Using blockchain to store academic accreditations

Authors: Luiz Gustavo Ferraz Aoqui (IBM) and Andrei Moskalev (Intela)

Overview

Fake academic diplomas have been around ever since receiving one generated value. In 1924 the United States Congress held hearings regarding diploma scams and from 1980s to early 1990s the FBI had a task force called DipScam whose goal was to track and take down these fraudulent operations [1].

With the rise of the Internet in the 2000s, diploma mills have become a billion dollar industry [2]. Axact, a company based in Pakistan but with business all over the world, sold over 215,000 fake diplomas in 2015, earning 51 million dollars in a single year [3] [4].

One of the aspects that render fake diplomas effective, and thus sought after by thousands of people, is the difficulty in distinguishing a genuine but not very well-known academic institution from a fraudulent one [5]. Globalization has magnified this issue, as more people are crossing borders in search for new opportunities. Companies, universities and other institutions need to deal with an increasing number of accreditation issuers.

Digitizing these documents is a natural next step required to cope with the threats of diploma mills. Cryptographic technologies such as digital signatures can help improve trust in the authenticity of a document, but they don't solve all the problems. Digital records must be stored in a secure and durable media, which is often centralized within storage providers or private owned infrastructure.

A blockchain network can be used to store these documents. It would offer a tamper-proof record of all transactions related to a diploma and it would have all relevant parties involved (educational institutions, learners, companies, government etc) working together, validating and endorsing these transactions to reach consensus regarding which documents are authentic and which ones should be rejected.

Documentorum

Documentorum is a project that aims to provide easy and fast validation of academic records and other types of documents. It will handle all stages of the verification process of a document, from confirming the issuing institution is properly accredited, to validating if a physical copy presented to someone matches the digital record stored in the blockchain.

The initial implementation uses a Hyperledger Fabric blockchain to store digital certificates issued upon completion of online courses, but it will soon expand to support other scenarios.

Challenges

In this section we list some of the challenges identified during the project.

Identity

- **How to identify learners across multiple institutions?** Learners will interact with multiple institutions along their life so they should be able to maintain their identity across them.
- **How do learners maintain their profile information in a blockchain environment?** Learners should be able to control what information is stored about them, even in an append-only environment such as a blockchain.

Privacy and data sovereignty

- **How do we control how data is shared?** Learners should be the sole owner of their own information and should be able to control *what* data is shared, with *who* this data is share, *when* the data will be initially shared and for *how long* this data will shared.
- **How do institutions share information about a learner without violating their privacy?** In order for the blockchain to be truly useful, all parties involved should be able to collaborate with each other (governments endorse diplomas issued by an university, an employer validates a candidate's qualifications etc.), but this exchange of information should respect the privacy of its users.

Public vs. permissioned

- **Should a system like this use a public or a permissioned blockchain?** Public blockchains allow full transparency and decentralization as all transactions are available to everyone to validate them and anyone can join the network. But this also limits the amount and type of information that can be stored. Transactions also often require a fee be paid in order for them to be submitted. A permissioned blockchain has more control over who can join be part of it, so data can be stored in-chain more liberally, but it risks becoming an oligarchy.

Off-chain storage

- **How and when to use off-chain storage for additional metadata abut a document?** Blockchains are known to have scalability issues, so it is a common practice to separate larger chunks of metadata (often in binary format, like PDFs and images) and store them off-chain, having only a reference or a cryptographic hash stored in-chain. This allows greater flexibility as to what formats of documents can be handled, but this leads to a hard dependency to whatever storage mechanism is used.

Conclusion

Blockchains have the potential to solve the issue of fake diplomas whilst also providing better privacy and streamlining processes that involves multiple independent parties.

But in order for this to become a reality there are still challenges that need to be solved.