

The Case for a Decentralized Science: Open Data Sovereignty, Distributed Cloud Resources, and Rigor

Poster No:

2552

Submission Type:

Abstract Submission

Authors:

Shady El Damaty¹

Institutions:

¹Opscientia, LTD., Singapore, Singapore

First Author:

Shady El Damaty

Opscientia, LTD.

Singapore, Singapore

Introduction:

The field of neuroimaging is approaching a critical mass of data production and consumption with the advent of massive consortium studies collecting high-spatiotemporal-resolution data from large cohorts of participants. The rate of data production has outpaced the ability of individual labs to store, analyze, and share large datasets with colleagues. The current web infrastructure was not developed with persistent identity, confidentiality, data sovereignty, and provenance for scientific reproducibility in mind.

Recent technological breakthroughs in cryptography and peer-to-peer protocols have ushered in the decentralized web 3.0. Distributed storage with IPFS, decentralized confidential cloud computation with the Ocean Protocol and iExec RLC, and DID-standard federated identities now power innovative web applications with built-in data ownership, versioning, and high bandwidth for on-demand sharing. Smart contracts running on the Ethereum blockchain allow for novel crowdsourced citizen science, complete transparency of science funding, universal programmable permissions for dataset access, and perpetual support for decentralized autonomous organizations (DAO). DAOs are community governed autonomous smart contracts that perform meta-analysis, data aggregation, and content publication with immortal hyperlinks that do not depend on a central web server or authority.

The release of community standards for neuroimaging research, such as BIDS, is well-timed with the refinement of web 3.0 services to bring big data neuroscience research on-chain, reaping the benefits of distributed cloud computation, storage, and security. The case for a decentralized science (DeSci) has never been stronger. Here we describe steps towards a decentralization of neuroscience research operating outside of traditional academia, vested interests, and scientific publication corporations.

Methods:

The current work reviews web 3.0 onboarding efforts during the Global Brainhack. Contributions to outstanding problems and barriers for web 3.0 were solicited from the community through the Brainhack GitHub Issues forum. All feedback, code contributions, and scholarly materials were made publicly available on the linked repository. Community onboarding for IPFS, public key encryption with Ethereum, and dataset tokenization on Ocean Protocol were explored and major obstacles to productivity annotated.

Results:

The Global Brainhack proved to be a potent catalyst for soliciting community activity to build infrastructure and launch a DAO. Outstanding obstacles for effectively bootstrapping a neuroscience DAO include a need for community education on web 3.0 applications, standards, and services. Peer-to-peer sharing of massive datasets was revealed to be limited by the availability of pinning services on IPFS, and thus require accessory funding or large datacenter participation to maintain availability of infrequently requested data. Potential integrations with git annex and datalad were identified as future opportunities.

Conclusions:

The maturity of web 3.0 services has reached a sufficient level to support decentralized autonomous neuroscience organizations on the Ethereum blockchain to perform data aggregation, computation, and publication. Coordination between dataset providers and data centers is required to maintain availability of massive datasets during the community bootstrap phase. The long-term outlook for a decentralized science is increasingly promising for catalyzing the next wave of scientific discovery in neuroimaging by citizen scientists.

Modeling and Analysis Methods:

Other Methods

Neuroinformatics and Data Sharing:

Databasing and Data Sharing ¹

Workflows

Informatics Other ²

Keywords:

Data analysis

Data Organization

Data Registration

Open Data

Open-Source Software

Other - decentralized autonomous organizations

¹²Indicates the priority used for review

My abstract is being submitted as a Software Demonstration.

No

Please indicate below if your study was a "resting state" or "task-activation" study.

Other

Healthy subjects only or patients (note that patient studies may also involve healthy subjects):

Healthy subjects

Was any human subjects research approved by the relevant Institutional Review Board or ethics panel? NOTE: Any human subjects studies without IRB approval will be automatically rejected.

Not applicable

Was any animal research approved by the relevant IACUC or other animal research panel? NOTE: Any animal studies without IACUC approval will be automatically rejected.

Not applicable

Please indicate which methods were used in your research:

Other, Please specify - qualitative outcomes, code review

Provide references using author date format

El Damaty, S. (2021), "The Case for a Decentralized Science." Brainhack Proceedings (submitting).