AI & Blockchains

These two terms are broad in terms of function. **AI** is a corpus of computer-based methods to apply statistical and network tools to analyze data. The idea is that AI augments data scientists to tackle big data challenges. The goal is for AI to create systems that learn and predict without human intervention. **Blockchains** live in a distributed Peer-to-Peer network architecture. Their purpose is to create an immutable record of events in a trustless network of participants.

In order for AI to gain knowledge to create a defined outcome, the more data the better. Where the data comes from, how it is validated and what formats are required is the hard part. For instance, in healthcare, genomic data, treatment results in episodes and outcomes, social determinants, drug efficacy and trials, all play a role. Creating an ontology that classifies data inputs is a major barrier to improving healthcare programs. Allowing access and meeting regulatory rules is another barrier. And finally, recording how this information is shared and used becomes paramount.

Blockchains, originating in their support of anonymous cryptocurrency transactions , have evolved to new non-crypto applications: supply chains, track and trace, recording of financial instruments such as content consumption, insurance and mortgages. Some are public where anyone can transact anonymously. Others are permission driven where participants are vetted and visible. The result is the same – a record of ledger of an event that requires an immutable record.

Can blockchains fulfill this role? Of course. So can a secure database. The question is how the latter can provide the same assurances for immutability of content. So far, it seems the weaker solution.

 One attribute of a blockchain method is that there are multiple copies, distributed across a network (mainnet), that requires a consensus of copy holders (in Bitcoin these are miners and in Ethereum these are validators) to agree to accept a record to be written to a block in the chain. It makes it difficult though possible, to alter the record being stored.

What is stored in the record leads to another issue: that what is being stored is valid and the issuer of the content is trusted. In the case of cryptocurrency, the problem is simple. Someone owns an alt-currency and transacts with another party, anonymous or otherwise. In the case of blockchains as a distributed method of storage, the transaction or event may have multiple parties in an ad hoc relationship whose cooperation requires multiple steps of engagement and recording.

That's why distributed ledger technology (DLT) plays an important role, whether on or off chain. Blockchain and DLT, together are a powerful tool for data exchange across a trustless network like the Internet.

Privacy

Data is just data until it becomes information. The Web is a perfect example. Many parties capture the intersection of participants in using networked services, especially with mobile users. What you view, click or engage with in an app is collected. When you go from one service to another is tracked. How long you use the service and frequency is captured. All this creates a profile.

The **G-MAFIA** (Google, Microsoft, Amazon, Facebook, IBM, Apple) and **BAT** (Baidu, Alibaba, Tencent) are all about profiling that translates into revenues. This information is free to them but not to third parties such as advertisers. It's the current concern of governments because of lacking regulation. What the EU is doing with GDPR is on one extreme, attempting to enforce an opt-in model over the default opt-out. On the other extreme, China wants to capture and apply profiles for the benefit of their social policies that do not favor privacy in any sense. What is common to both is that the data originator has no economic contract to permit its use by third parties.

There are initiatives that address how a data owner should be rewarded once their data is converted to information that has value. While this may be feasible in the G-MAFIA case (Facebook is considering creating a crypto-token as a means of exchange and reward), its too early to see where this goes. This subject is referred to as the **Privacy-Economy** and blockchain will play a role, at least in the form of record keeping available to the data creator. However, it is not a panacea.

Why? Blockchains do not address how to deal with data with different half-lives (minute, hour, days, forever). Since blockchains are intended to provide a permanent record, the chain can grow exponentially. Recording which farm lettuce came from is not the same as a mortgage.

How? Blockchains start with a root block and are linked to succeeding blocks using a mathematical tool – Merkle roots. Taking out a block destroys the chain. At some point the data has no time-sensitive value. With databases, older data can be archived to keep the data store efficient. With blockchains it's not feasible. With new improvements to blockchain protocols, this problem may not be serious but for now, beware.

Andre Szykier

UbiVault   Founder
BlockchainBTM   CTO
Aegis Health Security   Chief Scientist
andre@ubivault.com