Blockchain Solutions for Big Data Challenges

A Literature Review

Elena Karafiloski
Faculty of Computer Science and Engineering,
Ss. Cyril and Methodius University,
Skopje, Macedonia
elena.karafiloski@gmail.com

Anastas Mishev
Faculty of Computer Science and Engineering,
Ss. Cyril and Methodius University,
Skopje, Macedonia
anastas.mishev@finki.ukim.mk

***Abstract*— The popularity of Blockchain technology and the huge extent of its application, results with much ongoing research in different practical and scientific areas. Although still new and in experimenting phase, the Blockchain is being seen as a revolutionary solution, addressing modern technology concerns like decentralization, trust, identity, data ownership and data-driven decisions. At the same time, the world is facing an expansion in quantity and diversity of digital data that are generated by both users and machines. While actively searching for the best way to store, organize and process Big Data, the Blockchain technology comes in providing significant input. Its proposed solutions about decentralized management of private data, digital property resolution, IoT communication and public institutions' reforms are having significant impact on how Big Data may evolve. This paper presents the novel solutions associated with some of the Big Data areas that can be empowered by the Blockchain technology.**

Keywords—blockchain; big data; digital property; privacy; smart contracts; internet of things; healthcare

I. INTRODUCTION

Blockchain is an underlying technology of Bitcoin that soon emerged as the real discovery of Satoshi Nakamoto [1], thus making Bitcoin just the first of the many future Blockchain implementations. Technically, Blockchain is a distributed public ledger that contains all the transactions that ever executed in the system. It exists on a P2P network where every full node stores a copy of the Blockchain ledger. There is no central authority that manages the Blockchain database. This concept of obtaining a database only between the actual and equal users of the system sets the base for building so called: "decentralized trust". For the transactions to be validated and authorized, a consensus of nodes that agree upon the issue is required. The concept of decentralized trust comes as opposite solution to almost every system that we have built using the client-server architecture. By removing the central authority out of the system, there is no longer a mediator processing the actions and the data. That results with lower transactional costs, non-reversible transactions and no need for trust in the governments or private corporations. In this solution, Blockchain users don't even need to trust the other party included in the transaction. They should trust only the system and the code.

Blockchain is a database that stores all the transactions grouped in blocks. When new transaction is created, the sender broadcasts it in the P2P network to all the other nodes. The transaction is still new and not confirmed. As the nodes receive the transaction, they validate it and keep it in their transactional pools. To validate transactions means to run predefined checks about the structure and the actions in the transaction. Special node types called miners create a new block and include all, or some of the available transactions from their transaction pool. Then the block is mined, which is a process of finding the proof of work using variable data from the new block's header [27]. Finding the proof of work is continuous calculation of a cryptographic hash that fits the defined difficulty target. Mining requires a lot of processing power and the miners use a dedicated mining hardware. The miner that first finds a solution for its block is the winner. His candidate block becomes the new block in the chain. Because transactions are added in the mining block as they arrive, we can say that the latest block in the Blockchain contains the latest transactions.

When a new block is created (mined) it is time-stamped and propagated to the network. Every node receives the block, validates it, validates the transactions in it, and adds the block to his local Blockchain copy. The transactions included in the block become authorized and non-reversible part of Blockchain in the moment the block is accepted by majority of the nodes. Blocks can also be inspected as a way of transactional and financial clearing. Valid transactions are approved in groups, at certain periods of time. This solution was done to avoid conflicts and solve the double spending problem. In addition to transactions, every block stores some metadata and the hash value of the previous block. So every block has a pointer to its parent block. That is how the blocks are linked, creating a chain of blocks called Blockchain.

The ledger is publicly available for everybody to inspect the blocks and the transactions within. However, the users stay anonymous, identifying only by using their public key as an address. The transactions are encrypted too. Invalid transactions are rejected and are not included in the blocks. Attempts for malicious changes in the transactions will require repeated calculation of the proof of work for the attached block and all the blocks afterwards. These calculations are infeasible unless majority of the nodes in the network are malicious.

Using the Blockchain transactions, users can store or directly exchange the assets they own. In case of Bitcoin the asset is a digital currency, but the Blockchain transactions are not limited to that application. They can represent physical or digital property, smart contract between two or more parties, or any other data and document. Blockchain combined with smart contracts can be used to build a new generation of transactional applications that establishes trust, accountability and transparency at their core, while streamlining business processes and legal constraints.

In the following chapters we will review possible Blockchain solutions that can change the current model in which organizations collect and control massive amounts of data. In the second chapter we will present attempts for

978-1-5090-3843-5/17/$31.00 ©2017 IEEE

securing personal data using Blockchain and in the third chapter we will open the never resolved question for digital property in todays Internet. Then, in the fourth chapter we will see how Blockchain can redefine IoT implementations. In the end, we will focus on the healthcare data, as one of the most representable applications of Blockchain.

II. Decentralized Protection of Personal Data

Losing control of privacy is what happens when social media networks are constantly collecting users' personal data, actions and habits. While there are benefits of the data-driven services offered on personalized preferences, users still don't have a clear preview on which precise data is collected and for what purpose. Users lose total control of what happens with the data afterwards and they cannot withdraw the permissions. Usually there is a privacy setting page on most of the social media sites, where users can limit what other people see about them. What they cannot control and configure is what the social media corporation sees. People are used to privacy agreements provided in a non-user-friendly way, explaining superficial aspects about collecting personal data. But, collecting personal data doesn't stop within the site. Social Medias are following users' web browsing interests and the interactions they do with other web or mobile applications. The mobile apps that users install are collecting even more sensitive data like contact lists and location. Users are required to just accept the third-party access when installing a mobile application with no detailed information or option for partial acceptance.

Concerns about privacy of data grow as we are faced with the consequences of what others have seen or learned about us. In recent study about the state of privacy done in light of numerous high-profile data breaches, 74% of the respondents say it is "very important" to them that they be in control of who can get information about them, and 65% say it is "very important" to them to control what information is collected about them [2]. In the same study, fully 91% of adults agree that consumers have lost control of how personal information is collected and used by companies.

A proposed solution by Zyskind et al. [3] refers to the common privacy issues like data ownership, data transparency, auditability and fine-grained access control. This solution is an access-control management system that focuses mostly on mobile platforms and the inability of the user to revoke granted access to her private data. By installing a mobile app the permissions are granted indefinitely and the user has to uninstall the app and stop using the services if she wants to revoke the access. The goal of the new solution is for the user to be able to control and audit which data are stored and how they are used. As we mentioned, access should be revocable. So the technical idea is to store access policies to personal data on the Blockchain and then, let the Blockchain nodes moderate access to a DHT (distributed hash table).

The solution is composed of three entities: the user, the company that provides the service and the Blockchain. When the user wants to grant or revoke access of her personal data, Blockchain comes to actions as a mediator. Here Blockchain supports two transaction types: transaction for access and transaction for data. These transaction types are allowing access control management, data storage and data retrieval. When a user installs new application, a shared identity is created and sent to the Blockchain along with the configured permissions by the user's wish. All the granted permissions are listed in the so called Policy. The shared keys (the user's public key and the service's public key) and the policy are sent by an access transaction in the Blockchain.

The Bitcoin's Blockchain uses public key identity mechanism. All nodes in the system have public keys that are also called addresses. By identifying with the addresses, users act with pseudonyms, staying anonymous. In the proposed system, a new compound identity is introduced. The compound identity is a shared identity between the user and the service. The user is the owner of the key and the service is a guest. The compound key is composed of the signing key-pairs of both parties, so that the data will be protected from all the other parties in the system, except from the owner and all his guests.

The sensitive user data is encrypted with the shared encryption key and sent with a data transaction for storage. Blockchain sends the data to an off-blockchain key-value store (the DHT) and keeps only the hash value as a pointer to the data. The value set on the DHT is encrypted by the compound key. The pointer to the value is made known to the user and the service. The DHT only fulfills already approved read and write functions. Both, the user and the service can query the data using the pointer to the data. Every time the service accesses the data, its permissions are checked against the last access transaction. The user can revoke the permissions any time, or modify it, by initiating a new access transaction. To keep track of this, a web dashboard showing the user's current permissions can be easily developed.

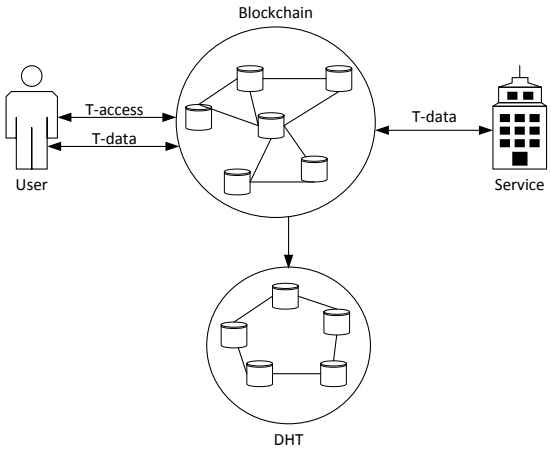

Figure 1 – Overview of Decentralized Permissions System

An adversary controlling any number of DHT nodes cannot compromise privacy of the sensitive data because they are encrypted. If the adversary obtains only one of the keys, the data is still safe. Because the personal data are not stored centrally, we don't require trust from a central institution. Furthermore, instead of direct access the system can use secure Multiparty Computation MPC protocol. This will be a better approach that will run computations directly on the network and obtain the final results instead the raw data. All the transactions of the service's requests are traceable, so the user can audit the frequency of the access.

Another research about a user controlled social media based on Blockchain is Ushare [4]. The vision of Ushare is that users should own their online presence by tracking the posts that they share and controlling the possibility for re-sharing. Using the P2P capabilities, Ushare creates a decentralized content distribution network. The asset that is managed by the Blockchain in this case is the data users post. The proposed

Ushare solution consist of: a hash table with encrypted content shared by a user, a system for controlling the maximum number of shares performed by user's circle members, a local personal certificate authority (PCA) that manages the user's circles and the Blockchain.

When a user shares some post with her circle, her PCA encrypts the data using the circle's public key. The encrypted data is stored into a distributed hash table. This DHT has three columns allowing the users to share the post that they see. Every time a user shares some post, the first column records a hash of the encrypted data of the post she sees and shares. The second column records the hash of encrypted data with her circle's public key. The third column stored the data item she encrypted. The reason for using the DHT in this second solution is the same - large size data like documents, images and videos need to be stored in a non-centralized way. The Blockchain stores only the transactions about the users' shares. It cannot store the actual data because downloading a full chain to all nodes will produce computability and time constrains. When a user creates a post, she sends new transaction to the Blockchain with her identity, the hash key of the encrypted data and a token that specifies allowed number of shares. Next, the user sends a separate transaction to every member of her circle with the encrypted data hash key. If another user who received the post, wants to share it, she send new transaction with her identity and the key of the data encrypted with this new user's circle key. Again, multiple new transactions are sending to the next users who can preview the re-shared transaction. The token number is descending with every share.

All the efforts for creating these two Blockchain solutions are because personal and sensitive data are not trusted in the hands of third-parties. As users create and post data, they should remain main owners of the data too. Regarding the surveillance that is done by tracking users' actions and interests, the users should at least have knowledge about it and certain benefits. The Blockchain can be a filter for permissions of accessing private data or it can implement a whole decentralized social network, as shown in the second solution.

III. Digital Property

A very specific implementation of Blockchain in the area of Big Data is associated with intellectual property of digital art. Artists, designers and creative workers can share easily on the Internet, but keeping the right with proper attribution or getting fairly compensated has proven difficult in the digital world that we know. There is not a transparent way to own something that can be so easily copied and fully replicated, with no sign of the original. A project called Ascribe is stating that the Internet was built with a crucial flaw regarding the ownership question [7]. Once the work is put online, or even sold online, the author loses control. That is why Ascribe's vision is to build the ownership layer of the Internet for digital content. They are creating a tool for authorship that will obtain visibility into where the arts work spreads.

The authors, the channels and the consumers on the Internet are practically in a mess. Creators don't have a stable way of getting fairly paid, and at the same time the consumers face content that is not available to buy. The process to license and legally sell the content is a painful one even for the distribution channels [8]. Common solution for sharing videos, movies, music, images and photographs, 2d and 3d digital graphics is still not found.

So, why is this the case? The World Wide Web started its work with simple hyperlinks and set the base for no visibility of what is the original and what is copied. After that, people thought of a way to contribute the author by mentioning it in the references and crediting the pictures they copied and used. But this system is far from flawless. People can always find a way to copy stuff and not give attribution, and the author will not know and will not be notified. Or, people can attribute the author but it's in one directional link, so the author will still not know that someone referenced on her work. Or worse, people can attribute someone else who is not the original author. So, this leaves a feeling that the main tool for digital content and digital sharing (the www) overlooked the need for digital property. But history says different. Project Xanadu was the first hypertext project, founded in 1960 by Ted Nelson [10]. This problem was approached then, by introduction to a royalty-based publishing scheme in a system that will provide storage and publication services. Author's attribution was built in this system. Bi-directional links were to be automatically set every time someone used other user's data. It turned out that this was a complicated, non-feasible technology, and the project was shut down.

Now, with the Blockchain technology, Ascribe tries to achieve the Xanadu goals by finding a solution for the digital property registry and the visibility of the copies. Regarding the visibility, their try is to find all the copies of the protected content that exist on the Internet. This can be done by crawling the entire internet and performing similarity match against creator's content. This is a machine learning similarity search problem. When the copies are found, the system performs automatic bi-directional links. Then it's up to author to decide if she will be asking for licensing fees or maybe a takedown request.

When it comes to selling intellectual property of digital art, it's not just selling a copy, it's selling the ownership and the right to use, modify or resell the content. To do this kind of ownership selling (not licensing selling) requires doing a legal contract, hiring a lawyer etc. The idea of using Blockchain to store and sell ownership of digital data will be simple as sending an e-mail with a signature that the user transfers the ownership of her content. The terms of service that Ascribe provides are done in consultation with specialized lawyers. The complexity of the legal licensing and ownership processes is handled by just accepting the terms of service. Blockchain is publicly displayed trusted ledger and it will secure all users' copyrights. Time-stamping of the transactions can be used as evidence in court, in case of an ownership dispute.

Regarding the Blockchain implementation, Ascribe made its own protocol called SPOOL - Secure Public Online Ownership Ledger [11]. This protocol was made specifically for documenting transactions relating to ownership of digital property. Ascribe lets the artist set a fixed number of editions of a work which can then be transferred, guaranteeing each edition is authentic and from the artist. So when doing transfer transactions, the user can transfer ownership for one or many editions. The editions are under one work that is stored into BigchainDB [8], and hashed for the use in Blockchain. So when transferring ownership, the user creates a transaction for one of her works, includes the hash value, the edition and the new owner. She signs the transaction and sends it. Because Ascribe uses the Bitcoin's Blockchain, public explorers who know the work's hash can track the work's ownership and find all the addresses that own each edition.

Next implementation of Blockchain called Monegraph is a proof that we can build new, modern and user-friendly digital marketplace. The Monegraph name comes from the name "Monetized Graphics" as they are helping artists and owners claim the rights and commercial value of their digital media. From users' perspective, with Monegraph, it's easy to buy and sell fully licensed digital media directly with terms, rights and prices the authors control. Monegraph facilitates the process of licensing and receiving income for arts in many digital forms created by photographers, designers, illustrators and other media makers [12].

Monegraph allows the authors to create and customize a license contract that establishes the usage parameters for their media. There are four licensing types: artwork license is for non-commercial use, photo news license is for editorial, product image license is a commercial license and snapshot is license that gives all the right. The authors have public catalog like a portfolio of their work that is publicly available and for sale. The Blockchain keeps track of ownership history. But, keeping only the info about the owners may not be enough in a marketplace. There are a lot of sales contract data and products metadata that are equally important as the product. Because of size concerns we cannot include this data in the Blockchain. That is why Monegraph sees the need of a Blockchain ecosystem that can be implemented for other digital markets too. What the solution needs is an integration with other services. So the ownership data are kept in the Blockchain to remain trustfully traceable and irreversible. But other documents associated to the product can be stored in a document database like MongoDB or CouchDB. The documents can be public, weather encrypted or not. Regarding the digital art itself, it can be stored to a documents repository available by a HTTP or P2P access. For example the file can be stored on Amazon Simple Storage Service (S3). The ecosystem should find a way to link the Blockchain, the documents repository and the digital art storage.

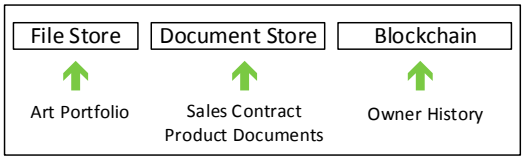

Figure 2 – Blockchain Ecosystem

The last property system that is going to be reviewed is called "Bitmark" and it allows transfer of both digital and physical objects [13, 14]. But storing information about physical objects like car, computer or a house, as an abstract property comes to a challenge how to record and organize the asset. We can easily have a fingerprint for digital data by applying a hashing algorithm over it, but when it comes to physical objects, there is new solution called "ObjectMinutiae" [15]. This is a framework for identification of physical assets based on unique surface-level texture patterns. Every asset in the Bitmark registry first gets recorded with its fingerprint, metadata and the registrant signature. New bitmark is created when the asset is set a new owner. Then the bitmark ownership can change by an ownership transfer transaction.

No matter what kind of solution is found now, the Internet has left a lot of materials with lost authorship. But, looking to the future, Blockchain can provide secure proof of the ownership by storing hash value of the digital art in a time-stamped transaction. Both Ascribe and Monegraph are using the Bitcoin network to keep proof of ownership of the digital art. They provide ways to verify the authenticity of artwork online and in real-time. Although Ascribe focuses more on tracking of sharing and selling of the digital art, Monegraph explores an ecosystem for licensing of the digital art. Still the concept is the same: the owner will possess the private key and the original copy of the hashed art. No-one can proof otherwise and no institution can change the data. The author can sell the art, and the new owner will now have the private key to the art. That is all the digital ownership has required, and now is possible.

IV. Internet of Things

The Internet of Things (IoT) is a fascinating developing concept, but there are major challenging aspects when it comes to having a secure ecosystem encompassing all building blocks of the IoT architecture. Using the present known technology to build IoT systems resulted with diverse protocols that were complex and with conflicting configurations. Current IoT ecosystems rely on centralized server-client paradigm. All devices are identified, authenticated and connected through cloud servers. A connection between devices goes through the Internet. Even if this solution work fine for now, it may not be able to respond to the needs of the larger IoT ecosystems in the future [16]. Continuing IoT development into more decentralized way was seen as true direction, but most of the technology was missing e.g. privacy and security in huge IoT P2P networks. The Blockchain technology may become the ideal component and fundamental element for tracking billions of connected devices, processing transactions and coordinating the devices. Blockchain will allow peer-to-peer messaging, file distribution and autonomous coordination between devices with no need for centralized cloud. In a vision of one IBM report [19] the Blockchain is the framework that is facilitating transactions and coordination among the devices. Each device will own its role and manage its behavior in the new Internet of Decentralized and Autonomous Things.

How IoT devices work now, is that that they are controlled by the user from a central point. The central point can be the user's mobile device. All the actions, commands and the rules are set by the user. While this is good for personal control, it is not automated in many ways. The real revolution can happen if all the devices get controlled by the Blockchain instead of the direct user control. This is possible using smart contracts. The smart contract is a set of conditions and business rules that must be met before a transaction is included in the Blockchain. The transaction can that is written in the Blockchain can be more complex than just an ownership transferring. Smart contracts have an integrated mechanism for conducting different contract types between the nodes. The smart contract is also autonomous and technically speaking it is a computer code that can be self-maintained and self-executed [21]. Once it comes to force, no human factor is needed to control it. Executing smart contracts is made possible by Ethereum which is a platform for creating Blockchain systems. Ethereum has its own network, nodes and miners, just like Bitcoin. But the Ethereum nodes are capable to execute every type of contract that comes to them [20].

Slock.it is a first implementation of IoT and Blockchain using the Ethereum platform [22]. So called Slocks are real-world physical objects that can be controlled by the Blockchain. They use the Ethereum Computer which is a piece of electronics that brings Blockchain technology to the entire home, making it possible to rent access to any compatible smart object and accept payments without intermediaries.

Slock.it is enabling anyone to rent, sell or share anything - without middlemen. With Slock.it, there is a vision for sharing economy in witch services like Airbnb apartments become fully automated. It functions in the following way: the owner of a smart item (Slock) creates a smart contract for its usage setting the price and the deposit. Users can find the Slock and then make a payment on the Ethereum Blockchain, thereby gaining permission to open or close that Slock, meaning to use it as agreed. The smart contract is automatically enforced, with the deposit returned to the user minus the cost of the rental. Practically this means that by installing smart lock on the apartment doors, the users can rent an apartment on the Blockchain. The smart contract will unlock it and make it available as the contract specifies. Beside smart doors, this system allows to rent, sell or share any smart object that has the Slock.it technology embedded. Bikes, cars and any object that can be secured by a physical lock is a potential use case.

As some researches think that the most secure approach to develop future IoT applications is on top of a stable existing Blockchain like Bitcoin [17], others argue that the Blockchain we know requires expensive computation and time delay before the transactions are mined into blocks. That is why in a paper by Dorri at al. [18] the authors propose a new secure, private, and lightweight architecture for smart home IoT, based on the Blockchain technology. Here, mining the blocks is seen as the first problem because IoT devices are resource restricted devices and cannot perform such an operation. Also IoT devices should act in the same second something is detected or ordered. The required time for blocks mining, in most of the cases cannot be acceptable. The IoT network will contain billions of devices. This figure presents much more nodes than Blockchain has experienced so far. That is way the proposed solution [18] contains three tiers: the local network, the overlay network, and the cloud storage. The local network contains all the smart home objects and a local computer acting as a local Blockchain that is constantly online. This local Blockchain is centrally managed by its owner. When there is a new smart device in the home, the user adds it in the Blockchain. All transactions related to a particular device are chained together. There is no standard mining, so when a transaction is received it is automatically put in a block and seen as valid. The overlay network is a peer-to-peer network that connects more smart homes and users. This network manages public keys of users allowed to access data for the smart homes, and public keys of smart homes that provide accessible data. The cloud storage is included as a solution for devices that may wish to store data in the cloud, so that a third party can access the data and provide certain smart services.

To conclude, some of the IoT challenges didn't even have a vision until Blockchain appeared. Using Blockchain, IoT can transition toward a network of devices that can interact with each other and with the environment without human intervention. The devices will also make smart decisions so many workflows will be automated in new ways, achieving significant time and cost savings.

V. HEALTHCARE

Big data in Health comes from a variety of sources, such as clinical trials, electronic records, patient databases, medical measurements and imaging. All this data comes in a wide range of formats and from different data streams. The data has to be assessed and interpreted in a timely manner to benefit patients. But clinicians need new tools to track, trace, and provide fast feedback for individual patients. Proper data management will also help the prediction strategies, interventions, health services, and health policies. Since the medicine is always in step with the technology, leveraging many innovations, we can conclude that healthcare data is in a need of crucial transformation, just as the Big Data is.

Tal Rapke in his visionary approach towards Healthcare [23] discusses that maybe change should come from a point where people own and can access their health and life record. Blockchain with its concept of work has a way to bring this consumer centered approach to the health sector. Data from results and procedures can be stored on the Blockchain that will not rely on one central storage facility. This will help governments and other enterprises be liberated from the liability that data is becoming. At the same time, data will reside on the latest secure technology and using verifiable cryptography. As owners of the data, the consumers will be empowered to decide who they share their data with. Health will be oriented more toward the consumer, but still in balance with the other important players in the health system.

In another paper by Gupta et al. [24], it is explained how Blockchain could enable an interoperable and secure electronic health records exchange in which health consumers are the ultimate owners. The proposed scenario is to store only the metadata about health and medical events on the Blockchain. Otherwise the Blockchain infrastructure will have to scale massively to support complete health records. So, metadata such as patient identity, visit ID, provider ID, payer ID, etc. can be kept on a Blockchain, but the actual records should be stored in a separate universal health cloud. For example, if a patient visits two hospitals today, they will store data about him in two databases that the patient does not own. If the hospitals have to communicate, they will use a mediator for a standardized communication like web services, e-mails or shared files repository. In a scenario where Blockchain is applied, the first hospital creates a record on the universal health cloud. Then the hospital creates a Blockchain transaction with the visit metadata and a URL to the record in the cloud. The patient signs this transaction with her key. When the patient now visits the second hospital, she must provide her key in order to read the Blockchain transactions. Only stakeholders with the patient's key can decrypt the transactions. So this is an example of how people can own the data and authorize needed access. Even smart contracts can be encoded into the blocks to carry instructions about insurance, emergency contacts, wills, etc. These smart contracts will be activated by events that the Blockchain can read from another web service.

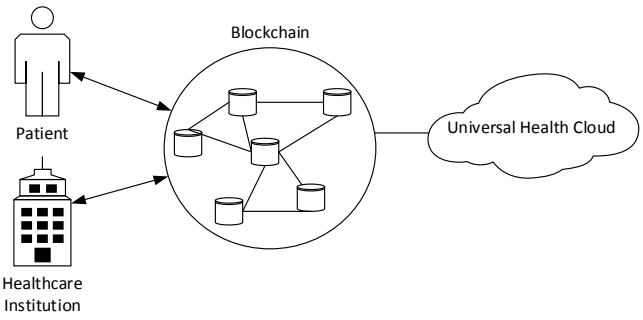

Figure 3 – Overview of Blockchain Healthcare System

Another research agrees that there is no sense in replicating all patients' data into Blockchain nodes [25]. Transactions in the blocks should contain a user's unique identifier, an encrypted link to the health record and a timestamp for when

the transaction was created. The transaction can also contain the type of data that is stored. Depending on the implementation, this could help querying and processing the accessed data. This Blockchain will contain history of all medical data, including formal medical records as well as data from mobile applications and wearable sensors. Keeping the data away from the Blockchain, in a Cloud pool, it could represent a good base for queries, mining, analytics and machine learning. This kind of analysis will not influence any patient privacy. This data would be encrypted and digitally signed to ensure privacy and authenticity of the information. The user would have an option to assign a set of access permissions and designate who can query and write data to his Blockchain. Further development of a user-friendly interface for the patient to review her health data and manage access privileges is totally doable.

Another case study for Blockchain in Healthcare, utilize Ethereum's smart contracts to create representations of existing medical records [26]. These contracts as we know are stored directly within individual nodes on the network. The proposed solution called "MedRec" structures the large amount of data into three types of contracts. The first one is Registrar Contract. It stores the participants' identity with all the needed details and of course, the public keys. This kind of identity registration can be restricted only to certified institutions. The second contract is the Patient-Provider Relationship Contract. It is issued when one node stores or manages data for another node. The main usage will be when there is a smart contract between the care provider and patient. The last one is Summary Contract which helps the patient to locate her medical history. As a result of this contract, all previous and current engagements with other nodes in the system are listed. MedRec also proposes a mining model that engages the healthcare community in the mining process. Medical researchers and health care stakeholders can mine in the network. The mining reward in fact can be some approved access to aggregate anonym's medical data.

To sum up, Blockchain offers a future that promises to center the individual in healthcare and help patients discover and manage their own medical records. But, a global standard is needed to store, access and share encrypted data on the cloud. Looking to the future, all of the proposed solutions have the potential to engage millions of individuals, health care providers and medical researchers. The solutions could influence enormous advancement in medical research.

VI. Conclusion

Blockchain presents many promises for the future of Big Data. The first one is that in many areas, users could be in control of all their data and transactions. They can trust that transactions will be executed exactly as the protocol commands removing the need for a trusted third party. This concept can influence Big Data to find a solution for storing and managing data in a distributed manner on a P2P network. Blockchain technology can be a new part of the surrounding ecosystem of tools that Big Data uses. Actually it can play a crucial role in security for user authentication, restricting access based on a user's need, recording data access histories and proper use of encryption on data.

Some challenges still remain, such as consensus models, the computational costs of mining blocks and validating transactions. Also, Blockchain applications offer solutions that require significant changes or complete replacement of existing systems. That is why the transition will not be easy and fast. But we are still in the early stages of Blockchain development, and these obstacles will eventually be overcome, opening the path for many exciting possibilities.

References

[1] Satoshi Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System", 2008

[2] Lee R and Maeve D, "Privacy and Information Sharing", Pew Research Center, 2016

[3] Guy Zyskind, Oz Nathan and Alex 'Sandy' Pentland, "Decentralizing Privacy: Using Blockchain to Protect Personal Data", Security and Privacy Workshops (SPW), 2015 IEEE

[4] Antorweep Chakravorty and Chunming Rong, "Ushare: user controlled social media based on blockchain", International Conference on Ubiquitous Information Management and Communication, 2017

[5] Thanh Bui and Tuomas Aura, "Application of Public Ledgers to Revocation in Distributed Access Control", 2016

[6] Ahmed Kosba, Andrew Miller, Elaine Shi, Zikai Wen, Charalampos Papamanthou, "Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts", Security and Privacy, IEEE, 2016

[7] Trent McConaghy and David Holtzman, "Towards An Ownership Layer for the Internet", ascribe GmbH, 2015

[8] Trent McConaghy, Rodolphe Marques, Andreas Muller, Dimitri De Jonghe, T. Troy McConaghy, Greg McMullen, Ryan Henderson, Sylvain Bellemare, and Alberto Granzotto, "BigchainDB: A Scalable Blockchain Database", 2016, ascribe GmbH, Berlin, Germany

[9] "PRS for Music takes legal action against SoundCloud streaming service", The Guardian, 2015

[10] Roy Rosenzweig, "The Road to Xanadu: Public and Private Pathways on the History Web", The Journal of American History, 88.2, 2001

[11] Dimitri de Jonghe, "SPOOL Protocol", https://github.com/ascribe/spool, 2016

[12] Melanie Swan, "Blockchain Blueprint for a new Economy", O'Reilly Media, Inc, 2015

[13] Christopher Hall, Casey Alt, Lê Quý Quốc Cường, and Sean Moss-Pultz, "Bitmark: The property system for the digital environment.", 2016

[14] Casey Alt, Sean Moss-Pultz, Amy Whitaker, and Timothy Chen "Defining Property in the Digital Environment", Bitmark 2016

[15] Tzu-Yun Lin, Yu-Chiang Frank Wang, Sean Moss-Pultz, "ObjectMinutiae: Fingerprinting for Object Authentication", 2015.

[16] "Internet of Things: Privacy & Security in a Connected World", FTC Sta¬ff Report, 2015

[17] Marco Conoscenti, Antonio Vetro, Juan Carlos De Martin, "Blockchain for the Internet of Things: a Systematic Literature Review", 2016

[18] Ali Dorri, Salil S. Kanhere, and Raja Jurdak, "Blockchain in Internet of Things: Challenges and Solutions", 2016

[19] "Device democracy-Saving the future of the Internet of Things", Executive Report, IBM Institute for Business Value, 2015

[20] Dr. Gavin Wood "Ethereum: A Secure Decentralised Generalised Transaction Ledger", 2014

[21] Vitalik Buterin "A next generation smart contract & decentralized application platform", Ethereum White Paper, 2016

[22] Christoph Jentzsch "Decentralized Autonomous Organization to Automate Governance", 2015

[23] Tal Rapke, MD "Blockchain Technology & the Potential for Its Use in Healthcare", 2016

[24] Nitesh Gupta, Anand Jha, and Purna Roy, "Adopting Blockchain Technology for Electronic Health Record Interoperability", 2016

[25] Laure A., Linn Martha B., Koo, M.D, "Blockchain For Health Data and Its Potential Use in Health IT and Health Care Related Research", 2016

[26] Ariel Ekblaw, Asaph Azaria, John D. Halamka, MD, Andrew Lippman, "A Case Study for Blockchain in Healthcare: "MedRec" prototype for electronic health records and medical research data", 2016

[27] Andreas M. Antonopoulos, "Mastering Bitcoin", O'Reilly Media, Inc, 2015