

(PoS) blockchains randomly selected participants are obliged to hold a certain amount of coins or tokens to participate in the validation process, the stake of each participant defines the likelihood of wining the validation.

Blockchains can also be defined by the governance model:

- **Public Blockchains:** Participant nodes can freely join or leave the network, there is no control or veto to add or remove a node and no central authority for data verification, transactions are transparent to users and outsiders. Trust and privacy are not enforced by code or processes. Examples: Bitcoin and Ethereum.
- **Private Blockchains:** A central authority regulates the enrolment process and life cycle of participant nodes, all participants are acknowledged and the governance process provides a high degree of trust among the parts. Applies mostly to enterprise use cases and consortiums like the CDBC (Central Bank Digital Currency). Hyperledger and Ethereum Enterprise are the main players.

The first successful use case for the application of the blockchain technology, Bitcoin, developed by an individual or collective known as Satoshi Nakamoto released in 2015 a network protocol and digital coin currently in use by individuals and institutions, storing economic value and moving financial assets across borders worldwide. Bitcoin sparked a wave of initiatives on decentralisation and new plaforms such as Ethereum are flourishing creating an ecosystem for the digital economy and the Web 3.0 (Nakamoto, 2009).

4 OUR PROPOSAL

In this section we provide a detailed discussion about the main needs of a researcher related to data generated by experiments, we also contextualise the universal principles of the blockchain that are relevant to the improvement of experiments reproducibility and present a decentralised architecture for an application supporting the recording of experiments

4.1 Approaching Data

On conducting experimental preclinical studies, a researcher needs to preserve the integrity of generated data avoiding deletion or tampering or even fraudulent manipulation, preferably the researcher should safely access data anytime, anywhere and not only on lab premises. For accuracy of reporting is also fundamental that a researcher could track changes in data reported and most important could openly publish data with transparency.

4.2 Blockchain Properties

Universal principles featured in blockchain implementations:

- **Immutability:** An intrinsic property of blockchain based systems, records can only be appended to the database. Once a transaction record is committed to a valid block it cannot be modified or deleted. This provides the proper safeguard and integrity for experiment's data, avoiding further changes or even frauds that could harm research results.
- **Resiliency:** Blockchain architectures are based on peer to peer networks where each node plays a egalitarian role, holding the full blockchain database, providing resilience and fault tolerance, this features provides a layer of resilience and availability against failures and malicious attacks, protecting research records from deliberate or accidental deletions or corruption.
- **Traceability:** Every transaction committed to the blockchain is timestamped and all the records are openly available for tracing, providing transparency for research data along with the history of every new version.
- **Decentralisation:** There is no central authority to manage the network operations, nodes can freely join or leave the network, the rule set that coordinates the network is based on code. Incentive mechanisms regulates the economics of the network, providing rewards and inhibiting malicious attackers. Therefore, research data could be freely and openly shared with no need for a central repository controlled by an authority or institution.
- **Programmability:** The interactions with the blockchain can be controlled by special programs called Smart Contracts, these programs resides on the blockchain and their execution is enforced by a set of embedded rules, as in traditional analogical contracts. In the context of pre-clinical research Smart Contracts can establish the proper input and workflow of the data, checking for validity, authorship and expiration dates.

4.3 Blockchain-Based Architecture

To provide a fully decentralised architecture to support the handling of experiment's data, we have designed a DApp with four main components, as follows (see Fig. 1): Front-end, smart contract, distributed digital ledger (DLT) and decentralised storage (DStorage).

The interactions between the DApp front-end and the other components is executed through API calls. The front-end presents a web based user interface for sign up, authentication and data input. (1) A smart contract deployed on the blockchain is invoked for business logic processing. (2) Every transaction is permanently recorded on the DLT. (3) Uploaded files are stored on the DStorage which is implemented on a IPFS file system. (4) Every stored file receives an unique CID (content identifier) that is sent back to the DApp and written on the DLT for file hash check and retrievals.

The DApp web engine and user interface files are fully deployed on the DStorage as well, eliminating any point of dependency on central services or authority.

The user, typically a experimental researcher, access the application through the Internet and authenticates using his or her private key. Upon authentication the researcher inputs experiment's data following a proper protocol or guideline.

Inputs (e.g., authorship, study design, materials, statistical models, lab and environmental conditions, images and reports) are validated through a smart contract and the transactions are committed permanently, this step is backed by the immutability property of the blockchain and assures experiment's data integrity, avoiding data tampering, violation or fraud.

Files generated by the experiment are uploaded through the application to the DStorage and are versioned and protected, every file generates a CID that is written on the blockchain, future file retrievals are based on the CID.

Resiliency and availability of transactions and files is achieved through the decentralised nature of both the blockchain and the IPFS file system, the complete record of an experiment's data and files can be traced back to any point in time since the beginning of the experiment, data is openly exposed and shared with stakeholders through the public blockchain.

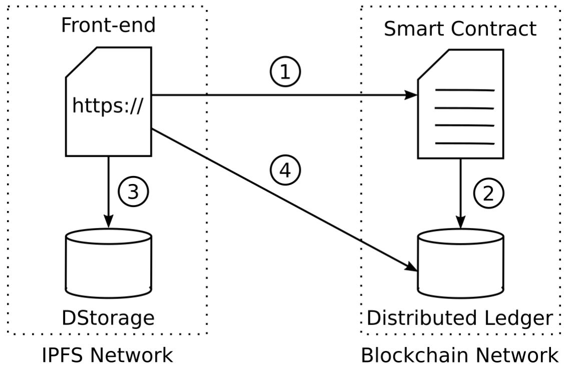

Figure 1: Proposed Architecture.

5 CONCLUSIONS

Our work aims to present the important role that decentralisation can play in the improvement of reproducibility of preclinical research through the application of its core principles on data generated by experiments.

By providing an DApp architecture based on blockchain and a distributed file system, it is possible to achieve consistency and validation for data input through smart contracts, safeguarding data records permanently in the blockchain and files in the DStorage, also allowing traceability and version control of the data and files in any point in time. Both the blockchain and the IPFS network provide availability and fault tolerance for records and files.

Since it relies totally on a decentralised model, no central authority can influence, dictate or veto the publishing of experiment's data, promoting transparency that leads to better practices of the whole experiment life cycle.

Future works can explore incentive mechanisms to attract institutions, publishers, researchers and reviewers to use the proposed architecture and also discuss interoperability with the current centralised applications.

ACKNOWLEDGEMENTS

This work was supported by the Research and Development Programme at UNIJUI; the Coordination for the Brazilian Improvement of Higher Education Personnel (CAPES); and, the Brazilian National Council for Scientific and Technological (CNPq) under grant 309315/2020-4.

REFERENCES

Begley, C. G. and Ioannidis, J. P. (2015). Reproducibility in Science: Improving the Standard for Basic and Preclinical Research. *Circulation Research*, 116(1):116–126.

Benet, J. (2022). IPFS - Content Addressed, Versioned, P2P File System (DRAFT 3). Last accessed on 30 Nov 2022.

Collins, F. S. and Tabak, L. A. (2014). Policy: NIH plans to enhance reproducibility. *Nature*, 505(7485):612–613.

Curtis, M. J., Bond, R. A., Spina, D., Ahluwalia, A., Alexander, S. P. A., Giembycz, M. A., Gilchrist, A., Hoyer, D., Insel, P. A., Izzo, A. A., Lawrence, A. J., MacEwan, D. J., Moon, L. D. F., Wonnacott, S., Weston, A. H., and McGrath, J. C. (2015). Experimental design and analysis and their reporting: new guidance for publication in BJP: Editorial. *British Journal of Pharmacology*, 172(14):3461–3471.

Daniel Hellwig, Goran Karlic, A. H. (2020). *Build Your Own Blockchain*. Springer Nature B.V., 1st edition.

Freedman, L. P., Cockburn, I. M., and Simcoe, T. S. (2015). The Economics of Reproducibility in Preclinical Research. *PLOS Biology*, 13(6):e1002165.

Ioannidis, J. P. A., Greenland, S., Hlatky, M. A., Khoury, M. J., Macleod, M. R., Moher, D., Schulz, K. F., and Tibshirani, R. (2014). Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet*, 383(9912):166–175.

Landis, S. C., Amara, S. G., Asadullah, K., Austin, C. P., Blumenstein, R., Bradley, E. W., Crystal, R. G., Darnell, R. B., Ferrante, R. J., Fillit, H., Finkelstein, R., Fisher, M., Gendelman, H. E., Golub, R. M., Goudreau, J. L., Gross, R. A., Gubitz, A. K., Hesterlee, S. E., Howells, D. W., Huguenard, J., Kelner, K., Koroshetz, W., Krainc, D., Lazic, S. E., Levine, M. S., Macleod, M. R., McCall, J. M., Moxley III, R. T., Narasimhan, K., Noble, L. J., Perrin, S., Porter, J. D., Steward, O., Unger, E., Utz, U., and Silberberg, S. D. (2012). A call for transparent reporting to optimize the predictive value of preclinical research. *Nature*, 490(7419):187–191.

Manolagas, S. C. and Kronenberg, H. M. (2014). Reproducibility of Results in Preclinical Studies: A Perspective from the Bone Field: Reproducibility of Results in Preclinical Studies. *Journal of Bone and Mineral Research*, 29(10):2131–2140.

Nakamoto, S. (2009). Bitcoin: A peer-to-peer electronic cash system.

NIST (2015). Secure hash standard (shs). Technical Report Federal Information Processing Standards Publications (FIPS PUBS) 180-4, U.S. Department of Commerce, Gaithersburg, MD.

Percie du Sert, N., Ahluwalia, A., Alam, S., Avey, M. T., Baker, M., Browne, W. J., Clark, A., Cuthill, I. C., Dirnagl, U., Emerson, M., Garner, P., Holgate, S. T., Howells, D. W., Hurst, V., Karp, N. A., Lazic, S. E., Lidster, K., MacCallum, C. J., Macleod, M., Pearl, E. J., Petersen, O. H., Rawle, F., Reynolds, P., Rooney, K., Sena, E. S., Silberberg, S. D., Steckler, T., and Würbel, H. (2020). Reporting animal research: Explanation and elaboration for the arrive guidelines 2.0. *PLoS biology*, 18(7):e3000411–e3000411.

Raval, S. (2016). *Decentralized Applications*. O'Reilly Media, Inc., 1st edition.

Szabo, N. (1994). Smart contracts. Last accessed on 28 Jan 2021.

Vasilevsky, N. A., Brush, M. H., Paddock, H., Ponting, L., Tripathy, S. J., LaRocca, G. M., and Haendel, M. A. (2013). On the reproducibility of science: unique identification of research resources in the biomedical literature. *PeerJ*, 1:e148.