

The 12th International Conference on Emerging Ubiquitous Systems and Pervasive Networks
(EUSPN 2021)
November 1-4, 2021, Leuven, Belgium

Emerging Concepts Using Blockchain and Big Data

Fadi Muheidat^{a*}, Dhaval Patel^b, Spandana Tammisetty^b, Lo'ai A. Tawalbeh^c,
Mais Tawalbeh^d

^aCalifornia State University- San Bernardino, CA 92407., USA, ^bUniversity of The Cumberland, Williamsburg
KY 40769, USA, ^cTexas A&M University-San Antonio, TX, 78224, USA

^dJordan University of Science and Technology, Computer Engineering Department, Irbid, 22110, Jordan

Abstract

Blockchain and big data are the emerging technologies that are on the highest agendas of the firms. These are significantly expected to transform the ways in which the business as well as the firm run. These are on the verge of increasing the expectation of the distributed ledgers, which would keep the firms away from struggling challenges. The concept of big data and Blockchain has been used up by various other concepts that would help secure and interpret the information. The ideal solutions offered by these technologies shall address the challenges of big data management as well as for analytics. In addition to that, Blockchain provides its own consensus method, which is the primary means to create an audit trail. This enables users to verify all transactions. The audit trail is a means of verifying the correctness and integrity of every transaction, regardless of who owns the asset. The Blockchain can also verify that different parties of a transaction are following an agreement and not breaking the agreement. Moreover, there have been continuous arguments in the concept of Blockchain at which the bitcoin is fundamental and there are several popular blockchain approaches developed which would deliver performance, security as well as privacy. Apart from this, the use of Blockchain plays a major role in adding an extra data layer for the big data analytics process. Big data is considered secure which cannot be further forged with the network architecture. The current paper shall be discussing the emerging concepts that are using Blockchain as well as big data.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the Conference Program Chairs

Keywords: Blockchain; Bigdata; Data analytics; Smart Cities applications; Healthcare;

* Corresponding author. Tel.: +1-210-784-9632,

E-mail address: fadi.muheidat@csusb.edu

1. Introduction

Blockchain is a distributed ledger system based on a decentralized environment that protects large amounts of data generated in the organization. Thus, act as a global record-keeping database and a central storage and verification service without the limitations of storing data offline [1]. By storing large amounts of data in a distributed ledger, large corporations can ensure transparency of transactions, manage all financial accounts and other transactions, avoid frauds, manage, and resolve disputes over assets, and prevent other criminals from gaining access to data [2]. This system of global record keeping, centralization, and financial security can be used to allow businesses to acquire, distribute quickly, and exchange services, such as access to digital assets, as well as to create and manage global social networks, among other things [3].

Blockchain can also verify that different parties of a transaction are following an agreement and not breaking the agreement. When two parties have a valid transaction that they signed with a trusted third party, they can now create that transaction into the Blockchain. The second party can see it on the ledger and can verify whether it is valid, but it is the third party who now has the power to go back and edit a transaction. That is, the third party can go in at the beginning of the block and undo a modification or it can reverse the change in the Blockchain that the first party attempted to initiate [4]. The third-party can also go in at the beginning of the block. It can simply take that block and add it to the existing Blockchain, it can make a duplicate, or whatever it is that the third party has to do, put it somewhere it can see and see that other side for which it has access, and so forth.

However, big data is the Blockchain's use to create a storage network, databases that can combine extensive data to create new data collections. This could make new kinds of business models, which are very relevant to big data segments, such as digital currencies and digital identity applications. Both technologies provide a way to make various extensive data collections without the limitations of being stored offline or using a data centre. Big data uses data to be stored and processed as data and makes data collections more stable and reliable. However, some businesses need to consider data security to adopt Blockchain to protect their assets and privacy needs. To protect their assets, digital services in the Blockchain have a built-in mechanism to ensure that they don't lose access to the data they make their money on, such as Blockchain [5].

Most Big Data applications use Blockchain to ensure data quality, accessibility, and security, thus giving businesses the advantage of investing funds in new kinds of business models. One of the benefits of Blockchain is that it allows users to create smart contracts on their computers that give them ownership rights to the data. This makes Blockchain a safe space for these kinds of businesses, allowing them to create new kinds of products and services, services of various financial and social uses. While using Blockchain is not bad for companies to develop new types of business models, this still means that this space is currently dominated by a few centralized companies and only a few distributed ones [6].

2. Blockchain and Big Data in the Smart City

It has been found that with rapid urbanization, there have been developments of smart cities that need efficient solutions for the administration and energy optimization. The integration of big data and internet technologies with the capability of delivering infrastructure solutions needed for smart cities. Blockchain technology caters to the needs of having space, energy efficiency, and maintenance of IoT devices. With this, the third-party auditors come into existence and refer to the centralized frameworks that are subjected to the issues in the cloud environment [7]. With this framework, the nature audit is handled and lets the owners in auditing. Apart from this, it allows batch verifications of auditing proofs which ensures the security of the privacy. Herein the blockchain-based infrastructure offers the Spatio-temporal smart services that offer sustainable IoT-based smart megacities.

The generation of big data collects and analyses the prediction of events. This framework consists of device-to-device communication that allows Blockchain to perform offline operations. Thus, it can be stated that blockchains shall encourage rapid growth in smart cities with their flexible integration. The functionality of blockchain-based on big data shall be delivering an approach for securing the transactions data with IoT nodes. The concept of a smart city illustrates how to deliver security to the citizens [8]. In this regard, the integration of technologies shall help in

identifying the prone areas to crime hubs and thereby predict the crime locations. Thus, the historical, as well as geographical data provided by the predictive analytics feature of big data, shall help the city in creating a safer environment. Overall, the use of the technology shall provide an intelligent way of generating information for the improvement of smart city facilities (see Fig. 1).

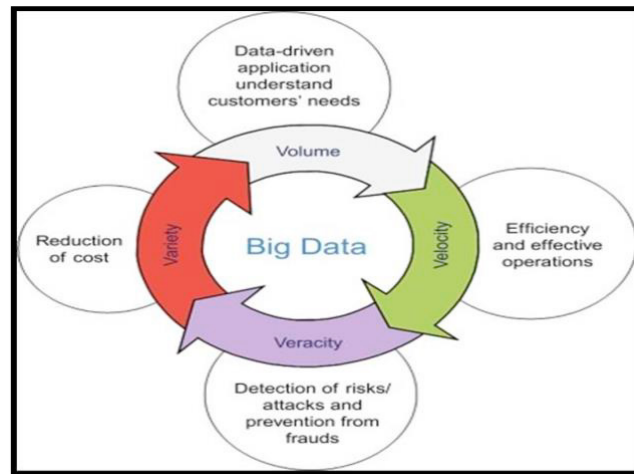


Fig. 1. Big data in cybersecurity across smart city applications.

3. Blockchain and Big Data in Healthcare

The Healthcare industry is considered as intricate and complex as it has undergone radical innovations in the areas such as clinical trials as well as cloud computing adoption. Blockchain technology comes with the potential of eradicating the issues and reforming healthcare by keeping the patient at the crux of the medical ecosystem. The key value of blockchain technology across healthcare is estimated to be \$2.12 billion in the year 2019 and it is currently expected to reach \$3.49 by the year 2025 [9]. The applications in healthcare are experimented by the firms and with effective tracking of payment options enhancements and decentralization of patient data, Blockchain has significantly affirmed as an irreplaceable tool.

The technology is known for its immutability as well as security. The transaction shall be recorded in the form of blocks which would not allow it to revamp and alter. The technology develops the cryptographic techniques which offer the cybercity structure in terms of data sharing. This allows seamless patient data management and while recording the data cloud computing shall store data. The patient shall easily manage the data access with the mode of smart contracts. Blockchain-based firms such as Viant offer the solution for the supply chain of pharma giants to eradicate the issue. The application is typically based on technology tracing which would tune with the medical industry supply chain [10]. Thus, with effective technology integration, collaboration and monitoring can be enhanced with saving of timing and effort at the core. Big data analytics comes with greater capabilities for analyzing the patient treatment, response behaviors as well as medicine usages. These results help in providing treatment at a low cost as well as develop the tools for the people's benefit (see Fig. 2).

4. Blockchain and Big Data Application in Education

The integration of Blockchain and big data in the educational sector shall improve the performance of the institutions and student learning. The technology has the potential to empower learners, enhance the security as well as Efficiency of the educational institutions. The suppliers of Blockchain across the education solutions which includes digital credential partnerships shall build the educational institutions along with the defined stars. However, it can be stated that the blockchain-in-solution integrations for generating significant positive impact shall require unity across the private and public sectors to coordinate the best efforts [11]. Additionally, Blockchain should not be

perceived as a mere threat to educational institutions and must be embraced as an innovative technology that would add value to the educational processes, make the learning more engaging and improve trust and enhance privacy. The concept of big data integration with the editorial system aims at improving the results. It helps in monitoring the actions of students such as the way of answering and skipping the questions [12].

Customized programs are also available to the students which are known as the blended version of learning. This makes the class interesting and reduces the chances of dropouts for college students. Big data is used for analyzing and monitoring the way in which the graduates are performing across the job market. With big data, institutions shall be more likely to predict the applicants and thereby analyze the factor affecting the application process [13]. The knowledge provides the institutions with the scope of adjusting the recruitment strategies and accordingly allocate the funds. The influx of data analyses the information about the schools and speeds up the application process for the context of international students [14]. Big data thus comes with the potential of revolutions the learning sector and smarter students shall create a positive impact on society. Thus, the application of big data and Blockchain can be considered huge for the learning sector.

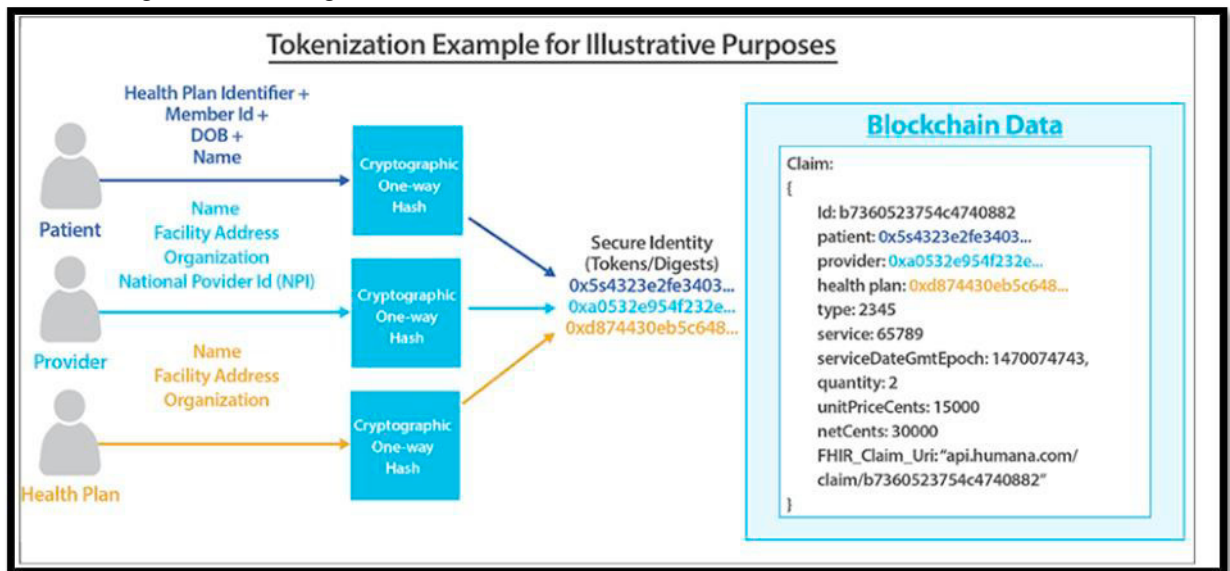


Fig. 2. Blockchain in healthcare.

5. Blockchain and Big Data in the Finance Industry

Blockchain Technology focuses on automating the processes which are required for meeting the needs for proper regulation prior to execution. There has been a growing demand for compliance within the financial industry which emphasizes to shift for securing the ongoing technologies. All these can be used for enhanced storage, tracking, and retrieving of data. Volume, velocity, and variety of data which most financial institutes depend on are found to be much more overwhelming. All the game-changing technologies which have come up with data can provide some indisputable value. The use of some of the leading technologies like Big Data and Blockchain in the financial sector can help the banking industry to go just beyond the mere option of cashless payment. Machine learning systems can help in automatic learning and improvement for identification of possible risks which help in better understanding and avoiding [11]. Most financial companies emphasize big data as it will again help understand processes like trading, fraud, and risk. Fintech companies can help in trading certain kinds of algorithms, which can again outperform the competitors. Supervised machine learning can give a higher accuracy in comparison to any other mathematical model. This is because trading decisions depending on stocks and models. Blockchain technology can help automate certain kinds of processes that can again help meet proper regulations prior to execution. There has been increasing demand for compliance within the financial industry that can help store, track, and retrieve the data [15]. Blockchain does not come up with any central point of vulnerability that needs to be exploited to make the present technology

safe and transparent. Both big data and machine learning can help the business to have an idea regarding the customers on the given level prior to analytics and forecasting.

Data-driven insights can also help in carrying out data analysis and minimize possible human errors.

The real-time application of using big data and Blockchain in a banking industry that would significantly reduce the cost of service while also decreasing risk. Money lending platforms mostly use this approach to ensure the data generated by users is secure, transparent, and reliable. When there is no risk involved, it can determine transactions that may or may not be fair. Many of these platforms are developed by large banking companies that can also validate and analyze payments. The data generated by these platforms are made available for a broad range of use, from collecting interest and credit history data and asset classification to the analysis of financial structures to credit quality data [1].

The application of big data analytics in peer-to-peer platforms could be applied to store and transfer big data using computational storage where Blockchain detects fraud and security flaws, which would reduce the risk associated with such fraud. Besides, these platforms can also be used to determine borrowers' security risks through credit histories, loan agreements, and transaction history, among others [16]. When users transfer funds to each other on P2P platforms, the payment protocol is a new business model with real-time payment processing based on Blockchain. Blockchain would facilitate payment processing without intermediaries or intermediaries having their banking accounts or credit card accounts. It could benefit the ecosystem because of reduced latency and increased reliability [16].

The data generated by online transactions and user accounts on the P2P platform is stored in a big file on the Blockchain, storing data as large as the entire bank's infrastructure. This data could be analyzed by applying computer vision algorithms, data mining, machine learning, and automated classification of documents using classification algorithms. Once processed, the dataset would be made available on an open, transparent, data-sharing platform where every user can see real-time data, that data is of interest to them, and what they can do with it. The information is also available in the data mining environment, where the Blockchain would serve as a database from which other platforms can discover the data in any order [16].

The open nature of the blockchain system in the payment network environment can also be exploited to implement privacy features. For example, suppose a user provides data to a payment network and cannot provide full details about their payment. In that case, the payment network will keep the privacy state variable between their users. To do this, the payment network would implement an encrypted message which the user can send to another payment network or their network as well, in which the message is encrypted, and the encryption key is protected by a particular layer of security measures that prevent anyone from looking at the news [16].

6. Blockchain and Big Data in the Retail Industry

Blockchain technology can be used to reduce the possible scenarios created by using the database. It will help in enhancing food safety, increase profit and minimize overall waste. Blockchain technology and big data can help in fraud prevention, manage loyalty points, complain about consumer data, accept payment with cryptocurrency, and lastly, track down inventory. Blockchain technology was created to prevent any kind of fraudulent activity for digital currency exchange [17]. Some of the leaders in the retail industry are making use of Blockchain technology to prevent certain items from reaching the desired customers. Blockchain technology aims to provide allocation and redemption within real-time to different consumer and loyalty programmers. Blockchain can help carry out the management of data privacy for different customers by using smart contracts [18]. This technology can be used for cryptocurrencies which can be again used for value or data. This mainly comes into the picture for retailers who can handle cross-border payments and micro-payments. Blockchain technology can give complete visibility to each and every point to the supply chain and retailers. This will help them in tracking shipments for performing stocks with accuracy.

7. Decentralized Storage System in Video Streaming Applications like Netflix

The online video streaming applications mostly rely on the decentralized nature of Blockchain and big data to process and store the video. The application itself is based on an online storage system and is powered by Blockchain. The network uses an open distributed architecture based on peer-to-peer networks and allows users to transfer file contents in real-time. This can enable a system to store files on a system accessible by a user at a fixed time; that is, the file contents may only be changed on a local machine at a specified time. Video data is processed using these local machines using a peer-to-peer network based on the Ethereum blockchain network. As a result, data storage does not depend on any central authority. The application can also display data in a local file store on the system accessible by a user at a fixed time [16].

In the video software platform, the user interfaces are powered by decentralized blockchains and can be distributed across a network. For example, the user-interface can be implemented using smart contracts on the Ethereum blockchain network. The applications could contain applications developed on the Ethereum blockchain network. The file-store application also depends on smart contracts and distributed application-layer services based on Blockchain. Therefore, the file-store application can be integrated with smart contracts and combined with an online video streaming platform to allow the user to display a video in live streaming.

When content creators stream their content on Netflix using the internet, they need to secure the data stored in Netflix's large file. But, since their files contain sensitive content, they cannot trust their files not to be stored in the cloud. To protect their files from being stolen, they need to encrypt them to prevent other users from viewing their files. This is achieved by using a blockchain that is encrypted using a public key and stored in a key-file distributed on the Blockchain. To secure the tickets used to encrypt the Blockchain, an algorithm is employed independently of the original data's encryption algorithm. The data stored in the Blockchain is then encrypted by a key that is stored on the Blockchain using the block ledger that encodes the file to the file [18].

The volume of data generated by Netflix increases every that needs big data storages, like videos and TV shows. When users stream their content, they have to trust to their files not to be stolen and not to be accessed by others, especially by criminals, but it cannot be prevented. To provide this trust, the size of Netflix's huge file is becoming larger and larger, so what if the extensive file contents are stored in local cloud storage. They can send any content to this cloud storage securely and retrieve their content without anyone knowing it. Then, they can stream their content to their own remote devices by linking to their data without anyone knowing about it [18].

8. Big Data Analytics for Logistics and Distributions using Blockchain-Predictive Analytics

Big Data analytics for logistics and distribution are proliferating because the demand to increase the storage and analysis capacity for this business category has grown manifold. The main challenges that the market faces are developing storage capacity, which comes at a price in terms of latency, and thus, more data is needed. Data analysis using Blockchain brings together all these aspects of data storage to be used in a safe, efficient, and secure manner [19]. To improve the situation, the companies are focusing on developing Blockchain technology, and Blockchain technologies are becoming popular among them because the cost of data analytics for logistics and distribution has come down significantly [20].

A company that has managed an application of Big Data analytics can use Blockchain to create transactions, access, and monitor every data produced by the application. All inputs need to be stored, verified, or verified and stored. A blockchain can validate all input parameters to be recorded in Blockchain that was previously unverified. If the input parameters for the application have been validated, the user can create a transaction that the company creates based on the inputs. Companies will need both the hardware and power to complete these transactions and verify transactions to make the Blockchain records, where all information is verified and stored in the Blockchain [20].

In most databases, there is a limited amount of space available for analyzing the data. In a Blockchain implementation, the processing will occur in a distributed manner and is not limited by the amount of space available to analyze the data. An example of an application for data analysis is the prediction of future economic trends based

on historical data. While prediction can be used for different purposes, a specific goal of prophecy would be to make investments and predict the futures of financial conditions by predicting economic developments, and thus forecasting the outcome of economic conditions to a more exact accuracy.

Big data in analytics is often described as applying algorithms that solve real-time problems of predicting future events. Using distributed computing systems can achieve such an application by using distributed machines. There are no resources available in a distributed device to operate the machines. There is no server location, a traditional distributed engine that means no physical area for computing machines. Using distributed machines is also referred to as multi-agent. A multi-agent device can perform calculations and analyze complex data. The measures can occur in multiple machines simultaneously, and this information can be presented to all the machines at once [20].

Big data and Blockchain in predictive analytics provide accurate predictions and information about future development. A centralized person who is not connected to any device can operate the machines independently and without using any physical space or server with a distributed machine. This is like how a data processor could operate a computer. Data to be processed, stored, analyzed, and applied on a centralized device are in the datastore. As it is centralized, someone can't move to another datastore location and re-analyze, store all such data on the database or even do such calculations in different areas [20].

9. Conclusion

The report has explored the emerging concepts that reusing the technologies to develop the creditworthy systems across the participants in the non-trustable environment. The secure, as well as autonomous-based technologies are based on self-regulated data as well as programmable smart contracts. The paper has provided insights on blockchain technologies and big data that depict the innovation on the applications. The concept of big data has further generated significant interest in the scientific domain over the past few years. With the decentralization feature of Blockchain, there has been significant potential to grow and improve the services. It has been found that the integration of the block and big data shall provide real-time data analytics to enable the financial institutes to settle the transactions in a quick fashion. From the extensive discussion, it can be stated that there is a need for devoted credit approaches that would further make the system creditworthy for the data-level storage operations.

References

- [1] M. K. F. Aldaboubi, "Big Data In Single Player Games," PhD Thesis, 2020.
- [2] K. Abbas et al., "Convergence of Blockchain and IoT for Secure Transportation Systems in Smart Cities," *Security and Communication Networks*, vol. 2021, 2021.
- [3] M. Aldwairi and L. Tawalbeh, "Security techniques for intelligent spam sensing and anomaly detection in online social platforms," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 1, p. 275, 2020.
- [4] S. M. Tadaka, "Applications of Blockchain in Healthcare, Industry 4, and Cyber-Physical Systems," in *2020 7th International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*, 2020, pp. 1–8.
- [5] G. Saldamli and A. Razavi, "Surveillance Missions Deployment on the Edge by Combining Swarm Robotics and Blockchain," in *2020 Fourth International Conference on Multimedia Computing, Networking and Applications (MCNA)*, 2020, pp. 106–112.
- [6] S. K. Sharma, B. Bhushan, A. Khamparia, P. N. Astya, and N. C. Debnath, *Blockchain Technology for Data Privacy Management*. CRC Press, 2021.
- [7] M. Sun and J. Zhang, "Research on the application of block chain big data platform in the construction of new smart city for low carbon emission and green environment," *Computer Communications*, vol. 149, pp. 332–342, 2020.
- [8] H. Yu, Z. Yang, and R. O. Sinnott, "Decentralized big data auditing for smart city environments leveraging blockchain technology," *IEEE Access*, vol. 7, pp. 6288–6296, 2018.
- [9] K. Rabah, "Convergence of AI, IoT, big data and blockchain: a review," *The lake institute Journal*, vol. 1, no. 1, pp. 1–18, 2018.
- [10] A. A. Siyal, A. Z. Junejo, M. Zawish, K. Ahmed, A. Khalil, and G. Sourso, "Applications of blockchain technology in medicine and healthcare: Challenges and future perspectives," *Cryptography*, vol. 3, no. 1, p. 3, 2019.
- [11] Q. Liu and X. Zou, "Research on trust mechanism of cooperation innovation with big data processing based on blockchain," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, pp. 1–11, 2019.
- [12] A. Alammary, S. Alhazmi, M. Almasri, and S. Gillani, "Blockchain-based applications in education: A systematic review," *Applied Sciences*, vol. 9, no. 12, p. 2400, 2019.
- [13] G. Chen, B. Xu, M. Lu, and N.-S. Chen, "Exploring blockchain technology and its potential applications for education," *Smart Learning Environments*, vol. 5, no. 1, pp. 1–10, 2018.

- [14] A. Mikroyannidis, J. Domingue, M. Bachler, and K. Quick, “Smart blockchain badges for data science education,” in *2018 IEEE Frontiers in Education Conference (FIE)*, 2018, pp. 1–5.
- [15] H. Hassani, X. Huang, and E. Silva, “Banking with blockchain-ed big data,” *Journal of Management Analytics*, vol. 5, no. 4, pp. 256–275, 2018.
- [16] N. Deepa et al., “A survey on blockchain for big data: Approaches, opportunities, and future directions,” *arXiv preprint arXiv:2009.00858*, 2020.
- [17] M. Li, S.-P. Guan, and R.-R. Du, “Research on the application of customs supervision on cross-border import e-commerce retail business from the perspective of blockchain,” in *6th International Conference on Management Science and Management Innovation (MSMI 2019)*, 2019, pp. 74–78.
- [18] Y. Zhang, F. Xiong, Y. Xie, X. Fan, and H. Gu, “The Impact of Artificial Intelligence and Blockchain on the Accounting Profession,” *IEEE Access*, vol. 8, pp. 110461–110477, 2020.
- [19] D. M. F. Mattos, F. Krief, and S. J. Rueda, *Blockchain and artificial intelligence for network security*. Springer, 2020.
- [20] J. W. Roepert, “Digital Supply Chain—Die Digitalisierung der Supply Chain mit Hilfe von IoT, Machine Learning, Blockchain, Predictive Analytics und Big Data,” in *Logistik—die unterschätzte Zukunftsindustrie*, Springer, 2020, pp. 83–98.