Technology against COVID-19

A Blockchain-based framework for Data Quality

Imane Ezzine
Ecole Mohammadia d'Ingénieurs
Mohammed V University in Rabat)
Rabat, Morocco
imaneezzine@research.emi.ac.ma

Laila Benhlima
Ecole Mohammadia d'Ingénieurs
Mohammed V University in Rabat)
Rabat, Morocco
benhlima@emi.ac.ma

Abstract—The effects of COVID-19 have quickly spread around the world, testing the limits of the population and the public health sector. High demand on medical services are offset by disruptions in daily operations as hospitals struggle to function in the face of overcapacity, understaffing and information gaps. Faced with these problems, new technologies are being deployed to fight this pandemic and help medical staff governments to reduce its spread. Among these technologies, we find blockchains and Big Data which have been used in tracking, prediction applications and others. However, despite the help that these new technologies have provided, they remain limited if the data with which they are fed are not of good quality. In this paper, we highlight some benefits of using BIG Data and Blockchain to deal with this pandemic and some data quality issues that still present challenges to decision making. Finally we present a general Blockchain-based framework for data governance that aims to ensure a high level of data trust, security, and privacy.

Keywords—Covid-19, Blockchain, Big Data, Data Quality, data governance.

I. INTRODUCTION

COVID-19 has shaken the whole world, not only in the health system, but also in the economy, education, transport, politics, etc.

Researchers, businesses, and innovators from around the world have come together to find a way to stop the spread of this pandemic by using innovative technologies and putting it at the service of governments and authorities in order to track and contain the outbreak.

In recent weeks, many initiatives to set up tracking applications, dashboards on the state of the spread of the virus worldwide, mobile health self-monitoring application and others, have been implemented and demonstrated their effectiveness in this fight. These applications are based on innovative technologies such as Blockchain, Artificial Intelligence, Big data and other more.

In fact, Blockchain has been deployed in patient tracking applications during the confinement, supply and delivery of drugs, etc. On the other hand, Big Data found the perfect fields to reveal itself via analytics, the collection of data from heterogeneous databases and with different formats and even more to combine this data to have dashboards about the spread of the virus in different parts of the world.

Many papers have been published online in recent months speaking about these technologies and their roles in the fight of COVID-19. But, there is a concern that prevents to declare the total effectiveness of these applications, it is the Data Quality. Now with the big data emergence, the volume of data is increasing, data is being produced at an increasing velocity, data types and formats have more variety, and data veracity is becoming more uncertain. To tackle these challenges, we propose a framework based on Blockchains to improve data quality. This framework will help ensure a high level of data trust, security, and privacy.

The rest of the paper is organized as follows. In the next section, we discuss the benefit of using some technologies such as Big data and Blockchain where they were used to deal with the COVID 19. Then, in section III, we are going to discuss some issues with the data quality that researchers may face in their fight against COVID-19. Then, in section IV, we present a use case about improving big data quality. In section V, some related works are presented. After that, we present our framework in section VI. Finally, we end with a conclusion and future works.

II. BIG DATA AND BLOCKCHAIN TO DEAL WITH THE CORONAVIRUS

A. Blockchain achievement in this pandemic

The Blockchain technology is a distributed registry that acts as a shared database, keeping all of its copies synchronized and verified. In a recent work [3], the Blockchain is formally defined as blocks containing messages, proof of work and a reference of the previous block and stored in a shared database, which is capable of carrying out transactions on the P2P network by keeping irreversible historical records and transparency".

In the context of health data management, some researchers have used this technology to reduce the propagation of the COVID-19 pandemic in the World and to be aware against a future pandemic.

Blockchain technology has the following advantages: the ability of a blockchain's registry to remain unchanged and indelible, to keep authorized users responsible for any transaction and to share data with appropriate authentication without third party intervention. Thus, it was wise to adopt such technology to keep track of health data related to COVID-19 and also to search for and contain infected people while bringing more confidence and integrity to this data that users use and share [4].

Several applications are using blockchain in these pandemic times. Here are some examples from around the world:

- Tracking Infectious Disease Outbreaks :

  The blockchain are used to track the surveillance of public health data, especially for epidemics of infectious diseases such as COVID-19. With increased transparency of the blockchain, this translates into more accurate reporting and effective responses.

2020 6th IEEE Congress on Information Science and Technology (CiSt) | 978-1-7281-6646-9/21/$31.00 ©2021 IEEE | DOI: 10.1109/CIST49399.2021.9357200

Blockchain can help develop treatments therefore early detection of symptoms before their spread to epidemics. In addition, it allows government agencies to keep track of the virus activity, patients, new suspected cases, etc[30].

- Donation Tracking :

  Blockchain is a solution for Trust in the case of major donation problems. With the help of blockchain capabilities, donors can see where the funds are most urgent and can track their donations until they receive notification that their contributions have been received by persons[2].

- Management Crises

  Blockchains could instantly alert the public to the coronavirus by global institutes like the World Health Organization (WHO) and could help provide governments with recommendations on how to contain the virus. It could provide a platform where all concerned authorities such as governments, health professionals, the media, health organizations and others can keep each other informed of the situation and prevent it from getting worse[2].

- Securing Medical Supply Chains

  Blockchain has proven effective in managing the supply chain in various industries [8]; similarly, it could be beneficial for tracking and tracing medical supply chains. Blockchain-based platforms can be useful for examining, recording and tracking the demand, supply and logistics of epidemic prevention equipment without allowing anyone to follow the process. This technology could help streamline medical supply chains, ensuring that doctors and patients have access to the tools when they need them, and preventing contaminated items to reach stores[29].

- Fake news

  Blockchains have been used to remedy the problem linked to fake news that is spreading across the globe and that constitutes a new challenge for the media environment, but also for businesses. In fact, quite a few companies have developed blockchain-based solutions to allow companies to guarantee the accuracy of the information they disseminate[2].

B. Big data against COVID-19

Big Data has been hailed by experts as one of the leading resources in the fight against COVID-19. Commonly, big data is the information asset characterized by such a high volume, velocity and variety that require a specific technology and analytical methods to extract a useful information to serve the decision making.

In the COVID 19 context, Big data can refer to the patient data such as medical records, physician notes, X-Ray reports, case history, list of doctors and nurses, and information of outbreak areas. Experts have found uses for Big Data and Big Data analytics platforms for various purposes, such as:

- Containment control and Social Distancing Analysis :

  In addition to strict quarantine measures which kept people in place at the peak of the epidemic, some countries like China have made widespread use of big data to contain the virus through applications where citizen must scan a code on his smartphone and enter his information to show where he has been for the past two weeks. The information is added to a database which can be checked to confirm if he has completed quarantine or not [31]

- Better interconnectivity across national data systems

  In the light of the outbreak of the virus, many countries in the world, tried to track their citizens who were traveling and came back. So, they tried to link medical records on the national health insurance database with customs and immigration records to identify and test people who had recently travelled from China, sought medical care, or showed signs of severe respiratory illness[14]. In these circumstances, the use of Spatial Data Science and location-based data streams is more important than ever. Now, we have the opportunity to put forward this technology to serve this global quest and flatten the curve.

- Data visualization to track COVID 19 outbreak

  To track COVID-19's spread in real time, governments, scientific institutions and companies created many "dashboards" for visualization of the disease by making resources available, including funds and the opening of large-volume data repositories,. The most frequently used, is that of the John's Hopkins' Center for System Science and Engineering (CSSE)who provides real-time visualization [15]. In these circumstances, the use of data from heterogeneous sources and location-based data streams is more important than ever and thanks to these dashboards, governments can control, understand and predict the dynamics of distributed COVID-19 in time and space, moreover, it has given authorities perspectives to better plan and respond to the dramatic pressure exerted on health infrastructure, emergency systems and the global economic system

- Big data analytics

  A new trend has appeared, where computer scientists, biologists' researchers and doctors collaborate and take advantage of this explosion of data on COVID 19 from hundreds of thousands of medical records from coronavirus patients into effective treatments and predictive analytical tools that could help lessen or end the global pandemic. This analytics can help to implement large-scale COVID-19 investigations, develop comprehensive treatment solutions. This would also help healthcare providers to understand the virus development to find an effective vaccine[16].

The next section is about the most encountered data quality issues in this fight against the COVID-19.

III. DATA QUALITY CHALLENGES

Despite the innovation and the advancement of medical big data, policy makers should approach data with care, as the

data in circulation has questionable quality. Indeed, much data is still lacking and the available data may not be exact or reliable and may contain substantial uncertainty, concerning, for example, the precise timing and natural history of the cases [8].

Researchers have found that the big data collected and used in research on COVID 19, presents data quality problems, such as pravicy, security and trust. We briefly introduce each of these issues in what follows.

A. Data privacy

There is growing concern about the way governments are using data to respond to the COVID-19 crisis. While the government's efforts are directed to slow coronavirus outbreaks, there are also concerns that gathering information about people's geo-location and other personal data to aid management of the pandemic risks infringing on the person privacy more than ever before. In fact, some Governments gave the authorities the right to require telecommunication companies to use or access mobile phone location information without user consent to hand over data of people with confirmed infections to track their location. The data has enabled the rapid deployment of a notification system alerting people on the movements of all potentially contagious persons in their neighborhoods or buildings. [3][7] and [9]. No one fears "technology for good", nevertheless, we must not relax the basic privacy requirements : maintaining anonymity, encrypting data and preventing our information from ending up in the wrong hands.

B. Data security

At the IT level, data quality and security controls must be ensured. Weaknesses in data integrity, which are common when data from personal digital devices is used, can introduce small errors , which, in turn, can have a disproportionate effect on large predictive models ladder. In addition, data breaches, insufficient or ineffective anonymization, and biases in the data sets can become major causes of distrust of public health services[10][7].

C. Data Trust

The current COVID-19 pandemic raises important questions about opening, sharing and using data, and highlights the challenges associated with data trust. Without data, we cannot understand the pandemic. Only using good data can we know how the disease is spreading, what impact the pandemic has on the lives of people around the world, and whether the counter measures countries are taking are successful or not.

On the one hand, social media has become a conduit for spreading rumors, deliberate misinformation, and wrong data. Many perpetrators are using sites such as Facebook, Twitter, YouTube, and WhatsApp to create a sense of panic fake news and confusion in the circumstances of coronavirus. The pressing issue is fake news that spread more rapidly in social media than the ones from reliable sources and damages the authenticity balance of news ecosystem and eliminating trust in the data and the information [11]. On the other hand, sensor nodes or agents can produce permantely a large quantity of data items. These data items describe the properties of certain entities or events for example check of fever in coronavirus with thermometer. Due to the possible presence of malicious source providers and inaccurate knowledge generated by intermediate agents, the information provided to the data users could be wrong or misleading [12]. Relying on trustworthy sources is always good advice, but now it is an absolute must[13].

In the next section, we explain how works the technology of blockchain, and then we present a use case about the use of blockchain to improve big data quality.

IV. BLOCKCHAIN FOR DATA GOVERNANCE:

A. The mecanism of Blockchain technology

Blockchain is a technology that redefines trust in the new generation systems. It doesn't need mediators like corporations and governments, since it almost always come as central entities that receive, process and store the transactions. There are no mediators that process the transactions using correct business logic and have full control on data privacy and security, so the trust is decentralized. Users just need to trust the system and the smart code that is shared between all the participants. From technical point of view, Blockchain is a distributed database that exists on a P2P network where every node in the network is at the same level as all the other nodes. Although nodes can come in many forms but there is no central node that is an authority. Every node stores a local copy of the Blockchain. If consensus of nodes agrees upon transaction's validity, then the transaction is considered valid. When a transaction is created, it has to go through validation and confirmation stages before it enters the Blockchain then it is broadcasted to the network. P2P nodes share the transaction between themselves almost in a real time. If valid, the node saves the transaction into its transaction pool. If not, it is immediately removed. Some of the nodes, called "miners", take all the available transactions from the pool and include them in a new candidate block. A proof of work concept calculates a random hash, which is a generated string, that is computed using the data as input, of the candidate block. The correct hash value must satisfy a defined difficulty target. This number is calculated using all block's metadata including the hash of the previous block. This is the key to Blockchain security. If someone tries to change a transaction from the past, the hash value of the block that contains the transaction must be calculated again and all hash values for the blocks that come afterwards must be calculated again too. This can't be done, unless more than half of the nodes in a network are infected. Once a new block is created, it is broadcasted to the network. All nodes receive the block, validate it with all the transactions in it. Transactions that are included in the created block, are then removed from the pool[21].

B. Fake news use case

The objective of this use case is to have as final result Blockchains containing hashes of data and pointing to a database containing this reliable data.

Fig.1 represent a process of detecting fake news. First, when receiving an information, we begin with the verification of the reliability of the source of this information. There are different criteria that can be taken in consideration to evaluate a data source reliability, such as : reputation, data update, data type (documents, streaming, images …etc). If the source is reliable, then we go to the next step; otherwise, the data is automatically rejected. The next step will be checking the data

accuracy by using algorithms based on Machine learning in order to detect the fake news. If it's not a fake news, this data will be stored and a Blockchain will be created with a hash of this data(Fig.2), otherwise, it's a fake news and it twill be rejected. As a final result, we will have a trusted Blockchain that point to the data stored in the database.

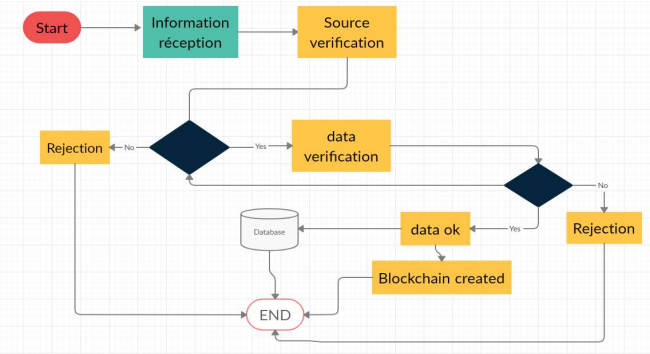

Fig.1: process of controlling fake news

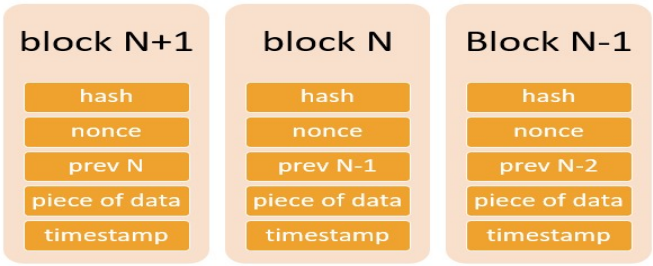

Fig.2: proposed Blockchain structure

C- Data governance for data quality

It is critical to verify the data integrity since the creation. In fact, the data is exposed to different processes, procedures, transformations and uses, which impact data quality. This data must be tracked and supervised to monitor the change of data across the process, otherwise poor data quality will create significant problems for companies and negatively impacts on business decisions. Data governance provides organizational approach to data and information management that formalizes a set of policies and procedures to encompass the full life cycle of data, from acquisition to use and to disposal [22]. It educates data consumers regarding data usage, meaning and quality levels, to build trust and encourage data utilization. But that trust can't be built without robust data quality controls to ensure reliable and accurate data and turn data into value. With increased data utilization among business users, data governance is even more critical to ensure that information is easily understood, appropriately accessed and effectively leveraged. Business users are increasingly using analytics to explore operational data, generate insights and drive business decisions. Consequently, companies should implement a data governance framework. These frameworks ~~contend~~ implement traditional checks for data quality to ensure the completeness, consistency and conformity of data [23]. In fact, adopting data governance framework, enables data understanding, which helps users trust their data. The data must be trustworthy, to generate reliable analytics results that enable the business to make good decisions. However, the existing data governance faces challenges in the light of Big Data cases.

In the next section, we present some related works on big data governance.

V. RELATED WORKS

With the emergence of the Big Data, in the last years, ushered in several new research, challenges and applications as well as exacerbated data quality problems [24]. Big Data is often characterized by five Vs [25], especially, variety and veracity that pose special challenges for data quality management in Big Data applications. In fact, data trustworthiness, privacy and security have been very quoted in the last few years [6]. Soares defines big data governance in a clear and comprehensive manner as follows: Big data governance is part of a broader information governance program that formulates policy relating to the optimization, privacy, and monetization of big data by aligning the objectives of multiple functions [20].

Usually, implementing data governance means improving decision making and to protect the needs of stakeholders [28]. in the few researches about big data governance, authors propose frameworks that consider implementing Strategy, organizations, policies processes and standard of organizations [27], [20],[32], others, that establish difference between traditional data and Big Data governance and the fact that is possible to use the same components (i.e. data quality, compliance, data life cycle, master data management, data privacy)[26],[33][34]and [35]. Big data faces challenges such as ensuring secure access to data, policies to govern in the light of fast generation of data, stream data, and making value from a trustful data. These studies didn't give importance to deal with similar issues.

To our knowledge, there are a very few studies on regulatory issues and big data governance [17]. Therefore, we think that it is wise to consider it as potential solution in order to improve Big data quality and specially data trust.

VI. OUR FRAMEWORK FOR BIG DATA QUALITY

The big data governance framework can present a way to improve the quality of data. In addition to the quality level of Big data, data governance will improve the strategy for the protection the private information like personal data and disclosure data alongside data security by setting up mechanisms against attacks.

In this research, and in order to improve the quality of big data, we opted to set up a Big data governance framework (Fig.3) that relies on a database based on Blockchains to guarantee data trust.

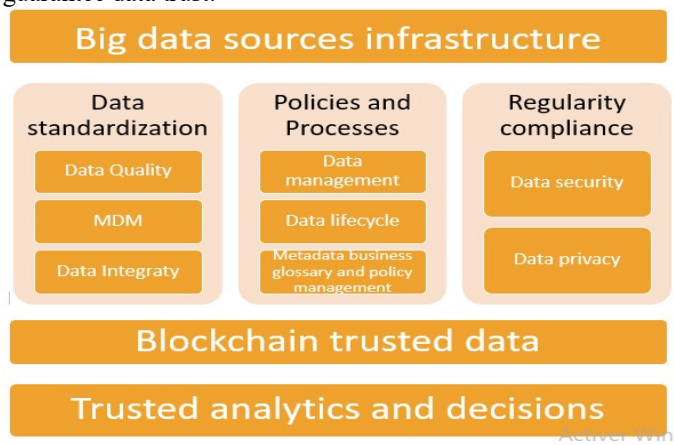

Fig.3: Our proposed Big Data governance framework

The proposed framework consists of the following layers:

- Big data sources infrastructure: The exponential growth in the number of data sources will continue to make it difficult to collect reliable information. It includes transaction data, data from social networks, content and machine data. This can lead to uncertainty, such as the origin, quality, source and accuracy of the data.
- Data standardization: This layer includes :
  - Data quality: this component is responsible for defining, monitoring and improving data accuracy, completeness and timeliness. It has the ability to parse, standardize, validate and match enterprise data.
  - Master data management: It refers to all the methods, tools, concepts and processes to ensure that the Master data is correctly identified, of high quality, free of errors and usable without any risk. Master Data in a central repository provides a single, authoritative, view of the data held by the company. It eliminates the costly inefficiencies of data silos. The MDM solution supports certain business initiatives with functions to identify, link and syndicate information on products, customers, stores sites..etc. For example, in the CRM system, customer can be considered as master data.
- Processes and policies: This layer includes:
  - Data management: It is an operational concept focused on the implementation and coordination of policies and procedures. Data managers manage essential data resources, including making data decisions, making recommendations, and developing policies.
  - Metadata, business glossary and policy management : define both metadata and governance policies with a common component used by all integration and governance engines.
- Data lifecycle management : it manages the existence of the data from its reception or creation until it makes more sense to keep these massive volumes and therefore the deletion and archiving of this data from the business system.Regularity compliance: These include privacy and security and involves hiding data in applications to protect sensitive data, monitoring repositories to prevent data breaches, protecting data from external or internal attack and take compliance. It must also be taken into account that appropriate policies and procedures must be followed (created and defined in the Processes and policies component) to prevent the misuse of Big data, taking into account regulatory and legal risks when managing social media, geolocation, biometrics and other forms of personally identifiable information. At this level, we are considering the use of mechanisms based on artificial intelligence and machine learning to allow high levels of cybersecurity and to detect fakes news.
- Trusted Blockchain data: this layer represent a data lake that contain the Blockchains created from trusted data collected. In fact, each data has been checked in all processes and rules established to realize the big data governance. We have to precise that if we store all the data in the Blockchain, we will end up with a problem of storage. The Blockchain stores only the hash of the data and parts of the data. If the hash changes that means that our data has been changed. Since the hash is very small, it takes small space in storage in the Blockchain, and the cost of transaction is relatively slow. We assign the id (hash) of the Blockchain transaction to our raw data. In addition, depending on the parts of the data stored in the Blockchain we get some transparency, because the data will be publicly accessible. In fact, the main goal is that these Blockchains will serve as a database, and guarantee the trustworthiness of the data. All occurring events are registered with accuracy, so the trustworthiness of the stored data will be good.

- Data analysis and decision making: Analytical applications rely on a Big Data platform to process and analyze information. In turn, the analytics engines of the Big Data platform rely on a reliable database in order to return precise and usable results and to integrate this information into other business systems. In our case, the database is build up from Blockchains that contain the data processed with the different processes developed in the other components.

The Framework will provide visible benefits such as:

- Reinforced security, obtained by locating critical data, identifying owners and users of data, assessing and correcting risks related to critical data
- Better data quality: by defining, monitoring, maintaining data integrity, the decisions for the company will be running based on accurate and complete information, rather than misleading or outdated data.
- Greater operational efficiency: thanks to processes and procedures allowing faster and easier data management
- Reduced data management and storage costs: Instead of having isolated software platforms, centralized control mechanisms offer the potential for optimizing the cost of data management (increasingly important in the age of data sets explosion)

- A decrease in the number of security breaches, thanks to better training on data resource management

VII. CONCLUSION AND FUTURE WORKS

As the world increases in its fast adoption of new technologies in the fight against the COVID-19, the use of Big Data cannot replace the development of efficient health infrastructures and the development of strict systems and protocols for examination and monitoring. However, AI and Blockchains based applications have been used with success in some countries. Nevertheless, remains the biggest challenge of ensuring security, privacy and trust in data. In the first part of this article we presented a set of different perspectives on key design elements, challenges, opportunities, and best practices for big data and Blockchain technologies.

While the proposed framework is not exhaustive, it does provide a basis to improve data quality in the field of Big Data This framework is a starting point. We are working to implement the Blockchain layer of the solution..

As future work, we seek to set up a method for evaluating data sources, and we will take stock of MDM policies by joining them with Blockchains.

REFERENCES

[1] *R. A. Addi, A. Benksim, M. Amine, M. Cherkaoui, " COVID-19 Outbreak and Perspective in Morocco" Electron J Gen Med. 2020.*

[2] *D. C. Nguyen, M. Ding, P. N. Pathirana,and A. Seneviratne,, "Blockchain and AI-based Solutions to Combat Coronavirus (COVID-19)-like Epidemics: A Survey," preprint.*

[3] *A. Azim, M. N. Islam and P. E. Spranger, "Blockchain and novel coronavirus: Towards preventing COVID-19 and future pandemics," published.*

[4] *T. P. Mashamba-Thompson and E. D. Crayton," Blockchain and Artificial Intelligence Technology for Novel Coronavirus Disease 2019 Self-Testing", published.*

[5] *H. M. Yassineaand Z. Shah, How could artificial intelligence aid in the fight against coronavirus?, EXPERT REVIEW OF ANTI-INFECTIVE THERAPY2020, VOL. 18, NO. 6, 493–49.*

[6] *Z. Allam and D. S. Jones , "On the Coronavirus (COVID-19) Outbreak and the Smart City Network: Universal Data Sharing Standards Coupled with Artificial Intelligence (AI) to Benefit Urban Health Monitoring and Management ," published*

[7] *B.Tang, N. L. Bragazz, Q. Li, S. Tang, Y. Xiao and J. Wu "An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov)"Infectious Disease Modelling,Volume 5, 2020, Pages 248-255*

[8] *S. Saberi, M. Kouhizadeh, J. Sarkis and L. Shen, 'Blockchain technology and its relationships to sustainable supply chain management, InternationalJournal of Production Research(2019), 57:7, 2117-2135,*

[9] *M. Ienca and E. Vayena,"On the responsible use of digital data to tackle the COVID-19 pandemic", Department of Health Sciences & Technology, Swiss Federal Institute of Technology in Zurich, Zurich, Switzerland.*

[10] *S. Tasnim, M. M. Hossain and H. Mazumder, "Impact of Rumors and Misinformation on COVID-19 in Social Media " J Prev Med Public Health. 2020;53 (3): 171-174*

[11] C. Dai, D. Lin, E. Bertino and M.t Kantarcioglu, « An Approach to Evaluate Data Trustworthiness Based on Data Provenance",Secure Data Management*, 2008, Volume 5159*

[12] *C. M. Pulido, B. Villarejo-Carballido, G. Redondo-Sama and A. Gómez,« COVID-19 infodemic: More retweets for science-based information on coronavirus than for false information" ,International Sociology2020, Vol. 35(4) 377 –392*

[13] *Q. Pham, D. C. Nguyen, T. Huynh-The, W. Hwang, and P. N. Pathirana, "Artificial Intelligence (AI) and Big Data for Coronavirus (COVID-19) Pandemic: A Survey on the State-of-the-Arts," IEEE TRANSACTIONS ON ARTIFICIAL INTELLIGENCE, April 2020, Preprint*

[14] *D. Buhalisa and R. Leungb, "Smart hospitality—Interconnectivity and interoperability towards an ecosystem",International Journal of Hospitality Management Volume 71, April 2018, Pages 41-50*

[15] *N. Naudé, "Artificial intelligence vs COVID-19: limitations, constraints and pitfalls", AI & Soc (2020), published.*

[16] *C. J. Wang, C. Y. Ng, R. H. Brook, "Response to COVID-19 in TaiwanBig Data Analytics, New Technology, and Proactive Testing",*JAMA. *2020;*

[17] *A. Al-Badia, A. Tarhinia, A. Islam Khan, Exploring Big Data Governance Frameworks, Procedia Computer Science 141 (2018), Pages 271–277*

[18] *Alhassan, I., Sammon, D. and Daly, M., Data governance activities: an analysis of the literature, Journal of Decision Systems, 2016*

[19] *J. Hagmann, Information governance–beyond the buzz, Records Management Journal, vol. 23 (3) 2013, pp. 228-240.*

[20] *V. Morabito, Big Data Governance. Big Data and Analytics 2015,Pages 83–104.*

[21] J. Wang, M. Li, Y. He, H. Li, K Xiao, & C. Wang, *A Blockchain Based Privacy-Preserving Incentive Mechanism in Crowdsensing Applications. IEEE Access* ,2018,*, 6, 17545–17556.*

[22] *L. K. Cheong, V. Chang, The Need for Data Governance: A Case Study, 18th Australasian Conference on Information System The Need for Data Governance 5-7 Dec 2007*

[23] *W. Kristin, "A Model for Data Governance – Organising Accountabilities for Data Quality Management" (2007). ACIS 2007 Proceedings. 80.*

[24] *D. Rao, V. N. Gudivada, & V. V. .Raghavan. Data quality issues in big data. IEEE International Conference on Big Data 2015*

[25] *A. Noraini, I. S. Adli, S. Siti, and M. S.Suriani , Data Quality in Big Data: Areview, Int. J. Advance Soft Compu. Appl, Vol. 7, No. 3, November 2015*

[26] *W. Dai, I. Wardlaw, Y.Cui, K. Mehdi, Y. Li, and J. Long, Data profiling technology of data governance regarding big data: Review and rethinking, In Information Technology: New Generations Springer,2016, pp. 439-450)*

[27] *K. O'Neal, Big Data: Governance is the Critical Starting Point, 2012.*

[28] *P. Brous, M. Janssen, and R. Vilminko-Heikkinen, Coordinating Decision-Making in Data Management Activities: A Systematic Review of Data Governance Principles. Electronic Government, 2016, pp 115–125*

[29] *T. P. Mashamba-Thompson and E. D. Crayton, Blockchain and Artificial Intelligence Technology for Novel Coronavirus Disease 2019 Self-Testing,* Diagnostics *2020,* 10*(4), page 198*

[30] *Alam, Tanweer, Internet of Things and Blockchain-Based Framework for Coronavirus (COVID-19) Disease (July 25, 2020).*

[31] *C. Gros,R. Valenti,L. Schneider,K. Valenti,and D.Gros, Containment efficiency and control strategies for the Corona pandemic costs,2014*

[32] *H. Y. Kim, and J. S. Cho, Data governance framework for big data implementation with NPS Case Analysis in Korea, Journal of Business and Retail Management Research, 2018, vol. 12 (3), pp. 36-46.*

[33] *Datameer, Datameer Big Data Governance: Bringing open-architected and forward-compatible governance controls to Hadoop analytics,2016*

[34] *J. Caserta, Big Data Warehousing Meetup: Intro to NoSQL Databases Accessed ,2016*

[35] *S. Soares, Big Data Governance: A Framework to Assess Maturity, 2012*