

Introduction

Eduardo Oliveira

February 18, 2025

Abstract

Abstract

The scientific community faces a critical challenge: failed reproductions and replications, which undermine the validity and reliability of research findings. The lack of effective mechanisms to ensure reproducibility leads to decreased accuracy, increased costs, and diminished confidence in scientific discoveries. To address this issue, we propose a decentralized approach leveraging blockchain technology, smart contracts, and IPFS to enhance the adoption of Open Science principles. Our research investigates the question: How can blockchain technology, smart contracts, and IPFS facilitate data sharing and foster reproducibility in scientific research? By building an artifact that emulates an Open Science platform functions, we analyze the potential benefits and technical challenges of this integration, aiming to contribute to the development of a more transparent, reliable, and trustworthy scientific ecosystem.

1 Introduction

The reproducibility crisis in science refers to a situation where the results obtained from an experiment or study cannot be reliably replicated by others, even when using the same methods and materials as originally used. This issue is particularly concerning in fields such as biology, medicine, and physics, where accurate replication of results can significantly impact scientific progress and public trust.

In recent years, there has been a growing concern about the lack of transparency and reproducibility in scientific research. Many studies have shown that a substantial number of published findings fail to hold up under rigorous scrutiny when retested or replicated by other researchers. This raises serious questions about the reliability of scientific knowledge and can lead to wasted resources, misdirected efforts, and loss of public confidence in scientific research.

The main reasons for this crisis include issues with study design, data quality, measurement tools, and researcher practices. For instance, poorly designed studies may not capture important factors that affect outcomes, while flawed measurement tools can yield misleading results. Furthermore,

researchers' failure to report certain information or omitting key details in the initial publication can lead to incorrect assumptions about the experiment's outcome.

One potential solution is to promote transparency and reproducibility by using Open Science practices. This involves publishing detailed descriptions of methods, data, and materials alongside research findings. It also includes making raw data available for other researchers to verify results through replication studies or further analysis. This dissertation aims to explore how decentralized technologies, such as blockchain, InterPlanetary File System (IPFS), and smart contracts, can foster reproducibility and trust in scientific research by addressing these challenges in a platform.

The primary objective of this study is to evaluate the potential of decentralized technologies in tackling reproducibility issues and propose an artifact to be applied in scientific workflows and materials, such as reports, protocols, datasets, images and videos. In the context of this dissertation, 'decentralized technologies' refer to peer-to-peer networks that operate without a central authority or single point of failure, thereby promoting censorship resistance, resilience, and transparency [?].

The core focus will be on the following decentralized technologies and their potential applications in the realm of reproducibility in scientific research:

- **Blockchain:** A distributed, immutable ledger technology that can securely record and verify transactions across a peer-to-peer network. Blockchain's transparency, immutability, and security can be harnessed to track data provenance, ensure data integrity, and validate research processes [?].
- **InterPlanetary File System (IPFS):** A peer-to-peer distributed file storage system that uses content-addressed, peer-distributed files and networking protocols to create a permanent, decentralized network. IPFS can facilitate secure, tamper-evident data sharing and hosting, essential for preserving and reproducing research outputs [?].
- **Smart Contracts:** Self-executing contracts with the terms of the agreement directly written into code. Smart contracts can automate tasks, enforce agreements, and enforce data input, thus streamlining research processes and ensuring methodological rigor [?].

This dissertation will investigate the underlying principles and functionalities of these technologies, followed by an in-depth analysis of their potential to enhance reproducibility in scientific research. The practical implementation and evaluation of these technologies in a real-world scientific research scenario will also be explored to demonstrate their feasibility and assess their tangible benefits.

1.1 Blockchain and Open Science

The relation between reproducibility and Open Science is crucial, as it ensures that findings can be duplicated and verified without artificial constraints or obstacles. This allows researchers to build upon each other's

work, fostering a collaborative environment where knowledge is shared freely for everyone to benefit. By promoting transparency in research, the scientific community aims to reduce the risk of misconduct, increase accountability, and ensure that scientific progress can be sustained over time without setbacks due to irreproducible findings or biased data manipulation.

Examples of Open Science applications include:

- Making research articles openly accessible.
- Encouraging the use of open-source software in scientific computing and data analysis.
- Promoting transparent communication and reporting of methods and results.
- Ensuring that all data used for experiments or analyses is made available to others.
- Providing detailed documentation about experimental procedures.
- Following best practices like preregistration of studies before data collection.

By implementing these strategies, researchers can ensure their work is more accessible and usable by the broader scientific community, allowing for collaboration and building upon each other's discoveries without artificial constraints. This ultimately contributes to advancing the understanding and development of science in a more efficient and sustainable manner.

1.2 Technical Approaches: Blockchain, Smart Contracts, and IPFS

The use of blockchain, smart contracts, and IPFS can significantly enhance the reproducibility of scientific research through Open Science. Blockchain technology allows for secure, transparent, and tamper-proof record-keeping and data storage. Smart contracts ensure that complex rules are executed automatically when certain conditions are met, which can streamline processes such as peer review and submission of research articles. By utilizing these technologies, researchers can maintain a permanent and auditable record of their work, providing an additional layer of transparency and accountability.

IPFS enables the sharing and interconnection of large files across the network, facilitating collaboration among researchers working on complex projects. It also allows for decentralized data storage, reducing reliance on centralized servers and minimizing the risk of data loss due to technical failures or intentional sabotage. The combination of these technologies can create a robust infrastructure that supports the open-source development and sharing of scientific knowledge, ultimately enhancing research reproducibility and collaboration among scientists worldwide.

1.3 Potential Applications of a Decentralized Open Science Platform

A platform encompassing the aforementioned technologies could be applied in various scientific use cases:

- **Decentralized Repositories:** Researchers can store and share research findings in an open-access manner with clear guidelines for transparency and reproducibility.
- **Smart Contracts for Peer Review:** Automated peer review processes ensuring thorough manuscript vetting before publication.
- **Decentralized Collaboration:** IPFS can enable researchers to work together on large-scale projects while reducing reliance on centralized data storage.

In summary, the combination of blockchain, smart contracts, and IPFS has vast potential for improving research reproducibility through Open Science. By providing secure, transparent, and auditable records of research activities, these technologies can help reduce corruption in scientific research, increase accountability among researchers, and foster collaboration between scientists worldwide.

2 Research Context

Research Context

The scientific community has long recognized the importance of reproducibility in ensuring the validity and accuracy of research findings. Reproducibility refers to the ability of researchers to replicate another researcher’s study or experiment, with minimal changes, to verify its results and conclusions (National Science Foundation, 2019). However, despite efforts to address this issue, failed reproductions remain a persistent problem in scientific research.

Studies have shown that reproducibility rates vary widely across disciplines, but even among those with high reputations for rigor, the rates are often disappointing. For example, a study published in the journal *Science* found that only 22

The reasons for these poor rates of reproducibility are complex and multifaceted. Methodological flaws, such as sampling biases or inadequate control groups, can contribute to the failure to replicate results. Additionally, a lack of transparency in research practices, including insufficient details about methods, materials, and results, can make it difficult for others to replicate studies. Furthermore, limited resources, such as small research teams or those with limited funding, may not be able to conduct the rigorous experiments required to achieve high levels of reproducibility.

Inadequate data management practices are also a significant contributor to failed reproductions. Poor data storage and sharing practices can lead to errors or loss of data, making it impossible for others to verify results. This is particularly problematic in fields where data-intensive research is common, such as physics and biology (Katz Martin, 2018). The

consequences of these problems are far-reaching, not only undermining the credibility of individual studies but also compromising our collective understanding of complex phenomena.

Given the challenges posed by failed reproducibility, it's essential to develop new approaches that promote transparency, collaboration, and data sharing. Our research aims to contribute to this effort by exploring the potential of blockchain technology, smart contracts, and IPFS for promoting reproducibility in scientific research.

3 Motivation

The scientific community is facing a critical challenge in maintaining the integrity and validity of research findings. The failure to reproduce results has become an endemic problem, not only hindering the progress of individual studies but also compromising our collective understanding of complex phenomena. This issue has far-reaching consequences, including undermining the credibility of individual researchers, institutions, and even entire fields of study.

The root cause of this problem lies in the broader structural issues that pervade many fields. The growing reliance on data-driven approaches, computational modeling, and high-throughput experimentation has created an environment where results are increasingly difficult to verify or replicate across diverse domains of science. Moreover, the limitations of existing publication systems, funding mechanisms, and peer-review processes have failed to address this issue effectively, exacerbating the problem in fields as disparate as physics, biology, chemistry, and social sciences.

As a result, there is a pressing need for innovative solutions that can foster greater transparency, collaboration, and reproducibility in scientific research. Our proposed decentralized approach aims to address these challenges by providing a platform for researchers to share data, methods, and results in a secure, transparent, and auditable manner.