Thinking of a Blockchain for VIVO

Alexander Garcia Castro, Federico Lopez, Mike Conlon

VIVO is an example of a decentralized system; institutions publish VIVO data just by adhering to a simple data structure in the form of an ontology. Similar to a distributed ledger, VIVO is a decentralized database that is used to maintain a continuously growing list of records. These records aim to include a comprehensive list of scholarly outputs. Although outputs are often described in a single narrative, the published reviewed paper, the research has generated many other outputs, which may or may not be recorded. The Research Object (RO) is a container for the purpose of recording all scholarly outputs associated with a particular research effort. The paper, datasets, software, notes, are bundled in an RO. The published paper is often a small fraction of the content of an RO. VIVO represents works beyond papers, datasets, software and reports. More is needed to represent all items that might appear in an RO. But another clear gap exists in identifying the collection of works, traditional and non-traditional, as research objects that can then represent relationships between the works. The richness in relations across ROs is also unrepresented, time-based relationships, and logical precursors, for example.

We are designing a system based on Blockchain technology to keep the ledger for ROs; from the conception through the life cycle of the RO. Wikipedia describes blockchain as follows:

"By design, blockchains are inherently resistant to modification of the data. Once recorded, the data in any given block cannot be altered retroactively without the alteration of all subsequent blocks and the collusion of the network. Functionally, a blockchain can serve as an open, distributed ledger that can record transactions between two parties efficiently and in a verifiable and permanent way.The ledger itself can also be programmed to trigger transactions automatically."

We argue that all the value chain for ROs should be kept in a distributed ledger; in this way the RO is preserved and, modifications and transactions over the RO are kept in the ledger. Researchers are thus able to account for their products in a data-based ecosystem that makes it possible for third parties to develop specialized tools over the ROs, researchers and transactions. The Blockchain also preserves the metadata associated with transactions and modifications of ROs; the data is portable, once on the ledger the researcher and the ROs do not depend on a node for existing. The physical existence of the objects is left to specialized apps, e.g. github, figshare, dspace, dryad etc, using protocols, part of our design, to ensure provenance and traceability, thus asserting the existence of the object. Institutions are suppliers of metadata to the ledgers of their people. VIVO is enhanced as a portal, it becomes an integral tool for researchers to increase the value of the ledger, and provides tools for reaping value from the distributed collection of ledgers. In this way, the value of the commons is safely kept. Others may enhance the value of the commons by mining and contributing to the commons, by certifying elements of the commons, and by providing additional means for making use of the commons. The distributed ledger makes it possible to account for all the value that is currently unaccountable. In this context, the tragedy of the commons is not depletion of resources; it is to generate uncountable additional resources.