information

Article

Machine Learning and Blockchain: A Bibliometric Study on Security and Privacy

Alejandro Valencia-Arias [1,*], Juan David González-Ruiz [2], Lilian Verde Flores [1], Luis Vega-Mori [3], Paula Rodríguez-Correa [4] and Gustavo Sánchez Santos [3]

[1] Escuela de Ingeniería Industrial, Universidad Señor de Sipán, Chiclayo 14001, Peru; lilianverde@crece.uss.edu.pe
[2] Departamento de Economía, Universidad Nacional de Colombia, Medellín 050001, Colombia; jdgonza3@unal.edu.co
[3] Instituto de Investigación y Estudios de la Mujer, Universidad Ricardo Palma, Lima 15074, Peru; iepvlvega@gmail.com (L.V.-M.); sanchezsantosgustavo1@gmail.com (G.S.S.)
[4] Centro de Investigaciones Escolme—CIES, Institución Universitaria Escolme, Medellín 050001, Colombia; cies4@escolme.edu.co
* Correspondence: valenciajho@uss.edu.pe; Tel.: +57-3002567977

Abstract: Machine learning and blockchain technology are fast-developing fields with implications for multiple sectors. Both have attracted a lot of interest and show promise in security, IoT, 5G/6G networks, artificial intelligence, and more. However, challenges remain in the scientific literature, so the aim is to investigate research trends around the use of machine learning in blockchain. A bibliometric analysis is proposed based on the PRISMA-2020 parameters in the Scopus and Web of Science databases. An objective analysis of the most productive and highly cited authors, journals, and countries is conducted. Additionally, a thorough analysis of keyword validity and importance is performed, along with a review of the most significant topics by year of publication. Co-occurrence networks are generated to identify the most crucial research clusters in the field. Finally, a research agenda is proposed to highlight future topics with great potential. This study reveals a growing interest in machine learning and blockchain. Topics are evolving towards IoT and smart contracts. Emerging keywords include cloud computing, intrusion detection, and distributed learning. The United States, Australia, and India are leading the research. The research proposes an agenda to explore new applications and foster collaboration between researchers and countries in this interdisciplinary field.

Keywords: Internet of Things; 5G networks; artificial intelligence; PRISMA-2020; cloud computing; intrusion detection; smart contracts

Citation: Valencia-Arias, A.; González-Ruiz, J.D.; Verde Flores, L.; Vega-Mori, L.; Rodríguez-Correa, P.; Sánchez Santos, G. Machine Learning and Blockchain: A Bibliometric Study on Security and Privacy. *Information* **2024**, *15*, 65. https://doi.org/10.3390/info15010065

Academic Editors: Georgios Siolas, Georgios Alexandridis and Paraskevi Tzouveli

Received: 13 August 2023
Revised: 23 September 2023
Accepted: 25 September 2023
Published: 22 January 2024

1. Introduction

Two rapidly evolving research areas that significantly influence various fields are machine learning and blockchain technology. Machine learning, also called computer science, concentrates on creating algorithms and models that recognize patterns and draw conclusions from data without human interference. The immutable storage and recording of transactions are made possible by blockchain technology, a secure, decentralized data structure. Their potential to revolutionize industries and enhance various applications has generated significant interest [1,2].

Research on the future challenges and evolution of communication networks is ongoing. Salahdine and colleagues (2018) conducted research on advancements and complications associated with 5G, 6G, and beyond. Their study suggested the adoption of blockchain technology to establish a secure and decentralized computing model. This investigation focuses on the effectiveness of a blockchain-based platform for ensuring decentralized and secure computing operations.

To enhance the security and accessibility of sensor control systems, blockchain and machine learning were explored in the context of smart manufacturing. An architecture utilizing these technologies was proposed to ensure the security of sensor control systems in the industrial Internet of Things (IoT) [3]. Thus, blockchain technology serves as a valuable tool for enhancing security in advanced industrial environments.

Concerns over data security and privacy have grown as the Internet of Things (IoT) becomes increasingly prevalent. A study investigating the security threats, solution architectures, and application domains of IoT [4] emphasized the importance of addressing security issues within the IoT. Regarding the management and security of these networks, the introduction of 5G and beyond has presented opportunities for utilizing blockchain technologies. A thorough analysis of blockchain's application in 5G and beyond networks has been conducted, which divulges the current state of research in this area [5]. This illustrates the significance of integrating the potential of both technologies to tackle core issues in the arena of next-generation communication networks.

Due to its capacity to tackle issues and enhance applications in an ever-more digitized and interconnected world, this realm of research holds substantial weight in numerous fields. The convergence of the two technologies has demonstrated substantial potential in areas such as the Internet of Things (IoT). The significant research on the enormous IoT progression towards 6G underscores the value of machine learning in facilitating connectivity and informed decision making [6].

Security is critical in the thriving field of in-vehicle networks, which have undergone substantial growth alongside IoT. Researchers have investigated using blockchain and machine learning (ML) to enhance the security, reliability, and privacy of vehicle communications and have determined its potential advantages [2].

Blockchain technology consists of a network of cryptographically secured blocks that contain information concerning digital transactions. Owing to its versatility, it has been put to use in many sectors and has showcased its adaptability within Industry 4.0. Most recently, the technology has been harnessed as an intelligent traceability system in the wine industry to prevent counterfeiting in its supply chain [7]. It has similarly been employed in the food industry to ensure safety, quality, and regulatory compliance through a traceability approach [8]. In the manufacturing industry, companies are utilizing blockchain technology for collaborative innovation. Research by Wan et al. [9] suggests that blockchain applications enhance collaborative innovation and reinforce the favorable influence of social trust. Therefore, companies and governments must establish an ecosystem that facilitates blockchain technology adoption.

The technology has assisted in managing marine plastic waste, increasing public awareness of waste management by replacing paper documents and cash transactions with digital wallets and distributed ledgers, thereby reducing its impact on local economies. This technology has also improved efficiency and safety [10], making it suitable for sustainability management. This approach has been shown to be useful in marketing management, particularly in the supply chain and internal management of marketing operations. It can help professionals systematize their internal management and operations, including marketing campaigns [11].

Due to its ability to ensure data integrity and improve decentralization in the model training process, blockchain has attracted significant attention in the field of deep learning. Our review of blockchain's use in deep learning identifies open questions and areas for future research aimed at enhancing the efficiency and security of these systems [12]. Additionally, the interaction between blockchain and machine learning is explored in the context of artificial intelligence. A review and analysis of the advancements in the field were carried out, identifying research priorities and obstacles for the effective integration of blockchain technology in applications involving artificial intelligence [13].

Machine learning and blockchain fields have undergone significant growth owing to their potential to cater to a wide range of applications. Despite making significant progress, there continue to be research gaps in need of investigation through the use of bibliometrics

to assess the current state of scientific literature in these fields. The interpretability of models is an area requiring attention in machine learning research. As machine learning algorithms become increasingly complex and are utilized in mission-critical applications, comprehending how and why certain decisions are made is indispensable. This challenge is also pointed out in the examination of blockchain implementations in artificial intelligence, which accentuates the requirement to tackle the interpretability and transparency of blockchain-driven machine learning models [13].

Another research void linked to blockchain technology pertains to the scalability and effectiveness of blockchain networks in high-performance scenarios. As blockchain adoption expands into areas such as the IoT and vehicle networks, there is a growing need for solutions that can securely and efficiently handle large volumes of transactions. In fact, a review of quantum walks also highlighted challenges with efficiency and scalability when dealing with large datasets [14]. Therefore, the goal is to examine the research trends around the use of machine learning in blockchain, in order to structure a research agenda that allows for a focused implementation of future research:

RQ1: What are the years in which the use of machine learning in blockchain has shown the most interest?

RQ2: How is the number of scientific articles on the use of machine learning in blockchain growing?

RQ3: What are the main research references on the use of machine learning in blockchain?

RQ4: What is the thematic evolution derived from the scientific production on the use of machine learning in blockchain?

RQ5: What are the main thematic clusters on the use of machine learning in blockchain?

RQ6: What are the growing and emerging keywords in the research area of using machine learning on blockchain?

RQ7: What topics are positioned as protagonists for the design of a research agenda on the use of machine learning in blockchain?

2. Materials and Methods

To accomplish the objectives of this study, we have utilized an exploratory methodology limited to reviewing secondary sources of information, namely a bibliometric analysis that assesses the scientific literature pertaining to data security and privacy. The evaluation will follow parameters outlined in the international declaration PRISMA—2020 [15].

2.1. Eligibility Criteria

To conduct bibliometrics on machine learning and blockchain, we will initially include documents with titles and keywords that have basic metadata pertaining to these concepts. This metadata will serve as the foundation for identifying pertinent publications through the literature search. Second, documents that explicitly address the combination of machine learning and blockchain, as well as those that address these topics differently, will be reviewed to ensure a comprehensive coverage of the relevant literature.

The exclusion process includes three stages to guarantee appropriate selection of documents that fulfill the bibliometrics objectives. In the initial phase, we exclude documents that have erroneous or inadequate indexing to ensure the accuracy and quality of the collected data. The subsequent stage involves the exclusion of documents, particularly Systematic Literature Reviews, that lack access to the full text. This is because conducting a thorough review of the content is critical for bibliometric analysis. It is worth noting that this stage is solely applicable to systematic literature reviews, as bibliometric analysis primarily relies on document metadata.

Ultimately, in the third stage of elimination, conference proceedings and documents without complete indexing or inadequate information will be omitted. This exclusion stage ensures that only documents that are appropriately categorized and classified in databases and bibliographic sources are included in the analysis. These conference proceedings

were excluded as they typically remain in prototype phase and lack sufficient validation evidence, according to Dyzel et al. [16]. The aim of this study is to analyze the operation of blockchain and machine learning technologies in an operational environment to enhance understanding of current knowledge and identify future applications [17].

By adhering to the inclusion criteria and exclusion stages outlined in the PRISMA 2020 framework, we will ensure objectivity and consistency in selecting relevant studies for a comprehensive and dependable bibliometric analysis of machine learning and blockchain.

2.2. Source of Information

In this bibliometric analysis, the Scopus and Web of Science databases were chosen for their significance as the primary bibliographic sources in contemporary literary research. They are widely acknowledged to provide extensive coverage of scientific and scholarly publications, thereby guaranteeing the inclusion of an extensive range of relevant studies in the area of interest. Additionally, both platforms offer advanced instruments and functionalities for bibliometric analysis, including citation networking and normalized impact factor calculation.

This study emphasizes that Web of Science, Scopus, and Dimensions serve as distinct echo chambers with differing indexing policies, coverage, and scope when collecting bibliometric data [18]. Nonetheless, despite their discrepancies, these databases are deemed by the scientific community as the most utilized and trusted due to their emphasis on content quality and capacity to present a comprehensive and precise perspective of worldwide scientific production. Using Scopus and Web of Science as primary sources for bibliometric ensures obtaining a representative collection of publications and data, facilitating a substantial and meaningful analysis of machine learning and blockchain research.

2.3. Search Strategy

To achieve an efficient search in Scopus and Web of Science, we formulated two specialized search equations that closely aligned with the established inclusion criteria and were tailored to the distinct search attributes of each platform. These equations were meticulously crafted to comprehensively and relevantly encompass all publications pertaining to the topic of interest, namely, machine learning and blockchain. Terms, synonyms, and variations in concepts within the targeted research areas were carefully selected to ensure comprehensive and accurate coverage of the scientific literature. The implementation of specialized search algorithms, tailored to the specific parameters of Scopus and Web of Science, delivers an efficient and thorough compilation of pertinent studies, resulting in a meticulous and dependable bibliometric analysis of the subject.

For the Scopus database: ((TITLE (“machine learning”) AND TITLE (“blockchain”) AND TITLE (“security” OR “privacy” OR “cybersecurity” OR “data protection”)) OR (AUTHKEY (“machine learning”) AND AUTHKEY (“blockchain”) AND AUTHKEY (“security” OR “privacy” OR “cybersecurity” OR “privacy”))).

For the Web of Science database: ((TI = (“machine learning”) AND TI = (“block-chain”) AND TI = (“security” OR “privacy” OR “cybersecurity” OR “privacy”)) OR (AK = (“machine learning”) AND AK = (“blockchain”) AND AK = (“security” OR “privacy” OR “cybersecurity” OR “privacy”))).

2.4. Data Management

In this study, we utilized Microsoft Excel to extract and process data from Scopus and Web of Science. The merged database helped in evaluating the eligibility criteria by all authors. The VOSviewer® software version 1.6.19 was utilized to generate bibliometric graphs and analyze document-author relationships per the methodology outlined in [19]. This tool combination provided a comprehensive view of scientific production, aiding trend identification and discernment of field evolution.

2.5. Selection Process

As emphasized in the PRISMA 2020 statement and demonstrated in the study by [17], it is crucial to disclose the use of an internal automated classifier for study selection and the performance of internal or external validation to minimize missed studies or misclassification (refer to Figure 1). In this bibliometric investigation of machine learning and blockchain, the investigators utilized a jointly developed automated tool in Microsoft Excel as an internal aid. Each researcher independently applied the tool during the study inclusion and exclusion process to minimize the risk of omissions or errors in classifications, thereby ensuring the convergence and dependability of the outcomes. The implementation of the internal automation tool ensured improved accuracy and consistency in selecting studies, aiding the methodical and unbiased application of the inclusion criteria according to the PRISMA 2020 guidelines.

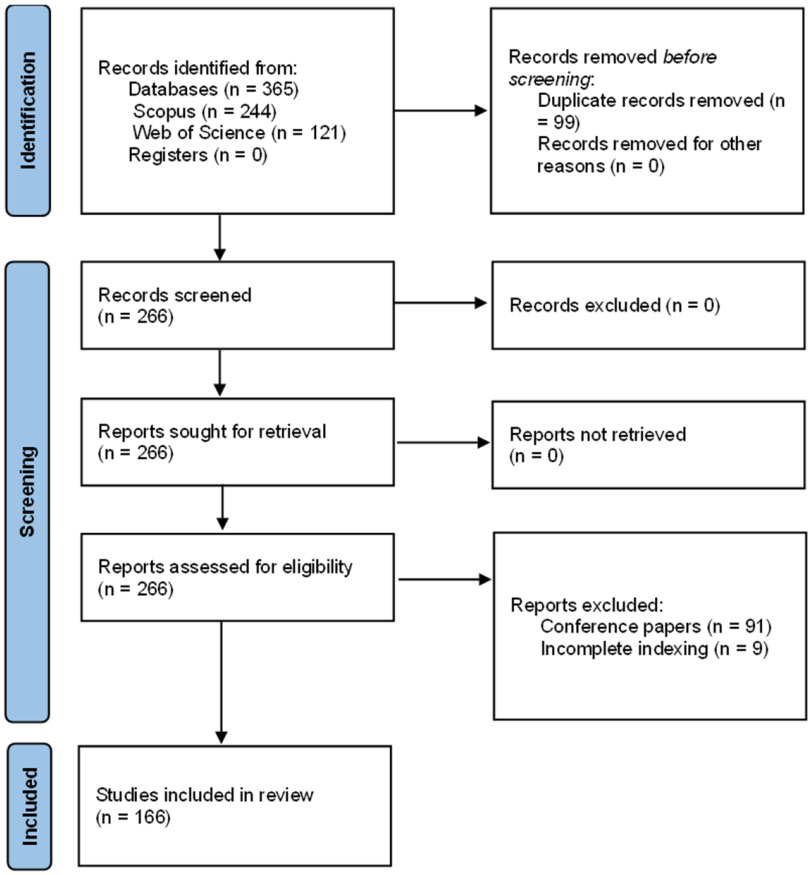

Figure 1. PRISMA flow chart. Own elaboration from Scopus and Web of Science.

2.6. Data Collection Process

For this bibliometric analysis of machine learning and blockchain, data were collected from selected reports using precise methods in accordance with guidelines established by [17]. Microsoft Excel was utilized as an automated tool for data collection from Scopus and Web of Science databases. All authors of this study acted as reviewers and independently verified the data for an objective and dependable evaluation of the publications. Furthermore, the authors performed a collective data confirmation process until complete convergence of results was obtained. The implementation of Microsoft Excel® (18.0) as an automated instrument, along with a collaborative strategy for data validation and confirmation, facilitated meticulous and precise gathering of data from the chosen bibliographic reports. This, in turn, bolstered the integrity and resilience of the bibliometric analysis.

2.7. Data Elements

We attempted to collect data for all articles that aligned with the research objective. For this purpose, we complied with the specialized search equations for each database, including those that mentioned both topics. However, if any information was missing or unclear in the articles, we excluded them as non-relevant to ensure consistency with the research's purpose and scope. Emphasis was placed on incorporating pertinent and lucid data and results to guarantee a thorough comprehension of the subject matter, ultimately facilitating the attainment of bibliometric objectives.

2.8. Assessment of Study Risk of Bias

This study followed a rigorous and systematic process to assess the risk of bias in the included studies. All authors utilized the same automated Microsoft Excel® tool for data collection, allowing for a collective and consensual evaluation of the risk of bias. This approach ensured the quality and integrity of the results by providing a consistent and objective assessment of the studies included in the bibliometric analysis. The use of the automated tool for assessing bias risk minimized potential biases and enhanced the validity and reliability of the bibliometric findings, bolstering the research's solidity and robustness.

2.9. Measures of Impact

This research utilizes a bibliometric approach for an objective and systematic analysis of scientific production on the topics of interest. Secondary sources provide information for studying scientific production through indicators such as publication and citation numbers, as well as the temporal evolution of keyword usage. Tools such as Microsoft Excel® and VOSviewer® are utilized for this analysis as well as to establish the thematic connections between documents. The approach is distinct from primary research as it refrains from utilizing effect measures, such as risk ratios or mean differences. Instead, the quantitative analysis and visual depiction of the thematic network is emphasized in order to identify patterns and trends using bibliometric indicators and visualization tools.

2.10. Synthesis Methods

Several techniques were utilized to identify qualified studies for each synthesis. This involved creating a table of intervention characteristics for each study and cross-referencing them against the planned groups for synthesis. Furthermore, techniques were employed to condition the data prior to presentation or synthesis, including managing absent summary statistics or data conversions. Specific approaches were adopted to generate tables and graphs displaying the results of individual studies and syntheses. Bibliometric measures of volume, quality, and composition were systematically computed via Microsoft Excel® for all documents that passed the three exclusion stages [20]. These measures enabled an impartial and methodical evaluation of relevant scientific output, fortifying this study's credibility and precision.

2.11. Assessment of Reporting Bias

Methods were used in this bibliometrics analysis of machine learning and blockchain to evaluate the potential for reporting bias in the reviewed scientific literature, which could create a risk of bias in the synthesis due to the lack of results. It is crucial to acknowledge that this study may exhibit a bias towards specific thesaurus-identified synonyms, such as IEEE, reflected in the search strategy, inclusion criteria, and data collection. Additionally, solely incorporating conference proceedings and documents with incomplete indexing may exclude valuable information necessary for gaining a thorough understanding of the topic being studied. Thus, when interpreting the findings and conclusions of the bibliometric study, it is crucial to consider these potential biases.

2.12. Assessment of Certainty

Methods were utilized to evaluate the certainty or confidence in the body of evidence for the obtained results. Dissimilar to primary studies, where certainty is evaluated on an individual basis, an overall assessment of certainty was performed in this bibliometric review. This was accomplished through the independent application of inclusion and exclusion criteria, as well as defining bibliometric indicators. In addition, this study reported potential biases in the methodological design and limitations that were discussed in the analysis. This comprehensive approach enables a thorough evaluation of the certainty of the evidence, taking into account the quality of bibliographic sources and possible limitations associated with bibliometric analysis.

It outlines the preliminary stage of identifying articles, displays the outcome of implementing the search tactics in each data source, and removes any redundant or duplicated materials.

3. Results

The study results are organized according to indicators of quantity and quality, examining key actors including authors, journals, and countries. A thorough analysis of the top publishers and most cited researchers is provided to determine their impact on the field. Structure indicators based on keywords are subsequently presented. Using the data collected, an analysis was conducted on the most frequently occurring keywords per year. Additionally, the research clusters and their growth were examined, along with an analysis of current and recurring topics. Based on these findings, this study proposes (1) a keyword classification based on their function, (2) an identification of research gaps, and (3) a future research agenda.

Next, the information collected through the bibliometric analysis is presented in terms of publications per year, main research referents, thematic evolution and keyword behavior in terms of their frequency and validity. In the historical and successive analysis, relevant information was found on the growth of scientific production in these areas of study. As can be seen in Figure 2, a quadratic polynomial growth was identified with 99.61% confidence, indicating a significant and accelerated increase in the number of publications over time. Specifically, the years with the highest number of articles published on machine learning AND blockchain are 2022, 2023, and 2021. These results provide a clear and quantitative view of the dynamism and current relevance of research in machine learning AND blockchain, which can be very useful for researchers, academics and professionals interested in these areas.

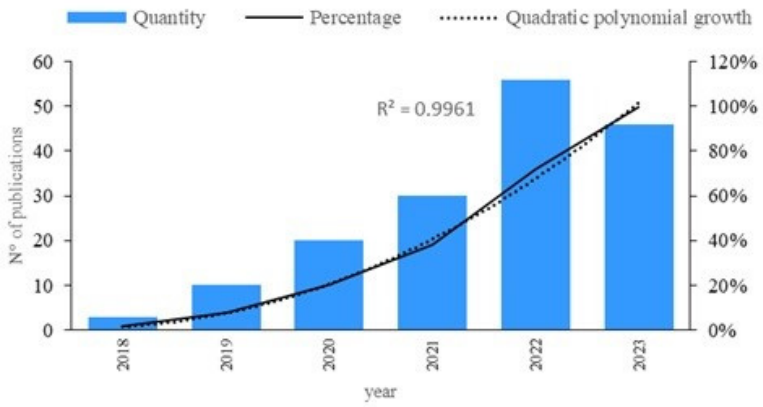

Figure 2. Publications per year. Own compilation from Scopus and Web of Science.

3.1. Analyze the Growth of Scientific Literature on Machine Learning and Blockchain

During the analyzed period of bibliometrics focused on machine learning and blockchain, there was a significant increase in scientific output during the years 2021, 2022, and 2023. Particularly in 2022, several studies pertaining to blockchain-based recommender systems

were prominent. These studies brought attention to the applications, challenges, and future opportunities of these systems, and were featured in *Computer Science Review* [21]. Another study, presented in *IEEE Transactions on Intelligent Transportation Systems*, explores the potential of utilizing blockchain and machine learning to safeguard vehicle networks [2].

In 2023, there was a rise in research focusing on blockchain and machine learning. Several studies explored the potential use of these technologies in future Smart Grids; one such study was published in Energies [22]. Additionally, another investigation examined the use of blockchain in the context of deep learning and discussed the associated challenges and opportunities [12].

In 2021, studies introduced a secure and private framework that combines blockchain and machine learning for smart cities powered by LoT. This framework was presented in an article published in *IEEE Transactions on Network Science and Engineering* [23]. Additionally, a study was conducted on solutions that use blockchain and artificial intelligence to combat epidemics similar to COVID-19. This topic is discussed in detail in an article published in *IEEE Access* [24].

These bibliographic references demonstrate the significance and diverse range of subjects covered in the scientific literature during the peak years of production in machine learning and blockchain. They illustrate how these technologies are being utilized and researched across various fields of study.

3.2. Analysis of Research References on Machine Learning and Blockchain

This study identified three primary sets of exceptional authors, illustrated in Figure 3. First, there are authors with high scientific productivity and impact, such as Hassija V and Chamola V. A second group includes authors with significant impact despite a low index of scientific productivity: Sikdar B, Saxena V, Goyal P, and Salah K. Finally, another group of highly productive authors with few citations is highlighted, primarily Zhang Y. These findings provide a comprehensive overview of the contributions and relevance of key authors in the fields of machine learning and blockchain. This allows for an appreciation of the various approaches and impact profiles in research.

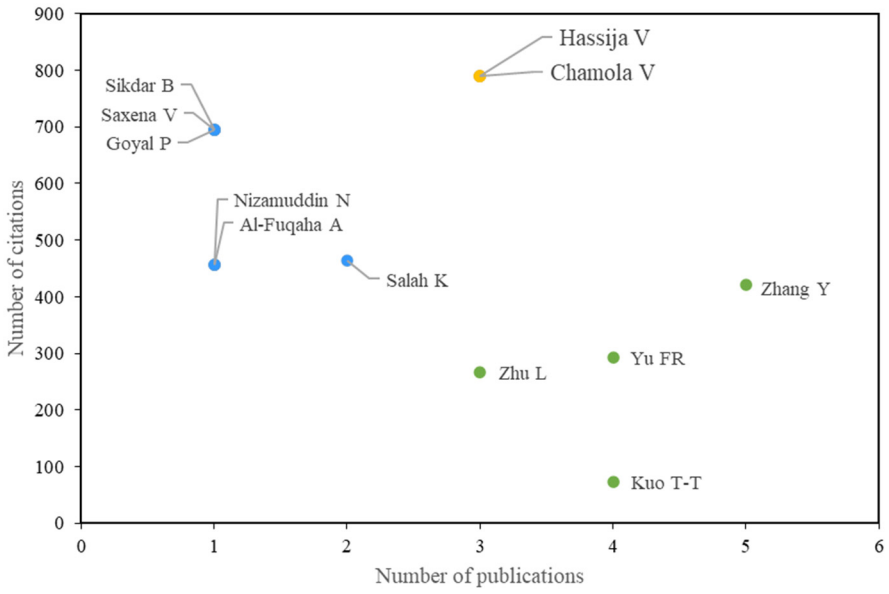

Figure 3. Main authors. Own work from Scopus and Web of Science.

Hassija V's significant research achievements are evident from an in-depth analysis of security in the supply chain that addresses application areas, security threats, and solution architectures [25], thus highlighting his importance in the scientific community. Furthermore, a comprehensive analysis on efficient and dependable drone communication was conducted in an article published in *IEEE Communications Surveys & Tutorials* [26].

Similarly, authors Sikdar B, Saxena V, and Goyal P demonstrated leadership in the field of research in their publication on Internet of Things (IoT) security in *IEEE Access*. They conducted a comprehensive review that explored application areas, security threats, and solution architectures [4].

Lastly, Zhang Y stood out as a productive scholar in the area of blockchain and federated learning. The authors addressed the use of blockchain and federated learning for data sharing while preserving privacy in the Industrial Internet of Things (IIoT) in their work published in *IEEE Transactions on Industrial Informatics* [27].

These investigations demonstrate the significance and impact of the authors' contributions in the fields of machine learning and blockchain, providing a comprehensive overview of their influence on scientific literature and progress in these research areas.

Similarly, Figure 4 identifies three prominent journal groups. First, there are journals that stand out for their scientific productivity and impact, including *IEEE Access* and the *IEEE Internet of Things Journal*. Another group of journals stands out for their impact despite having a low index of scientific productivity, such as *IEEE Transactions on Industrial Informatics and Internet of Things* (Netherlands). Finally, a group of reference scientific journals with high productivity but a low number of citations is identified, with a focus on *Sensors*. This study presents a comprehensive overview of the significance and relevance of leading journals in the fields of machine learning and blockchain, enabling the evaluation of distinct impact and productivity patterns in disseminating scientific research in these areas.

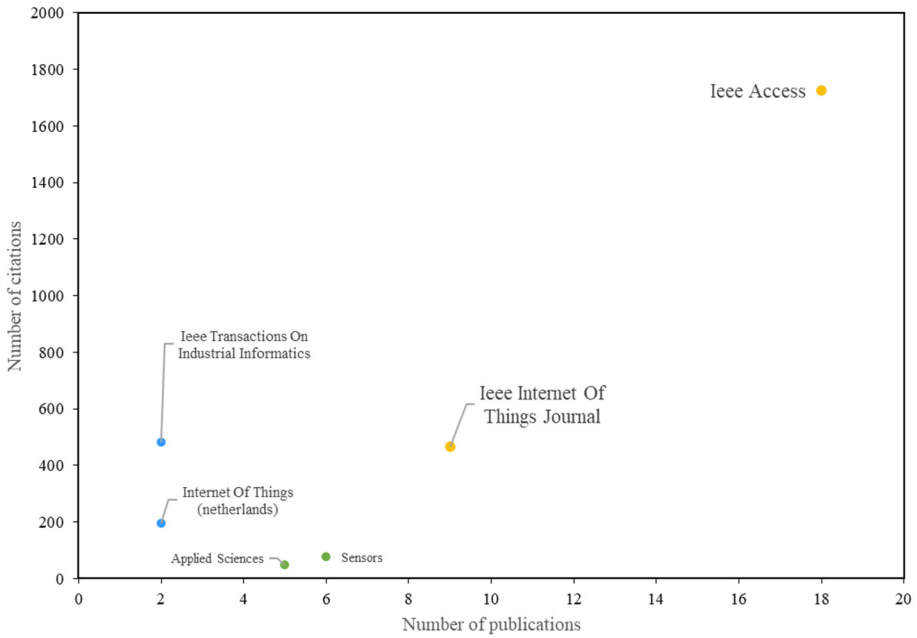

Figure 4. Main journals. Own elaboration from Scopus and Web of Science.

The contributions of these journals to the discipline are significant. Some of the papers published in *IEEE Access* offer a critical analysis of the implementation of blockchain in artificial intelligence, outlining both the obstacles and potential advantages in this area [12]. Moreover, *IEEE Access* conducted an investigation on security risks, application domains, and methods to circumvent such challenges to better understand the Internet of Things (IoT) [4]. In sharp contrast, the *IEEE Internet of Things Journal* has emerged as a reputable source of research on privacy and security in IoT.

In sharp contrast, the *IEEE Internet of Things Journal* has emerged as a reputable source of research on privacy and security in IoT. Several studies have investigated training support vector machines (SVM) to preserve privacy in encrypted data based on blockchain technology in smart cities [6]. Additionally, the *IEEE Internet of Things Journal* conducted a thorough examination of massive IoT enablement for 6G.

IEEE Transactions on Industrial Informatics is an essential publication on blockchain and machine learning applications for privacy in the Industrial Internet of Things (IIoT). Their research includes the use of blockchain and federated learning to protect data sharing privacy in IIoT [27].

Additionally, the Internet of Things magazine from the Netherlands plays a significant role in IoT security research. Several previous articles in this magazine have provided an objective analysis of the security issues in IoT and the utilization of machine learning, artificial intelligence, and blockchain technology as potential solutions [28].

Moreover, Sensorsha magazine featured a significant study on the incorporation of blockchain, IoT, and machine learning in smart manufacturing. Some publications by the organization focus on conducting research regarding multi-level quality control and implementing enhanced security measures for smart manufacturing through the collaborative use of machine learning and blockchain technologies [29].

These scientific journals are crucial for disseminating and advancing knowledge in the field of machine learning and blockchain. Moreover, they host important research that deals with diverse and fundamental topics in these fields, positioning them as prominent contributors to the scientific community.

Figure 5 identifies two main groups of highlighted countries. The first group consists of countries that stand out in both scientific productivity and impact, such as the United States, Australia, and India. The second group consists of countries that stand out for impact despite having a low index of scientific productivity, such as Singapore and Norway. The findings offer a comprehensive outlook on the contribution and significance of the prominent nations in the machine learning and blockchain fields. This enables an evaluation of distinct impact and productivity profiles in scientific study in these domains, making it a subject of noteworthy interest for analyzing trends and global partnerships in these disciplines.

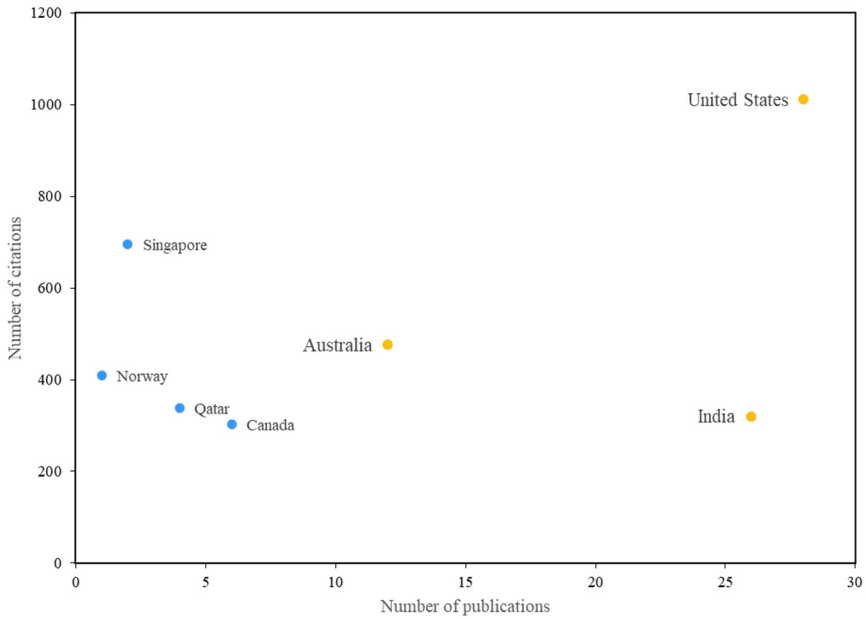

Figure 5. Main countries. Own compilation from Scopus and Web of Science.

In the realm of scientific production related to machine learning and blockchain, the United States, Australia, and India have emerged as noteworthy contributors, demonstrating both impact and productivity. These countries have generated significant and compelling research, enriching knowledge in this area. In particular, the United States has introduced a secure and privacy-preserving framework that integrates blockchain and machine learning for IoT enabled smart cities [30]. Similarly, the safeguarding of intelligent

energy networks through leveraging blockchain and deep learning was published in *IEEE Transactions on Industrial Informatics* [31].

Australia has also contributed significantly to the development of machine learning and blockchain in the context of Smart Grids. A published article in *Energies* [32] describes their research on distributed energy resources and the application of AI, IoT, and blockchain in Smart Grids. Moreover, in the *Journal of Network and Computer Applications*, a thorough study on the use of blockchain in 5G networks and other areas is presented [5].

India has served as a research benchmark for IoT security. The country has analyzed the challenges and solutions using machine learning, artificial intelligence, and blockchain technology [28]. Additionally, in the *Journal of Systems Architecture*, [29] presented a secure and trustworthy framework for sustainable smart cities based on blockchain and machine learning.

Singapore and Norway are considered benchmarks in research on machine learning and blockchain. In their study on IoT security in *IEEE Access*, [4] and their research on the use of blockchain and federated learning to share data while preserving privacy in the Industrial Internet of Things (IIoT) in *IEEE Transactions on Industrial Informatics* [27], respectively, these countries demonstrate their relevance in this field.

These nations have significantly contributed to the advancement and development of research areas on machine learning and blockchain through their high-impact research and scientific productivity.

In particular, the analysis of thematic evolution on these topics is discussed in Section 3.3.

In this bibliometric study on machine learning and blockchain, an exhaustive investigation was conducted to explore the thematic evolution in the scientific literature pertaining to this field. This study focused on the most frequently used keyword in each year between 2018 and 2023. The language used is clear, concise, and objective, with avoidance of biased or emotional language. All technical terms are explained, conventions and formats are correctly applied, and the grammar is error free. Figure 6 depicts the emergence of various crucial concepts, such as "artificial intelligence", during the initial year of 2018. Over the years, a shift was seen towards more relevant and detailed topics, including "Internet of Things" and "smart contract", which gained considerable prominence in recent research. These observations demonstrate trends and modifications in research methodologies over time, offering a dynamic outlook on the thematic growth of the fields of machine learning and blockchain in recent years.

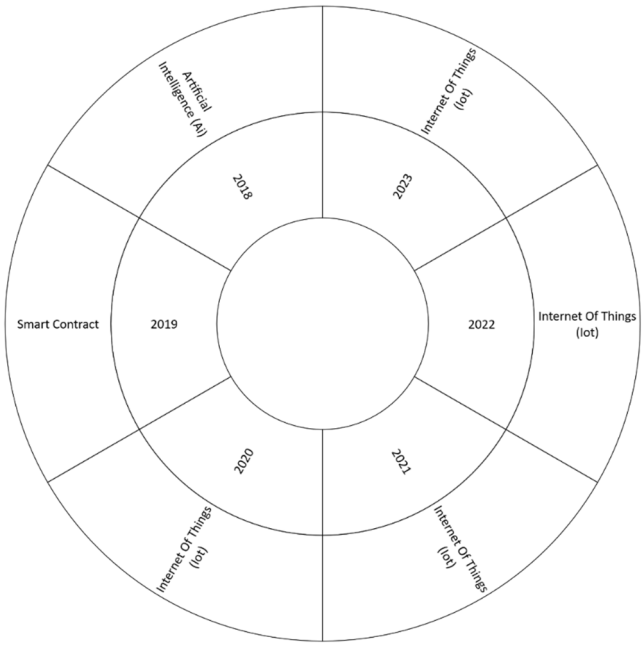

Figure 6. Thematic development. Own elaboration from Scopus and Web of Science.

The concept of artificial intelligence played a crucial role in early research. A significant study, which aided in strengthening this perspective, investigated its actual influence in the business sphere by conducting a business information survey [33]. This article, published in the *Business Information Review*, provides an in-depth overview of how artificial intelligence is transforming various business sectors and has become a central topic of scientific research. The initial interest in "artificial intelligence" provided a strong foundation for subsequent growth in the field and laid the groundwork for the development and expansion of knowledge in the fields of machine learning and blockchain.

In recent years, the concept of "smart contract" has become increasingly important in the fields of machine learning and blockchain technology. A review was conducted to present the open challenges in the application of blockchain for artificial intelligence, emphasizing the crucial role of smart contracts in this context [13]. These self-executing contracts enable the automation and secure execution of agreements and transactions, drawing the interest of researchers and professionals in the field. The augmented emphasis on smart contracts enriches the literature concerning the combined employment of blockchain and machine learning, spurring new research in this field [13].

Finally, the Internet of Things (IoT) has been extensively researched in recent years, as demonstrated by the number of studies conducted between 2020 and 2023 (refer to Figure 6). Numerous studies have explored the potential use of blockchain in networking beyond 5G [5]. Additionally, the healthcare sector has investigated the use of federated learning and blockchain in combination [34]. Similarly, blockchain and TinyML (machine learning on low-power devices) have been employed to enhance food supply chain security [28]. The analyses in this field demonstrate the IoT as a key platform for exploring inventive applications of machine learning and blockchain, which drives the thematic evolution of research into these technologies [5,34,35].

3.3. Analysis of Thematic Clusters on Machine Learning and Blockchain

Similarly, Figure 7 unveils the primary keyword co-occurrence network through six thematic clusters. The cluster in purple is the most prominent and includes terms such as "Internet of Things" and "Information Security". The green cluster follows, containing terms such as "federated learning", "deep learning", "intrusion detection", "transparency", and "privacy preserving". Other clusters in blue, yellow, red and light blue were identified, indicating further conceptual affinities found in the analyzed literature. These findings enable the visualization and comprehension of keyword relationships and clusters, presenting a comprehensive outlook on current trends and thematic approaches in research on machine learning and blockchain.

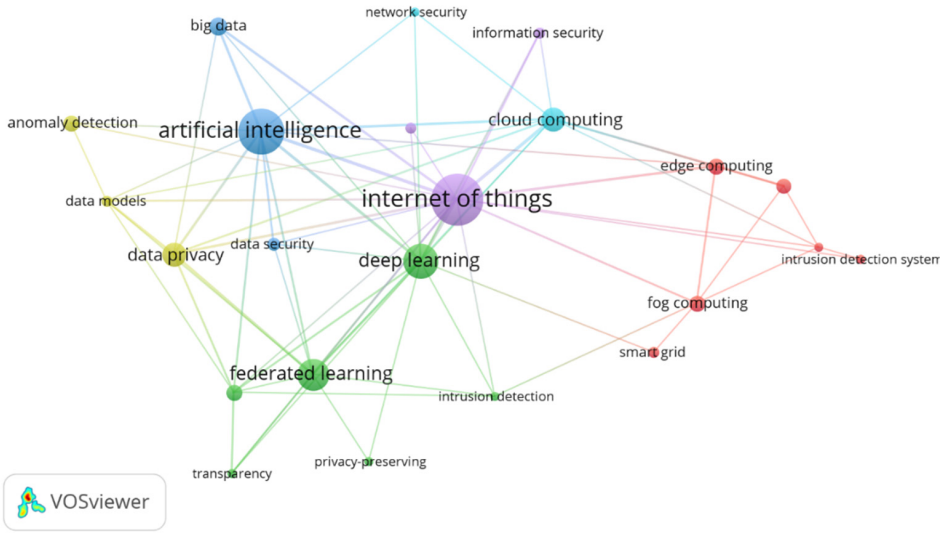

Figure 7. Keyword co-occurrence network. Own elaboration from Scopus and Web of Science.

The bibliometric analysis identifies several thematic clusters in the main keyword co-occurrence network, with the purple cluster being the most significant. The cluster comprises keywords such as "Internet of Things" (IoT) and "Information Security" that indicate a strong thematic affinity between the concepts. Previous research has highlighted the crucial significance of securing and maintaining privacy in the Internet of Things (IoT) through machine learning and blockchain technologies. Research has extensively analyzed the potential threats and countermeasures in this field [36], along with exploring the prospects of sustainable computing utilizing advanced IT technologies [37]. The purple cluster emphasizes the IoT and Information Security intersection as an important focus in the domain of machine learning and blockchain.

The second largest cluster, indicated by the color green, comprises keywords such as "federated learning", "deep learning", "intrusion detection", "transparency", and "privacy preserving". This cluster is noteworthy for its grouping of terms related to advanced machine learning and security technologies in the context of blockchain and machine learning. Research has provided an overview of the use of data science and AI in the FinTech industry, where technologies such as federated learning can play a crucial role [38]. Additionally, other studies have proposed privacy solutions for IoT in smart city environments that are blockchain-enabled and machine learning-based [39]. Finally, the study by [29] investigated methods to enhance the security of the food supply chain using technologies such as blockchain and TinyML. These investigations underscore the significance and intrigue surrounding the creation of solutions that fuse the effectiveness of machine learning with the advantages of blockchain technology to confront pivotal issues in diverse fields.

3.4. Frequency and Conceptual Validity Analysis around Machine Learning and Blockchain

A Cartesian plane is proposed with the frequency of keyword usage plotted on the *X*-axis and the validity of such usage on the *Y*-axis (see Figure 8). This classification system sorts keywords into four quadrants, each of which visually represents the corresponding keyword trends over time. In quadrant 2, keywords that are both rare and highly current are observed, indicating that these terms are emerging.

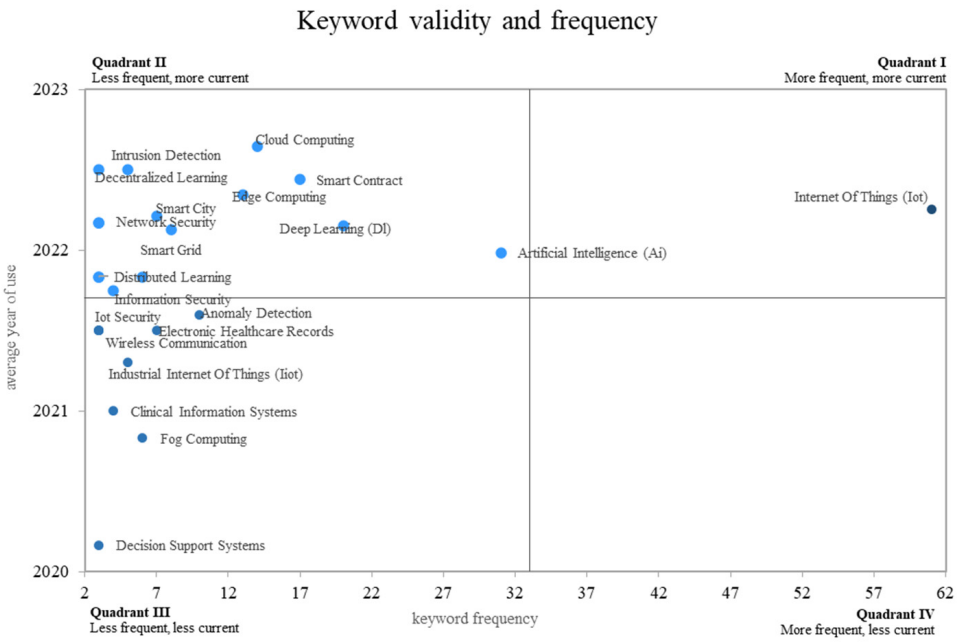

Figure 8. Validity and frequency of keywords. Own elaboration from Scopus and Web of Science.

Some emerging keywords include cloud computing, intrusion detection, decentralized learning, smart contracts, and edge computing. In contrast, quadrant 1 comprises growing and established concepts such as the Internet of Things. This chart offers insights into the

evolving trends in machine learning and blockchain research, illuminating emerging and established patterns. It aids in comprehending the thematic progression of the field and emphasizes the presently significant concepts.

The analysis of the bibliometric study on the Cartesian plane's quadrant 2 indicates the emergence of new concepts in the scientific field of machine learning and blockchain. This study identifies relevant emerging concepts, including "cloud computing", "intrusion detection", and "decentralized learning", which represent areas of great interest and rapid growth in the scientific community.

Firstly, "Cloud computing" has become a crucial concept in the age of digital transformation, enabling users to access and store data and computing resources via the internet. Several analyses have conducted extensive research on the Internet of Things (IoT) ecosystem and advanced technologies, exploring cloud computing applications and architectures [40]. This approach is essential for tackling the challenges of handling and processing vast amounts of data produced by IoT devices and machine learning applications.

Additionally, "intrusion detection" is a vital facet of system and network security. To address this, recent studies have proposed a novel energy exchange framework for smart networks based on deep learning and blockchain [41]. With the rise of cyber threats, it's crucial that we have sophisticated and efficient methods to detect and prevent intrusions. Deep learning techniques, combined with blockchain technology, offer potential benefits for enhancing security and trust in critical systems, such as smart power grids.

Finally, decentralized learning is a promising approach that aims to preserve data privacy and confidentiality during learning tasks in distributed environments. Researchers have suggested a trusted, blockchain-based distributed learning framework for multi-center medical imaging. With an increasing emphasis on data privacy and collaboration within the healthcare industry, distributed learning has emerged as an appealing option to enhance security and efficiency in sharing sensitive information [42].

In bibliometrics focusing on machine learning and blockchain, the "Internet of Things" stands out as a prominent and established concept, as highlighted in quadrant 1 of the Cartesian plane. The significance of this term lies in its present importance and future impact. The Internet of Things (IoT) pertains to the interconnecting of everyday objects and devices through the internet, facilitating the gathering and transmission of data for the enhancement of efficiency and individuals' quality of life.

Studies conducted a comprehensive survey of recent security trends in IoT, reflecting the topic's importance in the context of data protection and privacy in an increasingly connected world [43]. Additionally, extensive research has been carried out on anomaly detection schemes in IoT networks utilizing machine learning algorithms, indicating the growing attention given to this area in the field of cybersecurity and IoT system reliability [44].

The ongoing progress and evolution of technologies and applications within the Internet of Things (IoT) field ensures that it will remain a pertinent and auspicious focus of research, capable of revolutionizing diverse industries and enhancing people's global quality of life.

3.5. Classification of Keywords on Machine Learning and Blockchain According to Their Function

On the contrary, this literature review classifies the principal keywords identified as growing and emerging. The aim of this classification is to expand upon the associated tools, their main characteristics, and applications within the theme. Please refer to Table 1 for further details.

Table 1. Classification of keywords according to their function. Own elaboration from Scopus and Web of Science.

Keyword	Related Tools	Applications	Features
Cloud Computing	Amazon Web Services (AWS), Microsoft Azure, GCP	Data storage and access, web hosting	Scalability, on-demand resources
Intrusion Detection	Snort, Suricata, Bro/Zeek	Network security, cyber threat detection	Real-time monitoring, anomaly detection
Decentralized Learning	Federated Learning, Homomorphic Encryption	Collaborative AI models, decentralized data access	Privacy-preserving algorithms, Secure data processing
Smart Contract	Solidity, Ethereum, Neo	Decentralized finance, supply chain management	Self-executing contracts, trustless transactions
Edge Computing	Raspberry Pi, Nvidia Jetson Nano, Azure IoT Edge	Internet of Things (IoT), real-time analytics	Local data processing, low latency
Internet of Things	MQTT, CoAP, LoRaWAN	Smart home automation, industrial IoT	IoT device communication, long-range communication

3.6. Investigative Gaps

Based on the research findings and the analysis of various approaches, particularly the concept or keyword-based ones, this study identified several categories of research gaps, including thematic, geographic, interdisciplinary, and temporal gaps. These gaps pose challenges for future research and warrant further investigation to identify potential solutions, as demonstrated in Table 2.

Table 2. Research gaps. Own work from Scopus and Web of Science.

Categories	Gaps	Rationale	Questions for Future Researchers
Thematic Gaps	1. Integrating blockchain and machine learning techniques into specific applications.	Despite the growing interest in the convergence of blockchain and machine learning, there is a lack of research focused on the integration of these technologies into concrete applications, such as digital identity management, healthcare, or supply chains.	What are the most effective approaches for integrating blockchain and machine learning techniques into specific applications? How can the technical and privacy challenges associated with this integration be addressed?
	2. Exploring new machine learning architectures to improve efficiency in blockchain environments.	Since blockchain is inherently slower than traditional databases, research is needed to develop optimized machine learning architectures that can run efficiently in decentralized and distributed environments.	How can one design and evaluate machine learning architectures that are suitable for blockchain environments? What techniques can improve the efficiency and scalability of machine learning models in these environments?
Geographic Gaps	1. Lack of research on the impact of blockchain and machine learning in emerging economies.	Most research on machine learning and blockchain has focused on developed countries, leaving a gap in understanding how these technologies can address specific challenges and provide opportunities in emerging economies.	What are the potential use cases for blockchain and machine learning in emerging markets? How can these technologies address socio-economic and development challenges in these contexts?
	2. Limited representation of research in non-English speaking countries.	Much of the literature on machine learning and blockchain comes from Anglophone countries, which could lead to a geographical bias in the understanding and application of these technologies in different cultural and linguistic contexts.	What are the specific challenges of adopting blockchain and machine learning in non-English speaking countries? How can language and cultural barriers be overcome in the research and application of these technologies?

Table 2. *Cont.*

Categories	Gaps	Rationale	Questions for Future Researchers
Interdisciplinary Gaps	1. Research that addresses the convergence of blockchain, machine learning, and the Internet of Things (IoT).	While there is research on the combination of blockchain and machine learning, more work is needed to explore how these two technologies can be integrated with the Internet of Things to address challenges and create new solutions.	What are the most promising approaches for the convergence of blockchain, machine learning, and IoT? What applications and use cases are emerging from this convergence and how can they be optimized?
	2. Analysis of the ethical and privacy implications of the convergence of blockchain and machine learning.	The intersection of blockchain and machine learning raises ethical and privacy challenges, such as the management of personal data and the transparency of machine learning algorithms. More research is needed to address these issues.	How can ethical frameworks for the responsible use of blockchain and machine learning be developed together? What privacy considerations should be taken into account when developing solutions based on these technologies?
Temporary Gaps	1. Studies analyzing the long-term evolution of the theme in machine learning and blockchain.	Although a thematic evolution from artificial intelligence to the Internet of Things and smart contracts has been identified, a long-term perspective is needed to understand how this evolution will continue and what new trends will emerge.	What are the possible future research directions in machine learning and blockchain? How can the scientific community anticipate and prepare for emerging challenges in this field?
	2. Prospective research on the potential applications of integrating blockchain and machine learning in emerging industries.	As the technology advances, opportunities will arise in emerging industries such as renewable energy, agriculture, and manufacturing. More prospective research is needed to identify and evaluate these opportunities.	What are the emerging applications of the combination of blockchain and machine learning in industries such as energy and agriculture? How can these applications transform and improve these industries in the future?

3.7. Research Agenda

Based on the results, analysis, and derived components, we present the primary research agenda in machine learning and blockchain, shown in Figure 9, prioritizing the most studied, growing, and emerging topics.

In this sense, IoT has transformed our interaction with the world by facilitating the connection of devices and systems via the collection and exchange of data. In the realm of machine learning and blockchain, the integration of IoT has presented extensive research prospects. For further study, researchers can delve deeper into developing machine learning algorithms that are more scalable and efficient in processing and analyzing the vast amount of data generated by IoT devices. This would facilitate more informed and precise decision making. Furthermore, the utilization of blockchain technology can be investigated to safeguard security and trust in IoT data, hence ensuring the accuracy and reliability of information, and promoting secure collaboration between connected devices and systems.

On the flip side, artificial intelligence (AI) has demonstrated remarkable efficacy in diverse applications, and its fusion with blockchain technology has the potential to amplify its influence. Subsequent research should center on enhancing the interconnectivity and scalability of AI systems by leveraging blockchain technology. Developing more robust and secure AI models to detect anomalies in complex and highly dynamic data is crucial for safeguarding critical systems against attacks and errors. Additionally, the integration of AI and blockchain technology can potentially foster the development of smart cities, leading to optimized resource utilization, efficient urban planning, and enhanced quality of life for citizens.

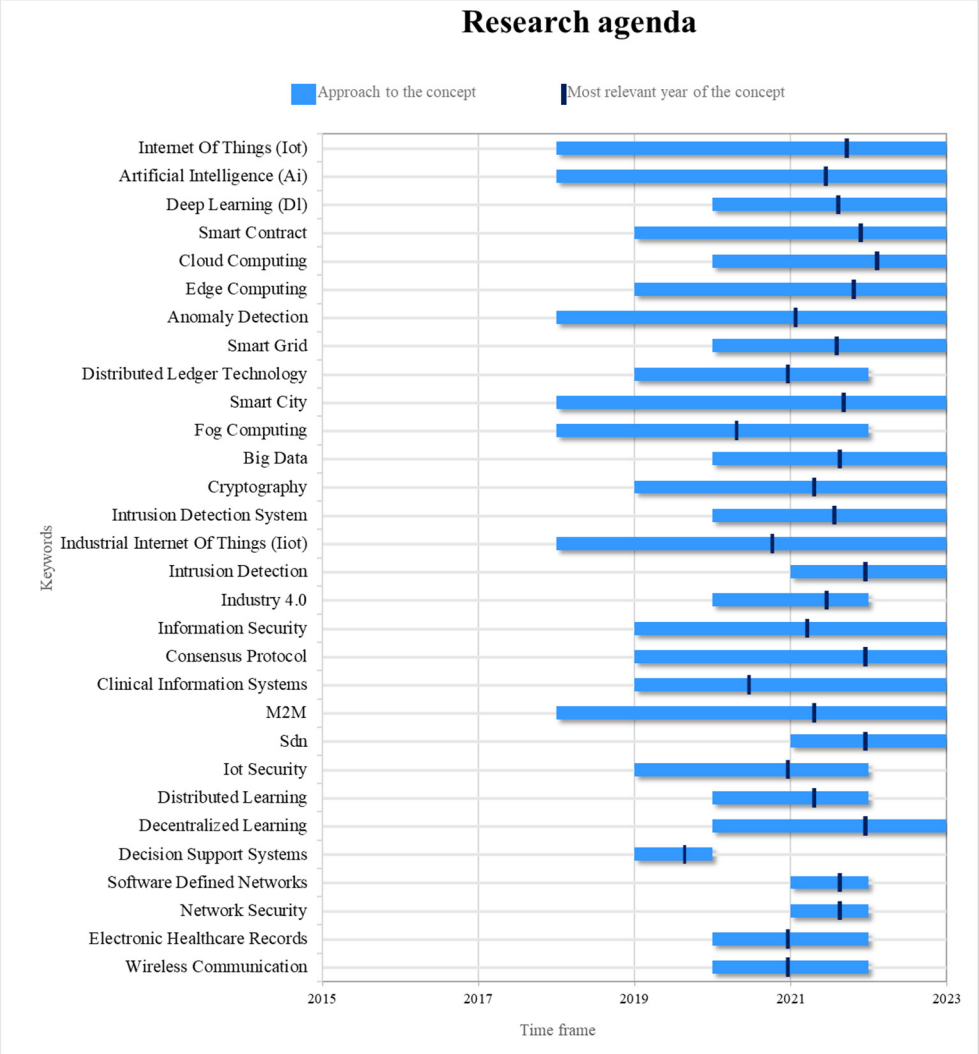

Figure 9. Research agenda. Compiled from Scopus and Web of Science.

In turn, anomaly detection is a crucial aspect of machine learning and blockchain. It facilitates the identification of atypical data patterns, ensuring the security and integrity of vital systems. Going forward, deep learning techniques and semi-supervised learning models can be explored for enhancing the precision and efficiency of anomaly detection algorithms. Furthermore, it is worth investigating how blockchain technology can enhance the dependability of anomaly detection outcomes by allowing for the tracing and verification of data and analysis results. Additionally, it is pertinent to examine how anomaly detection can be utilized in particular contexts, such as cybersecurity and predictive maintenance in industry, to tackle existing and burgeoning challenges in those fields.

Smart cities continue to be a growing field of research, and the integration of machine learning and blockchain technology can bolster their advancement. Potential future studies may explore how machine learning techniques can optimize urban planning and resource management in smart cities. This may lead to increased efficiency and sustainability in infrastructure and service usage. It is essential to examine how blockchain technology can enhance security and privacy in smart cities, safeguarding personal data and promoting trust in adopting smart services and technologies. Furthermore, it can examine how artificial intelligence can enhance urban mobility and data-driven decision making for providing tailored services that cater to citizens' diverse requirements.

IIoT has revolutionized the industry, automating and optimizing industrial procedures. Future research can explore how machine learning techniques can analyze and process

data generated by IIoT devices to increase efficiency and reliability in production and resource management. Additionally, the potential of blockchain technology in improving traceability and transparency in industrial supply chains can be investigated, ensuring the authenticity of data and preventing fraudulent practices. Similarly, it is crucial to investigate the advancements of AI and blockchain systems in order to enhance predictive maintenance in the industrial sector, resulting in increased availability and dependability of industrial assets.

Intrusion detection, a term that is currently gaining significance in the domains of machine learning and blockchain, pertains to the recognition and deterrence of malevolent activities in computer systems and networks. The protection of data and operations in digital environments is crucial. Future research could investigate how machine learning can enhance intrusion detection systems by developing more advanced and precise algorithms to detect patterns and anomalies in network traffic. Furthermore, incorporating blockchain technology into intrusion detection systems has the potential to enhance the immutability and transparency of activity logs. By doing so, it can increase trust in detection reports and encourage cooperation within the cybersecurity sphere to exchange information on emerging threats.

On the other hand, Software-Defined Networking (SDN) is an important emerging term in the fields of machine learning and blockchain. Its goal is to decouple the control plane from the data plane in communication networks, with potential benefits for flexibility and efficiency in network management. This is particularly relevant in the context of rising complexity and demand for connectivity. Future research can explore the potential for machine learning to enhance the management and allocation of resources within SDN networks, enabling a more flexible and self-governing adaptation of network infrastructure in response to changing demands. Furthermore, integrating SDN and blockchain technology could establish a secure framework to ensure the integrity and authenticity of network operations, prevent denial of service attacks, and promote the reliability of data in transit.

Deep learning is a term that has rapidly transformed the field of machine learning and blockchain. Its importance is attributed to its capacity to methodically process and scrutinize enormous amounts of intricate data, resulting in far-reaching advancements in speech recognition, computer vision, and natural language processing. In the future, there is potential for investigating how deep learning can enhance the efficiency and accuracy of consensus algorithms on the blockchain to enhance scalability in high-demand systems. Additionally, exploring the use of deep neural networks may bolster fraud and anomaly detection in blockchain transactions, ultimately augmenting security in the digital asset exchange. The future research on deep learning may tackle challenges such as improving the interpretability and explainability of deep learning models within the context of blockchain. This is crucial for comprehending the inner workings of AI systems and ensuring confidence in the outcomes.

Additionally, cloud computing is a significant term in the domain of machine learning and blockchain, denoting the delivery of computing and storage services via the internet. Its significance lies in the capacity to offer a scalable and adaptable infrastructure for executing applications and handling vast amounts of data. Subsequent research can concentrate on how cloud computing can elevate the efficiency and performance of machine learning algorithms, enabling accelerated and less expensive processing of intricate tasks. In addition, integrating blockchain technology into cloud computing could enhance the security and privacy of stored data, while providing greater control and ownership of information. Further research might examine how cloud computing can ease the deployment and operation of blockchain-based decentralized applications, which could encourage wider adoption of decentralized technologies in finance, healthcare, and supply chain management.

The Smart Grid, or the Intelligent Electrical Network, is an important emerging term in the field of machine learning and blockchain. It seeks to enhance electrical network efficiency, reliability, and sustainability through digital and communication technologies. It is necessary because of current energy sector challenges, including renewable energy

integration and efficient demand management. Future research can explore the potential for machine learning to enhance the management and control of the Smart Grid, enabling more accurate forecasting of energy demand and network behavior. Furthermore, the integration of blockchain technology into the Smart Grid could enhance the security and reliability of energy transactions between users and suppliers, thereby supporting the adoption of more equitable and decentralized market models. Future research could investigate the potential of artificial intelligence to optimize energy consumption and production within the Smart Grid, which could lead to increased efficiency and reduced costs in global energy supply.

Distributed ledger technology (DLT) plays a significant role in the field of machine learning and blockchain, as it provides a secure, decentralized system for recording and verifying transactions. Currently, distributed ledger technology (DLT), commonly known as blockchain technology, has gained widespread adoption in various industries and applications, including cryptocurrency and supply chain tracking. However, for DLT to maintain relevance, further studies are required to address challenges such as scalability and privacy. Future research could concentrate on developing more efficient consensus algorithms that promote higher performance and lower barriers to adoption in high-demand systems. Moreover, investigation can be conducted to ascertain the integration of machine learning techniques into DLT for security enhancement and detection of malicious attacks. This will bolster trust and safeguard sensitive information during transactions.

Apart from this, fog computing is gaining significance in the realm of machine learning and blockchain by decentralizing computing and data processing to devices in close proximity to the network edge. This enhances latency while reducing the workload on data centers. Future studies should investigate the integration of machine learning techniques into fog computing infrastructure to enhance intelligent and autonomous real-time decision making. Additionally, the potential of blockchain technology can be explored to enhance privacy and security in distributed computing environments. Future studies could address the scalability and reliability challenges of fog computing, allowing for broader adoption in Internet of Things (IoT) applications and smart systems. This, in turn, could position fog computing as a relevant and promising concept for the future of the machine learning and blockchain fields.

Finally, Industry 4.0 is a pertinent term in the domain of machine learning and blockchain. It denotes the inclusion of advanced technologies, such as artificial intelligence and blockchain, in the production and manufacturing processes. To ensure its relevance, upcoming research could concentrate on the implementation of machine learning for process optimization and failure prognosis in industrial equipment. Similarly, one may inquire about the potential of blockchain technology to enhance traceability and security in the supply chain, thereby enabling increased transparency and dependability in the exchange of goods and services. Furthermore, the combination of machine learning and blockchain technology can potentially simplify and decentralize industrial operations. This could reaffirm Industry 4.0's relevance as a pioneering concept in the digital transformation of companies and industry.

4. Discussion

4.1. Contrast with Other Studies

The utilization of blockchain technology and machine learning holds vast potential in various fields. It especially brings about a revolutionary transformation in the healthcare infrastructure. Nouman and Muneer [45] performed an extensive analysis of the literature, collecting data to predict heart diseases, and demonstrated that the Support Vector Machine (SVM) methodology offers exceptional precision. As found in this study, research hotspots include machine learning, blockchain, smart healthcare, the Internet, and the Internet of Things, based on keyword analysis. Li et al. [46] conducted a bibliometric visualization analysis through WOS to review the utilization of machine learning and blockchain technology in the smart medical industry. They identified the countries with the highest production and the main research topics in this field. The results revealed a

notable level of scientific production in India, China, and Saudi Arabia, with an increasing interest in artificial intelligence, sensors, and IoT-related topics. These findings align with the conclusions drawn in this article.

Cheng et al. [47] found that integrating two disruptive technologies can be effective and feasible. They conducted a literature review demonstrating the technologies' application in cancer survivorship care and effectively providing information on their integration. In terms of future research, the authors recommend exploring broader and deeper integration of the technologies most notable in cancer care, in line with the recommendations of this study. Karger's research [48] also reviewed these technologies to address research gaps and provide a comprehensive overview of this emerging field. According to their research, which suggests multiple interactions with other technologies and disciplines are feasible when combining machine learning and blockchain, it is highly probable that future research on this amalgamation will see dynamic growth.

Ekramifard and colleagues [49] conducted a systematic literature review examining the integration of blockchain and artificial intelligence. Their findings suggest that artificial intelligence algorithms can optimize the design and operation of blockchain. These findings align with the research emphasized in this article, which demonstrates the capability of machine learning in enhancing blockchain processes for cybersecurity threats. [36].

Blockchain technology, governed by intelligent machine learning algorithms, may enable automatic detection and defense against attacks. Several studies have analyzed blockchain management and the integration of machine learning in an IoT environment within 5G networks. According to the study conducted by Miglani and Kumar [17], collaborative applications of blockchain and machine learning are still in their infancy, presenting several research challenges that need to be addressed.

Comparable studies [50] have identified provisions and the need to establish a foundation for future research and recommendations on the implementation of blockchain technology. This is especially important considering the nature of literature reviews. Therefore, scholars have emphasized the capabilities of machine learning algorithms in incorporating Industry 4.0 technologies [51] and processing data through Big Data [52] to generate value [53] and achieve security objectives in supply chain management [54]. Such emphasis on the incorporation of technology and data is aimed at improving the logistics processes and ensuring security in supply chain management.

Unlike previous literature reviews on the integration of machine learning and blockchain technologies, this study offers a unique perspective on the importance of security and privacy in today's world. By providing a research agenda with potential for the future, the study highlights emerging themes that can be further explored, as well as related themes that are losing relevance and frequency. Therefore, this study precisely contributes to the approach of security and privacy, with integration of Industry 4.0 technologies, and future research directions. Resultantly, scholars gain insight into current trends, diminishing interests, and potential for future exploration.

4.2. Implications

The significance of blockchain and machine learning technologies investigated in this study is demonstrated by their remarkable increase in published works in recent years. This research enhances existing knowledge by providing an objective perspective on their progress and potential for integration. The logical and clear structure of this study also promote comprehensibility and renders technical terms easily understandable. The results of this analysis highlight the capacity of this research domain to address concerns related to cybersecurity in the context of digital commerce and logistics management, within various sectors.

Additionally, this study demonstrates the conceivable applications of combining machine learning and blockchain technology. This article focuses on a range of applications, including threat and anomaly detection, smart contract implementation for privacy protection, secure storage assurance, and network security fortification. Additionally, it highlights

the integration of machine learning with other Industry 4.0 technologies, such as the Internet of Things (IoT) and Big Data, promoting compatibility. This study presents a research agenda to the scientific community, identifying the most promising areas for the integration of machine learning and blockchain techniques. Objective evaluations are excluded except when clearly marked as such. The agenda offers a broad and trend-oriented view of the research, providing researchers with a clear perspective of the achieved and achievable objectives in this captivating interdisciplinary field.

Conducting a bibliometric study on machine learning and blockchain and analyzing its thematic evolution hold significant practical implications for guiding and enriching academic research, as well as for developing technologies and solutions in the field. Subjectivity has been avoided and a clear, concise and logical structure has been maintained throughout the sentence. The sentence adheres to conventional academic writing, with high-level, standard language and a formal register throughout. Consistent technical terms have been used, and balanced language has been employed throughout. The spelling, vocabulary, and grammar conform to American English. The identification of the thematic evolution, starting with a focus on artificial intelligence and transitioning to aspects connected to the Internet of Things and smart contracts, indicates a distinct research direction. This highlights areas of high relevance and interest for both the scientific community and industry.

The identification of the primary thematic cluster, indicating the conceptual relationship between terminologies such as the Internet of Things and Information Security, facilitates valuable insights into the most interrelated and promising research segments in the field. This, in turn, affects the detection of suitable areas for joint and collaborative research and the development of comprehensive solutions and applications that tackle the privacy and security concerns linked with the Internet of Things.

Furthermore, the analysis of keyword frequency and validity indicates the emergence of concepts such as cloud computing, intrusion detection, decentralized learning, smart contracts, and edge computing, along with the increasing significance of the Internet of Things. These findings are valuable to researchers as they identify emerging and expanding research areas that can inspire the exploration of new techniques, approaches, and applications in machine learning and blockchain.

Practically speaking, these findings enable researchers and professionals in the field to concentrate on areas of study and application that will have a significant impact on the scientific community and industry. Furthermore, this study can offer valuable insights for allocating resources and making informed decisions in machine learning and blockchain research and development. Ultimately, the findings can lead to advancements and innovation in critical domains such as security, privacy, efficiency, and scalability of machine learning and blockchain-based systems, thereby promoting progress in society as a whole.

4.3. Limitations

Although this literature review on machine learning and blockchain provides valuable insights, there are limitations to consider that may impact result interpretation and generalization. Specifically, the utilization of the Scopus and Web of Science databases limited the scope of included publications, potentially introducing bias and affecting the precision of bibliometric indicators. The employment of software such as Microsoft Excel® and VOSviewer® for examining and displaying data may have constraints on handling vast amounts of information and identifying intricate patterns in the keyword co-occurrence network.

Another limitation to consider is that the bibliometrics focused solely on publications listed in the Scopus and Web of Science databases, potentially excluding research published on other platforms or in open access journals. This may have limited the sample's representativeness and overlooked significant contributions to the fields of machine learning and blockchain. Due to the ever-changing nature of scientific research, certain findings and

trends identified in bibliometrics may change over time, necessitating periodic updates to obtain an accurate and current understanding of the research landscape in the field.

Another limitation of this study concerns its scope. Although we conducted a thorough review of the literature, we did not use a highly structured methodology, such as a systematic literature review, which could have produced more insights into the topic under investigation. Additionally, we purposely excluded conference proceedings, which could have provided different perspectives on this developing field of study. It is suggested that future research consider these conference proceedings to enhance the perspective offered in this study.

Despite its limitations, bibliometrics offers a thorough and organized perspective of the scientific output in machine learning and blockchain, establishing a firm basis for comprehending trends, areas of focus, and significant benchmarks in the field. The obtained results can provide valuable information to guide future research, facilitate collaborations, and advance knowledge and technology in this dynamic and interdisciplinary field.

5. Conclusions

Among the key findings, it is apparent that the years 2021, 2022, and 2023 have garnered the most interest in the fields of machine learning and blockchain. This suggests that the academic community has increasingly focused on these areas in the recent past. This study's results reveal that the number of scientific articles on machine learning and blockchain follows a quadratic polynomial growth trend, suggesting a significant increase in scientific production in this field. These findings indicate that further developments can be anticipated in the future.

Additionally, the predominant academic sources found in the existing literature are Hassija V, Chamola V, Sikdar B, Saxena V, and Goyal P, as well as the scholarly journals *IEEE Access* and *IEEE Internet of Things Journal*. Similarly, the United States, Australia, and India have notably contributed to scientific output concerning the subject matter of this research.

The conceptual perspective evolved thematically over time, shifting from an initial emphasis on artificial intelligence to more recent themes such as the Internet of Things and smart contracts, reflecting the varied research interests in machine learning and blockchain. The bibliometric study identifies crucial information regarding keyword frequency and credibility in this area of research. Consolidated terms such as the Internet of Things remain frequent over time, suggesting its significant role in research. On the other hand, it has been verified that newly arose terms such as cloud computing, intrusion detection, and decentralized learning have shown an increase in their frequency, indicating a rise in interest and focus on these areas. This presents opportunities for future investigations exploring the use and advancement of these new technologies and their integration with machine learning and blockchain to tackle current issues.

Overall, these findings indicate that research related to machine learning and blockchain is continuously evolving with a strong emphasis on innovative concepts such as the Internet of Things. This has led to a considerable diversification of research topics. Finally, the obtained conclusions establish a robust groundwork for constructing a research agenda that focuses on resolving major emerging problems, investigating innovative technology applications in diverse domains, and promoting international collaboration among researchers and nations in order to advance science in this highly regarded interdisciplinary field.

Author Contributions: Conceptualization, A.V.-A.; Methodology, A.V.-A., J.D.G.-R. and L.V.F.; Software, A.V.-A., P.R.-C. and G.S.S.; Validation, L.V.-M.; Formal analysis, L.V.F. and G.S.S.; Investigation, L.V.-M., P.R.-C. and G.S.S.; Resources, A.V.-A., L.V.F. and L.V.-M.; Writing—original draft, L.V.F. and P.R.-C.; Writing—review & editing, J.D.G.-R., L.V.-M. and G.S.S.; Supervision, J.D.G.-R.; Project administration, A.V.-A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The APC was funded by Universidad Señor de Sipán—USS.

Data Availability Statement: The data may be provided free of charge to interested readers by request through the correspondence email.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mrabet, H.; Alhomoud, A.; Jemai, A.; Trentesaux, D. A Secured Industrial Internet-of-Things Architecture Based on Blockchain Technology and Machine Learning for Sensor Access Control Systems in Smart Manufacturing. *Appl. Sci.* **2022**, *12*, 4641. [CrossRef]
2. Dibaei, M.; Zheng, X.; Xia, Y.; Xu, X.; Jolfaei, A.; Bashir, A.K.; Tariq, U.; Yu, D.; Vasilakos, A.V. Investigating the Prospect of Leveraging Blockchain and Machine Learning to Secure Vehicular Networks: A Survey. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 683–700. [CrossRef]
3. Mendis, G.J.; Wu, Y.; Wei, J.; Sabounchi, M.; Roche, R. A Blockchain-Powered Decentralized and Secure Computing Paradigm. *IEEE Trans. Emerg. Top. Comput.* **2021**, *9*, 2201–2222. [CrossRef]
4. Hassija, V.; Chamola, V.; Saxena, V.; Jain, D.; Goyal, P.; Sikdar, B. A Survey on IoT Security: Application Areas, Security Threats, and Solution Architectures. *IEEE Access* **2019**, *7*, 82721–82743. [CrossRef]
5. Nguyen, D.C.; Pathirana, P.N.; Ding, M.; Seneviratne, A. Blockchain for 5G and beyond networks: A state of the art survey. *J. Netw. Comput. Appl.* **2020**, *166*, 102693. [CrossRef]
6. Guo, F.; Yu, F.R.; Zhang, H.; Li, X.; Ji, H.; Leung, V.C.M. Enabling Massive IoT Toward 6G: A Comprehensive Survey. *IEEE Internet Things J.* **2021**, *8*, 11891–11915. [CrossRef]
7. Gayialis, S.P.; Kechagias, E.P.; Papadopoulos, G.A.; Kanakis, E. A Smart-Contract Enabled Blockchain Traceability System Against Wine Supply Chain Counterfeiting. In *Advances in Production Management Systems. Smart Manufacturing and Logistics Systems: Turning Ideas into Action. APMS 2022. IFIP Advances in Information and Communication Technology*; Kim, D.Y., von Cieminski, G., Romero, D., Eds.; Springer: Cham, Switzerland, 2022; Volume 663, pp. 477–484. [CrossRef]
8. Kechagias, E.P.; Gayialis, S.P.; Papadopoulos, G.A.; Papoutsis, G. An Ethereum-Based Distributed Application for Enhancing Food Supply Chain Traceability. *Foods* **2023**, *12*, 1220. [CrossRef]
9. Wan, Y.; Gao, Y.; Hu, Y. Blockchain application and collaborative innovation in the manufacturing industry: Based on the perspective of social trust. *Technol. Forecast. Soc. Chang.* **2022**, *177*, 121540. [CrossRef]
10. Gong, Y.; Wang, Y.; Frei, R.; Wang, B.; Zhao, C. Blockchain application in circular marine plastic debris management. *Ind. Mark. Manag.* **2022**, *102*, 164–176. [CrossRef]
11. Lemos, C.; Ramos, R.F.; Moro, S.; Oliveira, P.M. Stick or Twist—The Rise of Blockchain Applications in Marketing Management. *Sustainability* **2022**, *14*, 4172. [CrossRef]
12. Shafay, M.; Ahmad, R.W.; Salah, K.; Yaqoob, I.; Jayaraman, R.; Omar, M. Blockchain for deep learning: Review and open challenges. *Clust. Comput.* **2023**, *26*, 197–221. [CrossRef]
13. Salah, K.; Rehman, M.H.U.; Nizamuddin, N.; Al-Fuqaha, A. Blockchain for AI: Review and Open Research Challenges. *IEEE Access* **2019**, *7*, 10127–10149. [CrossRef]
14. Kadian, K.; Garhwal, S.; Kumar, A. Quantum walk and its application domains: A systematic review. *Comput. Sci. Rev.* **2021**, *41*, 100419. [CrossRef]
15. Page, M.J.; Moher, D.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, 71. [CrossRef] [PubMed]
16. Dyzel, V.; Oosterom-Calo, R.; Worm, M.; Sterkenburg, P.S. Assistive Technology to Promote Communication and Social Interaction for People with Deafblindness: A Systematic Review. *Front. Educ.* **2020**, *5*, 578389. [CrossRef]
17. Miglani, A.; Kumar, N. Blockchain management and machine learning adaptation for IoT environment in 5G and beyond networks: A systematic review. *Comput. Commun.* **2021**, *178*, 37–63. [CrossRef]
18. Stahlschmidt, S.; Stephen, D. From indexation policies through citation networks to normalized citation impacts: Web of Science, Scopus, and Dimensions as varying resonance chambers. *Scientometrics* **2022**, *127*, 2413–2431. [CrossRef]
19. Effendy, F.; Gaffar, V.; Hurriyati, R.; Hendrayati, H. Analisis Bibliometrik Perkembangan Penelitian Penggunaan Pembayaran Seluler Dengan Vosviewer. *Intercom* **2021**, *16*, 10–17. [CrossRef]
20. Durieux, V.; Gevenois, P.A. Bibliometric Indicators: Quality Measurements of Scientific Publication. *Radiology* **2010**, *255*, 342–351. [CrossRef]
21. Himeur, Y.; Sayed, A.; Alsalemi, A.; Bensaali, F.; Amira, A.; Varlamis, I.; Eirinaki, M.; Sardianos, C.; Dimitrakopoulos, G. Blockchain-based recommender systems: Applications, challenges and future opportunities. *Comput. Sci. Rev.* **2022**, *43*, 100439. [CrossRef]
22. Mololoth, V.K.; Saguna, S.; Åhlund, C. Blockchain and Machine Learning for Future Smart Grids: A Review. *Energies* **2023**, *16*, 528. [CrossRef]
23. Kumar, P.; Kumar, R.; Srivastava, G.; Gupta, G.P.; Tripathi, R.; Gadekallu, R.T.; Xiong, N.N. PPSF: A Privacy-Preserving and Secure Framework Using Blockchain-Based Machine-Learning for IoT-Driven Smart Cities. *IEEE Trans. Netw. Sci. Eng.* **2021**, *8*, 2326–2341. [CrossRef]

24. Nguyen, D.C.; Ding, M.; Pathirana, P.N.; Seneviratne, A. Blockchain and AI-based Solutions to Combat Coronavirus (COVID-19)-like Epidemics: A Survey. *IEEE Access* **2021**, *9*, 95730–95753. [CrossRef] [PubMed]
25. Hassija, V.; Chamola, V.; Gupta, V.; Jain, S.; Guizani, N. A Survey on Supply Chain Security: Application Areas, Security Threats, and Solution Architectures. *IEEE Internet Things J.* **2020**, *8*, 6222–6246. [CrossRef]
26. Hassija, V.; Chamola, V.; Agrawal, A.; Goyal, A.; Luong, N.C.; Niyato, N.D.; Yu, F.R.; Guizani, M. Fast, Reliable, and Secure Drone Communication: A Comprehensive Survey. *IEEE Commun. Surv. Tutor.* **2021**, *23*, 2802–2832. [CrossRef]
27. Lu, Y.; Huang, X.; Dai, Y.; Maharjan, S.; Zhang, Y. Blockchain and Federated Learning for Privacy-Preserved Data Sharing in Industrial IoT. *IEEE Trans. Ind. Inform.* **2020**, *16*, 4177–4186. [CrossRef]
28. Mohanta, B.K.; Jena, D.; Satapathy, U.; Patnaik, S. Survey on IoT security: Challenges and solution using machine learning, artificial intelligence and blockchain technology. *Internet Things* **2020**, *11*, 100227. [CrossRef]
29. Shahbazi, Z.; Byun, Y.-C. Integration of Blockchain, IoT and Machine Learning for Multistage Quality Control and Enhancing Security in Smart Manufacturing. *Sensors* **2021**, *21*, 1467. [CrossRef]
30. Kumar, P.; Gupta, G.P.; Tripathi, R. TP2SF: A Trustworthy Privacy-Preserving Secured Framework for sustainable smart cities by leveraging blockchain and machine learning. *J. Syst. Archit.* **2021**, *115*, 101954. [CrossRef]
31. Keshk, M.; Turnbull, B.; Moustafa, N.; Vatsalan, D.; Choo, K.-K.R. A Privacy-Preserving-Framework-Based Blockchain and Deep Learning for Protecting Smart Power Networks. *IEEE Trans. Ind. Inform.* **2020**, *16*, 5110–5118. [CrossRef]
32. Kumar, N.M.; Chand, A.A.; Malvoni, M.; Prasad, K.A.; Mamun, K.A.; Islam, F.R.; Chopra, S.S. Distributed Energy Resources and the Application of AI, IoT, and Blockchain in Smart Grids. *Energies* **2020**, *13*, 5739. [CrossRef]
33. Carter, D. How real is the impact of artificial intelligence? The business information survey 2018. *Bus. Inf. Rev.* **2018**, *35*, 99–115. [CrossRef]
34. Baucas, M.; Spachos, P.; Plataniotis, K. Federated Learning and Blockchain-enabled Fog-IoT Platform for Wearables in Predictive Healthcare. *arXiv* **2023**, arXiv:2301.04511. [CrossRef]
35. Tsoukas, V.; Gkogkidis, A.; Kampa, A.; Spathoulas, G.; Kakarountas, A. Enhancing Food Supply Chain Security through the Use of Blockchain and TinyML. *Information* **2022**, *13*, 213. [CrossRef]
36. Waheed, N.; He, X.; Ikram, M.; Usman, M.; Hashmi, S.S.; Usman, M. Security and Privacy in IoT Using Machine Learning and Blockchain: Threats and Countermeasures. *ACM Comput. Surv.* **2020**, *53*, 1–37. [CrossRef]
37. Park, J.H. Advanced IT-Based Future Sustainable Computing (2017–2018). *Sustainability* **2019**, *11*, 2264. [CrossRef]
38. Cao, L.; Yang, Q.; Yu, P.S. Data science and AI in FinTech: An overview. *Int. J. Data Sci. Anal.* **2021**, *12*, 81–99. [CrossRef]
39. Al-Qarafi, A.; Alrowais, F.; Alotaibi, S.S.; Nemri, N.; Al-Wesabi, F.N.; Al Duhayyim, M.; Marzouk, R.; Othman, M.; Al-Shabi, M. Optimal Machine Learning Based Privacy Preserving Blockchain Assisted Internet of Things with Smart Cities Environment. *Appl. Sci.* **2022**, *12*, 5893. [CrossRef]
40. Raj, A.; Shetty, S.D. IoT Eco-system, Layered Architectures, Security and Advancing Technologies: A Comprehensive Survey. *Wirel. Pers. Commun.* **2022**, *122*, 1481–1517. [CrossRef]
41. Ferrag, M.A.; Maglaras, L. DeepCoin: A Novel Deep Learning and Blockchain-Based Energy Exchange Framework for Smart Grids. *IEEE Trans. Eng. Manag.* **2020**, *67*, 1285–1297. [CrossRef]
42. Zerka, F.; Urovi, V.; Vaidyanathan, A.; Barakat, S.; Leijenaar, R.T.H.; Walsh, S.; Gabrani-Juma, H.; Miraglio, B.; Woodruff, H.C.; Dumontier, M.; et al. Blockchain for Privacy Preserving and Trustworthy Distributed Machine Learning in Multicentric Medical Imaging (C-DistriM). *IEEE Access* **2020**, *8*, 183939–183951. [CrossRef]
43. Harbi, Y.; Aliouat, Z.; Refoufi, A.; Harous, S. Recent Security Trends in Internet of Things: A Comprehensive Survey. *IEEE Access* **2021**, *9*, 113292–113314. [CrossRef]
44. Diro, A.; Chilamkurti, N.; Nguyen, V.-D.; Heyne, W. A Comprehensive Study of Anomaly Detection Schemes in IoT Networks Using Machine Learning Algorithms. *Sensors* **2021**, *21*, 8320. [CrossRef] [PubMed]
45. Nouman, A.; Muneer, S. A Systematic Literature Review on Heart Disease Prediction Using Blockchain and Machine Learning Techniques. *Int. J. Comput. Innov. Sci.* **2022**, *1*, 1–6.
46. Li, Y.; Shan, B.; Li, B.; Liu, X.; Pu, Y. Literature Review on the Applications of Machine Learning and Blockchain Technology in Smart Healthcare Industry: A Bibliometric Analysis. *J. Healthc. Eng.* **2021**, *2021*, 9739219. [CrossRef]
47. Cheng, A.S.; Guan, Q.; Su, Y.; Zhou, P.; Zeng, Y. Integration of Machine Learning and Blockchain Technology in the Healthcare Field: A Literature Review and Implications for Cancer Care. *Asia-Pac. J. Oncol. Nurs.* **2021**, *8*, 720–724. [CrossRef]
48. Karger, E. Combining Blockchain and Artificial Intelligence—Literature Review and State of the Art. In Proceedings of the Forty-First International Conference on Information Systems, Hyderabad, India, 13–16 December 2020.
49. Ekramifard, A.; Amintoosi, H.; Seno, A.H.; Dehghantanha, A.; Parizi, R.M. A Systematic Literature Review of Integration of Blockchain and Artificial Intelligence. In *Blockchain Cybersecurity, Trust and Privacy. Advances in Information Security*; Choo, K.K., Dehghantanha, A., Parizi, R., Eds.; Springer: Cham, Switzerland, 2020; Volume 79, pp. 147–160. [CrossRef]
50. Azimov, D. Analysis of the international experience of implementing blockchain technology. Access to science, business, innovation in digital economy. *ACCESS Press* **2021**, *2*, 138–149. [CrossRef]
51. Geldiev, E.M.; Nenkov, N.V.; Petrova, M.M. Exercise of Machine Learning Using Some Python Tools and Techniques. In Proceedings of the CBU International Conference Proceedings 2018, Prague, Czech Republic, 21–23 March 2018.

52. Petrova, M.M.; Sushchenko, O.; Trunina, I.; Dekhtyar, N. Big Data Tools in Processing Information from Open Sources. In Proceedings of the 2018 IEEE First International Conference on System Analysis & Intelligent Computing (SAIC), Kyiv, Ukraine, 8–12 October 2018. [CrossRef]
53. Petrova, M.; Popova, P.; Popov, V.; Shishmanov, K.; Marinova, K. Potential of Big Data Analytics for Managing Value Creation. In Proceedings of the 2022 International Conference on Communications, Information, Electronic and Energy Systems (CIEES), Veliko Tarnovo, Bulgaria, 24–26 November 2022. [CrossRef]
54. Azimov, D.T.; Petrova, M. Determination of The Efficiency of Implementing Blockchain Technology into the Logistics Systems. *Bus. Manag.* **2022**, *4*, 52–67.