earth

Article

Blockchain Projects in Environmental Sector: Theoretical and Practical Analysis

Matteo Vaccargiu [1,2,*,†,‡] and Roberto Tonelli [1,‡]

[1] Department of Mathematics and Computer Science, University of Cagliari, 72 Via Ospedale, 09124 Cagliari, Italy; roberto.tonelli@unica.it
[2] Division of Computer Science, University of Camerino, 9 Via Madonna delle Carceri, 62032 Camerino, Italy
* Correspondence: matteo.vaccargiu@unica.it
† Current address: Palazzo delle Scienze, Via Ospedale 72, 09121 Cagliari, Italy.
‡ These authors contributed equally to this work.

Abstract: The growing interest in environmental sustainability issues and, at the same time, the advantages offered by blockchain technology have strong connections to each other. This study explores the application of blockchain technology across various environmental domains, such as air quality, climate change impacts, and resource management. The research utilised a dual approach, combining a bibliometric analysis with VOSviewer and a topic analysis using BERT models to assess the discourse within both the scientific literature extracted from Scopus and practical blockchain projects obtained from GitHub. The findings reveal that food security, energy, and sustainable agriculture are predominant topics in academic discussions, with a noticeable increase in focus from 2017 onwards. Practical projects are focused on transparent tracking and decentralised management. The overlap between academic and practical spheres is evident in the shared focus on energy and environmental management, demonstrating blockchain's growing role in addressing global environmental challenges. This study underscores the importance of integrating theoretical research with practical implementations to harness blockchain's full potential in promoting sustainable environmental practices.

Keywords: blockchain; Earth; energy; sustainability; bibliometric analysis; topic analysis

check for updates

Citation: Vaccargiu, M.; Tonelli, R. Blockchain Projects in Environmental Sector: Theoretical and Practical Analysis. *Earth* **2024**, *5*, 354–370. https://doi.org/10.3390/earth5030020

Academic Editor: Samuel Asumadu-Sarkodie

Received: 14 July 2024
Revised: 9 August 2024
Accepted: 12 August 2024
Published: 14 August 2024

1. Introduction

In recent years, blockchain technology and Distributed Ledger Technology (DLT) have emerged as significant tools capable of revolutionising various sectors by providing enhanced transparency, security, and efficiency [1]. These technologies are particularly relevant in environmental contexts, where they can significantly contribute to the monitoring, reporting, and verification of data in applications ranging from air quality management to biodiversity conservation. Their application directly supports the achievement of the Sustainable Development Goals (SDGs), particularly those related to environmental protection and sustainable resource management [2]. The urgency of addressing environmental challenges such as air pollution, climate change, water scarcity, and the sustainable management of natural resources is more pronounced than ever before.

Blockchain technology, defined as a sharing register that ensures the immutability and transparency of data transactions [3], offers a robust framework for tackling complex environmental issues. Its decentralised nature prevents tampering and ensures the integrity of environmental data [4]. This is crucial for tracking atmospheric changes, managing the environmental impacts of climate dynamics, and ensuring the sustainable use of water resources. Furthermore, blockchain can facilitate enhanced resource management strategies, aid in disaster risk reduction, and support sustainable urban development by enabling more effective coordination and management of environmental policies and practices [5].

This paper explores the application of blockchain technology across several environmental domains, including air quality and pollution control, climate change impacts, water quality management, and the interplay between land use and environmental health. Based on a bibliometric and topic modelling approach, we tried to answer the following research questions:

RQ1: What are the main discussion topics within the **scientific literature** regarding the use of blockchain in critical areas such as Earth sciences, climate change, and environmental health? How have these discussions evolved over time?
RQ2: What themes emerge from **practical blockchain projects** in these fields, and how do they develop over time?
RQ3: Is there an **overlap** between the themes explored in **academic research** and those implemented in **practical projects**?

To answer these questions, we performed a bibliometric analysis using VOSviewer and a topic analysis with a BERT (Bidirectional Encoder Representations from Transformers) model with the goal of mapping the current landscape of blockchain applications in environmental research and highlighting how this technology is being used to address pressing global challenges. By bridging the gap between theoretical research and practical applications, this paper seeks to inform and inspire stakeholders from various sectors, particularly those without an IT background, about the potential of blockchain technology in promoting environmentally sustainable behaviours. Finally, in order to support global sustainability initiatives, we would like to promote the wider usage and application of blockchain technology to address environmental problems.

This work aims to contribute to the study of the main topics of interest related to the use of blockchain technology in projects related to environmental issues from both theoretical and practical points of view, providing a methodology that can also be replicated in other studies so that researchers, developers, politicians, stakeholders, and all those interested in environment-related projects can take into consideration the contribution that blockchain technology can provide in this area.

The paper is organised as follows: In Section 2, we present the existing literature connected to this study; then, in Section 3, we discuss in detail the methodology followed to answer the research questions. The results are presented in Section 4 and discussed and validated in Sections 5 and 6, respectively. Finally, in Section 7, conclusions and future research developments in this area are discussed.

2. Related Works

The exploration of blockchain technology in environmental and sustainability domains has attracted the attention of the scientific community, focusing on its application across various critical areas, such as climate change, energy sustainability, and environmental management. This section outlines the key contributions from the recent literature that align with our methodological approach, such as bibliometric studies, mapping studies, literature reviews, topic analyses, etc., to analyse the scientific literature related to the use of blockchain in the environmental sector, practical projects on the same topics, or both. In this section, we also provide a comparison between these studies and our objectives and methods to highlight the novelty of our research.

Jin et al. [6] addressed the integration of blockchain into environmental management frameworks, demonstrating its potential through bibliometric analysis while noting the scarcity of practical implementations. Their work underlines the foundational stages of blockchain applications in this field, which aligns with the preliminary findings of this study through bibliometric mapping. They used VOSviewer for the analysis, as in our work, but they focused only on bibliometric studies.

O'Donovan et al. [7] conducted an extensive review of blockchain applications within the energy sector, emphasising the gap between theoretical research and practical applications. Their insights into real-world blockchain initiatives offer a critical perspective that complements the practical component of this study, where real-world blockchain projects from GitHub were analysed. Comparing it with our work, the methodology is similar, but they only focus on applications in the energy sector.

Joshi et al. [8] systematically reviewed the literature on blockchain's impact on sustainable development, linking it to the United Nations Sustainability Development Goals. This study extends their thematic analysis by using VOSviewer and topic modelling to examine how these themes are discussed in the recent scientific literature and practical projects. Their methodology is similar to ours, but their focus is more on the SDGs.

Popkova et al. [9] explored the conceptual and empirical applications of blockchain for climate change and clean energy, which are in accordance with the areas of interest of this research. The discussion of the role of blockchain in promoting green initiatives and sustainable investments provides a comparative basis for evaluating the results of this study's bibliometric analysis and review of GitHub projects.

Böckel et al. [10] examined the potential of blockchain to support circular economy approaches. Their analysis of the challenges and opportunities mirrors the dual analytical approach of this study, where both academic and practical perspectives are considered to assess blockchain's impact on environmental sustainability.

Furthermore, Gawusu et al. [11] and Wang et al. [12] provided insights into the integration of blockchain with renewable energy sources, noting significant research interest and practical developments in this area. These findings are critical, as they align with this study's focus on energy-related topics within the blockchain discourse.

Dorfleitner et al. [13] and Arshad et al. [14] contributed empirical and theoretical insights into blockchain applications that specifically target climate protection and sustainability goals. Their discussions on the operational success factors and the strategic implications for policymakers provide a valuable framework for the discussions in this paper.

Mulyono et al. [15] conducted a bibliometric analysis on agricultural infrastructure funding, highlighting an increase in research paralleled by advancements in blockchain for sustainable development. This aligns with our approach to exploring environmental sector funding mechanisms, particularly in agriculture. Similarly, El jaouhari et al. [16] examined the role of ICTs like blockchain in decarbonising the agriculture supply chain, emphasising how these technologies enhance food safety and sustainability. Both studies resonate with our findings, but our research extends these themes by integrating practical blockchain applications with theoretical insights to offer a comprehensive analysis of blockchain's impact across the environmental sector.

Argumedo-García et al. [17] conducted a bibliometric analysis of the application of technologies in humanitarian supply chains, focusing on the interaction between critical technologies and their real benefits in disaster relief operations. Their study highlights significant gaps in research related to human-made disasters and health emergencies, showing a path for future studies that integrate advanced technologies like blockchain and artificial intelligence. This aligns with our work in assessing the role of blockchain in environmental sustainability but expands our perspective by connecting technology with direct humanitarian impacts. Similar to our methodology, they used VOSviewer for bibliometric mapping, but their focus was on a more niche application within humanitarian contexts, providing a comparative insight into how blockchain could be similarly transformative in environmental applications.

Pesqueira et al. [18] explored the application of blockchain in combating fraud in healthcare packaging, emphasising blockchain's potential to enhance traceability and security. Their systematic review, using a bibliometric analysis with VOSviewer, focuses on healthcare, specifically on counterfeit medications. Comparing this to our work, while they concentrate on healthcare fraud, our study extends blockchain's applicability to broader environmental sustainability, employing similar bibliometric methodologies but across a wider array of topics.

El Ouarrak and Hmioui [19] analysed the impact of blockchain on enhancing supply chain resilience through a bibliometric and visualisation study, identifying key areas where blockchain technology supports secure supply chain management and resilience, especially in times of global disruptions like COVID-19. Their approach, utilising VOSviewer for analysis, aligns with our use of bibliometric tools but focuses more narrowly on supply chain resilience, while our study broadens the application to various environmental sustainability efforts.

Finally, González-Mendes et al. [20] explored the intersection of blockchain technology and sustainability, employing a bibliometric analysis to uncover how blockchain can enhance supply chain sustainability and support transparent, sustainable business practices. Their research, while employing methods of analysis with VOSviewer similar to ours, primarily focuses on the economic and operational impacts of blockchain in business settings, offering a perspective that complements our broader environmental focus.

In Table 1, we report a summary of the related works considered in this study. While the reviewed literature lays a robust foundation for understanding blockchain's role in sustainability and environmental management, this study contributes an integrated analysis of scientific articles and practical projects, providing a different approach and methodology that show the full spectrum of the potential of blockchain in these critical areas. Moreover, the majority of the reviewed articles are focused only on a specific area of environmental sustainability, whereas our work tries to cover all aspects of interest related to the use of blockchain in this sector. So, summarising, this paper seeks to bridge the identified gaps between theoretical advancements and practical implementations, providing a comprehensive overview of the current state and future potential of blockchain technologies in environmental sustainability.

Table 1. Summary of related works.

Topic	Related Works	Similarities	Novelty of Our Study
Environmental Management	[6]	The use of VOSviewer for bibliometric analysis	A topic analysis of the literature and the introduction of practical implementations of blockchain, enhancing theoretical models with real-world applications and sustainability impacts.
Sustainability	[8,20]	A bibliometric analysis of sustainability and blockchain with VOSviewer	We offer a framework that includes both SDG implications and practical blockchain applications; also, the topic modelling methodology is different.
Supply Chain	[19]	Blockchain in supply chain resilience	Applies blockchain to environmental sustainability, extending beyond commercial supply chains to include ecological impact assessments.
Healthcare	[18]	Focuses on blockchain for traceability in healthcare	Adapts blockchain solutions to environmental fraud prevention, showcasing the adaptability of blockchain across sectors.
Circular Economy	[10]	Considers both theoretical and practical approaches following the PRISMA guidelines	The general findings are not only focused on the economy, and the methodology followed is different.

Table 1. *Cont.*

Topic	Related Works	Similarities	Novelty of Our Study
Renewable Energy	[7,9,11,12]	Insights into blockchain with renewable energy sources, also considering GitHub projects (see [7,9])	It covers more general environmental applications and not only energy topics, and they use a manual analysis method in contrast we propose topic modelling.
Climate Protection	[9,13,14]	Blockchain applications targeting climate protection with also a small focus on GitHub practical projects (see [9])	Our topic extraction methodology is different, and we also cover more aspects of environmental sustainability.
Agriculture	[15,16]	Blockchain's role in agriculture for funding and sustainability	Our study merges theoretical insights with practical applications, providing new methodologies for assessing and implementing blockchain in agricultural sustainability.
Humanitarian Aid	[17]	Bibliometric and visualisation approach to blockchain analysis in humanitarian contexts using VOSviewer	We cover more aspects of sustainability, and we perform a more in-depth analysis.

3. Methodology

This work aims to provide an overview and statistics about topics related to theoretical research and practical applications of blockchain technology in environmental projects. In this section, we offer a detailed look at the dataset used and explain the method we employed to extract and analyse the topics.

3.1. Dataset Overview and Statistics

For the bibliometric mapping analysis and for the topic modelling of the scientific literature, we considered a dataset extracted from Scopus using the following research query: *("blockchain" OR "DLT") AND ("earth" OR "Air quality" OR "pollution" OR "Environmental impacts" OR "climate change" OR "Water quality" OR "Sustainable urban development" OR "Soil system" OR "Natural disasters" OR "human-made disasters").*

We selected Scopus as our primary database due to its comprehensive coverage of interdisciplinary research and its robust analytical tools, which are essential for a holistic and accurate mapping of current trends and developments in blockchain technology applied to environmental systems.

Based on the above research query, we obtained a set of 1262 documents published from 1995 until June 2024, of which 1238 were in English, 17 in Chinese, 5 in Japanese, 2 in Spanish, and 1 in Bosnian.

We considered only the English documents for the analysis. Then, we also excluded 112 proceedings' data, keeping only research papers for the analysis. After that, we excluded another six articles without an abstract. We obtained, at the end of the process, 1120 papers.

In Figure 1, we have the distribution of the articles over time. As we can observe from the plot, the trend of interest in these themes has been growing steadily and continuously since 2017, with a peak in 2023. The mean number of citations for these articles is 11.98, and the most cited article is titled "Internet of things (IoT) and the energy sector", cited 460 times.

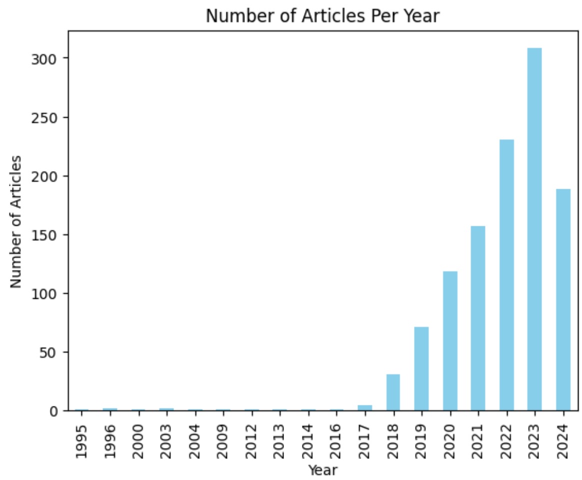


Figure 1. Number of articles published by year.

For the practical project analysis, we considered a dataset extracted from GitHub, which is a web-based platform that provides hosting for software development and version control. For these reasons, it is a good space to research blockchain projects related to the topics of sustainability, Earth, pollution, water, air, etc., and study how developers work on these issues. In particular, we extracted all repositories and all issues that were obtained with the string *"earth + blockchain"*. We thus obtained a dataset of 1000 issues and one of 59 repositories, 1 of which was eliminated because it was empty and lacked a description.

In Figure 2, we have the distribution of the opened issues over time. As we can observe from the plot, we have two peaks, one in 2019 and one in 2023. The mean number of citations for these articles is 11.98, and the most cited article is titled "Internet of things (IoT) and the energy sector", cited 460 times. Of these issues, 442 are still open, while 558 have been closed. The mean lifespan of issues is 135.31 days. The mean number of comments on an issue is 24.32, and the one most commented on has 1907 comments.

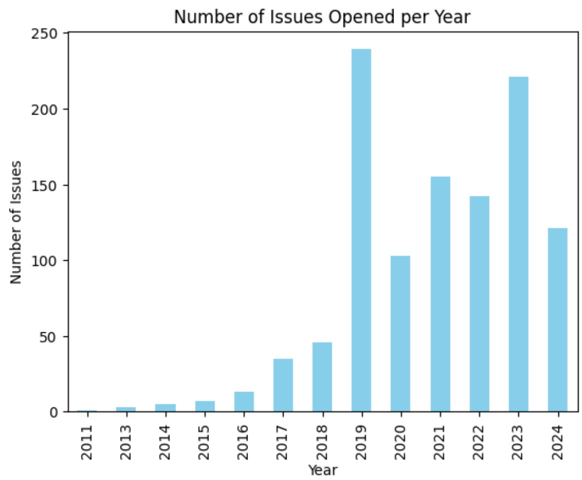


Figure 2. Number of issues opened by year.

In Figure 3, we have the distribution of the created repositories over time. The mean size of these repositories is 8894.78 KB, and the biggest is 156,808 KB. Moreover, we have nine different programming languages used for blockchain projects in the Earth field, and the top three most used are JavaScript (14 times), TypeScript (4 times), and Solidity (4 times). For this result, it is important to highlight that only 31 repositories declare the programming language used. Finally, the mean number of open issues per repository is 1.57, and the one that has 70 issues open is a content delivery network (CDN) that uses Ethereum and IPFS.

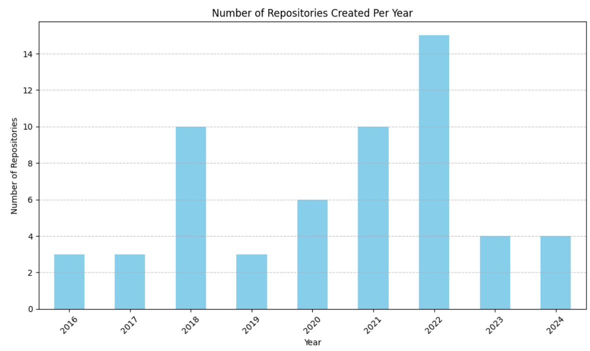
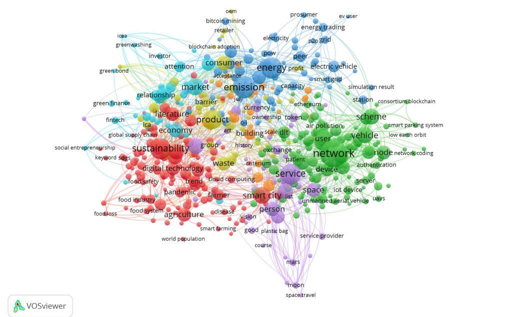


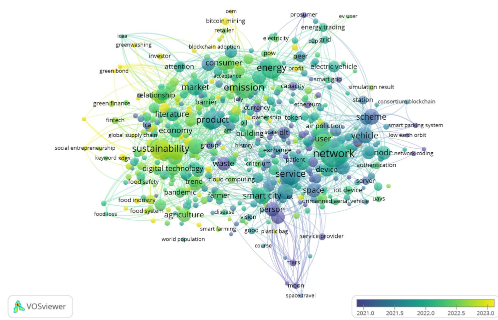
Figure 3. Number of repositories created by year.

3.2. Theoretical Project Analysis

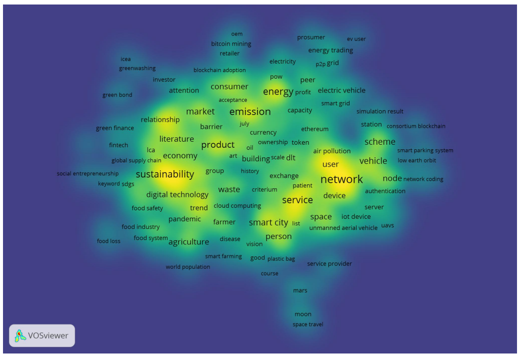
The primary descriptive analysis was a bibliometric map on the Scopus data created using VOSviewer, a software tool for constructing and visualising bibliometric networks [21]. Considering both the titles and abstracts of the paper, we performed both full and binary counting. Full counting helps capture all occurrences without reducing the weight of multiple contributions within the same item. It gives more weight to terms that appear frequently, which is helpful in the study of the influence or prevalence of certain topics. On the other hand, binary counting is useful to avoid the overrepresentation of a specific word. This is useful for analysing the breadth of topics in the dataset and minimising the influence of prolific terms. The results obtained are shown in Figures 4 and 5.



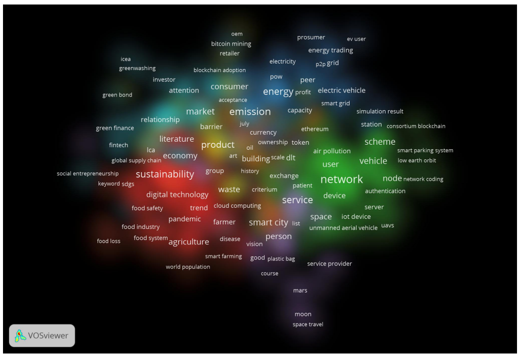
(**a**) Network visualisation with the full counting method analysis



(**b**) Overlay visualisation with the full counting method analysis

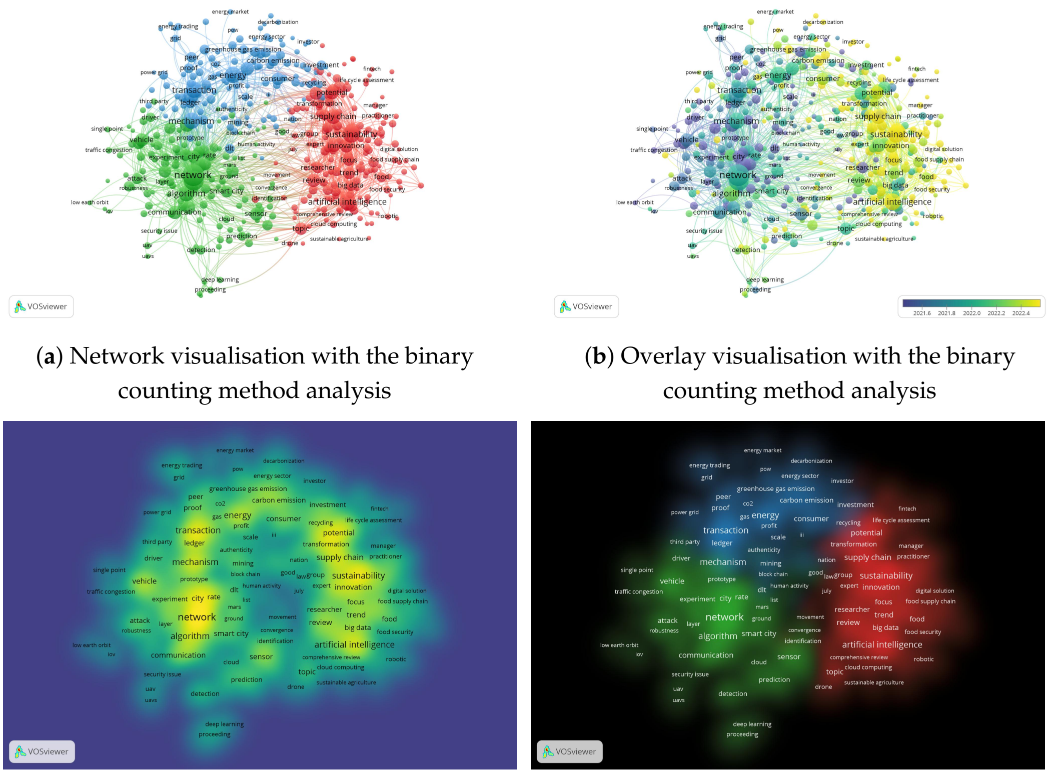


(**c**) Item density visualisation with the full counting method analysis



(**d**) Cluster density visualisation with the full counting method analysis

Figure 4. VOSviewer map analysis plots using the full counting method.



(**a**) Network visualisation with the binary counting method analysis

(**b**) Overlay visualisation with the binary counting method analysis

(**c**) Item density visualisation with the binary counting method analysis

(**d**) Cluster density visualisation with the binary counting method analysis

Figure 5. VOSviewer map analysis plots using the binary counting method.

Before analysing the research topics in more detail by applying a natural language processing (NLP) model, we preprocessed the data using Berteley's preprocess function, which is designed to systematically prepare textual data. It accepts a list of documents and performs a series of predefined cleaning and normalisation steps on the text. By incorporating these steps, the function ensures that the input text is standardised, potentially enhancing the performance of subsequent NLP tasks.

Then, once the data had been cleaned up, we performed a topic analysis using the BERT (Bidirectional Encoder Representations from Transformers) model [22], which is a pretrained transformer-based neural network model designed to understand the context of a given text by bidirectionally processing it. Specifically, we used *BERTopic* (https://maartengr.github.io/BERTopic/index.html (accessed on 1 August 2024)), a model that leverages contextual embeddings from BERT to identify and cluster topics within a collection of text documents. To encode the input text into a fixed-size hidden representation, BERT's architecture incorporates a multi-headed self-attention mechanism and a feed-forward neural network. To predict missing words in a given sentence, the BERT pretraining strategy involves training the model on a sizeable corpus of unannotated text. Specifically, in our case, we applied an unsupervised approach, considering two different embedding models. The first one, called *SciBERT*, is trained on a dataset of scientific articles [23], while the second one, called *ClimateBERT*, is trained on a dataset related to climate, sustainability, and the environment [24]. This choice is based on the fact that, by doing so, we obtain topics extracted from two different points of view. In addition, the choice of models was made based on the c_v score measure [25]. After the first process of topic extraction, we employed the feature "*reduce_outliers*" of BERTopic, which helped us reduce the unclassified documents, distributing them into clusters based on the class-based TF-IDF (c-TF-IDF).

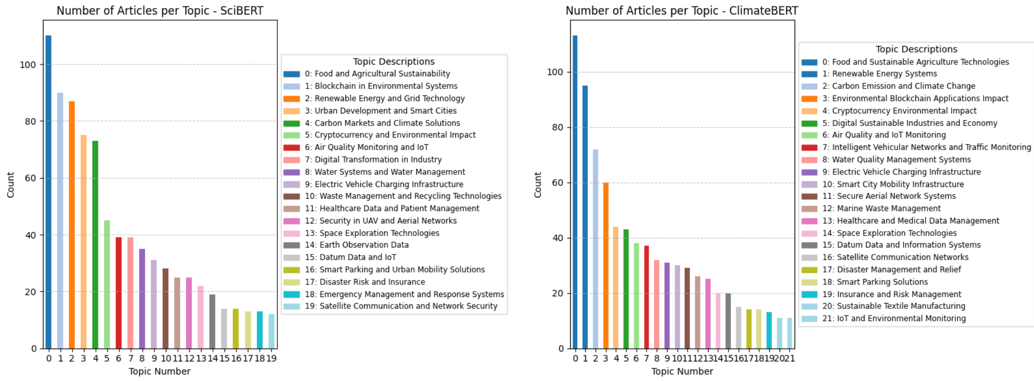
In Table 2, we have the number of topics and the c_v scores obtained with the two embedding models before and after the outlier reduction.

Table 2. Comparison of embedding models.

Embedding Model	Num Topics	c_v Score with Outliers	c_v Score with Outliers Reduced
SciBERT	20	0.6027	0.5802
ClimateBERT	22	0.5880	0.5811

Once we obtained the list of topics with the top 10 words associated with each one, we interpreted them for labelling the topics using Chat-GPT 4, followed by a manual check by the authors to validate the results. The effectiveness of the use of Chat-GPT for topic labelling was demonstrated by Colavito et al. [26].

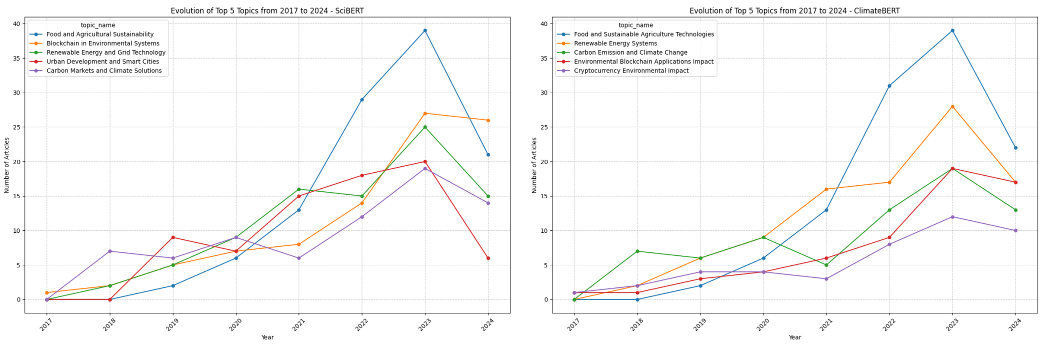
The results obtained are shown in Table 3 and in Figures 6–8.



(**a**) A barplot of the topics obtained using SciBERT as the embedding model

(**b**) A barplot of the topics obtained using ClimateBERT as the embedding model

Figure 6. A barplot of the topics obtained by the two topic extraction processes.



(**a**) A plot of the evolution over time of the top 5 topics obtained with sciBERT

(**b**) A plot of the evolution over time of the top 5 topics obtained with climateBERT

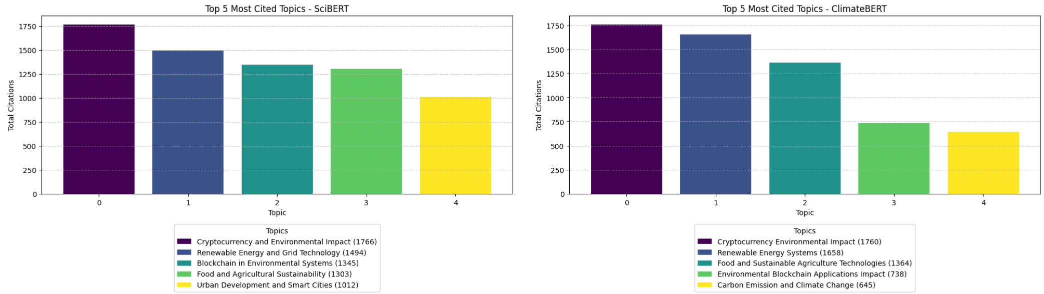
Figure 7. A plot of the evolution over time of the top 5 topics obtained by the two topic extraction processes.

Table 3. A comparison of topic extraction with SciBERT and ClimateBERT.

Topics Extracted with SciBERT		*Topics Extracted with ClimateBERT*	
Topic Name	**Count**	**Topic Name**	**Count**
Food and Agricultural Sustainability	116	Food and Sustainable Agriculture Technologies	119
Blockchain in Environmental Systems	171	Renewable Energy Systems	119

Table 3. *Cont.*

Topics Extracted with SciBERT		*Topics Extracted with ClimateBERT*	
Topic Name	**Count**	**Topic Name**	**Count**
Renewable Energy and Grid Technology	115	Carbon Emission and Climate Change	89
Urban Development and Smart Cities	100	Environmental Blockchain Applications Impact	140
Carbon Markets and Climate Solutions	87	Cryptocurrency Environmental Impact	52
Cryptocurrency and Environmental Impact	53	Digital Sustainable Industries and Economy	99
Air Quality Monitoring and IoT	47	Air Quality and IoT Monitoring	52
Digital Transformation in Industry	98	Intelligent Vehicular Networks and Traffic Monitoring	48
Water Systems and Water Management	42	Water Quality Management Systems	40
Electric Vehicle Charging Infrastructure	31	Electric Vehicle Charging Infrastructure	34
Waste Management and Recycling Technologies	33	Smart City Mobility Infrastructure	51
Healthcare Data and Patient Management	32	Secure Aerial Network Systems	40
Security in UAV and Aerial Networks	42	Marine Waste Management	31
Space Exploration Technologies	28	Healthcare and Medical Data Management	29
Earth Observation Data	29	Space Exploration Technologies	28
Datum Data and IoT	34	Datum Data and Information Systems	37
Smart Parking and Urban Mobility Solutions	14	Satellite Communication Networks	19
Disaster Risk and Insurance	13	Disaster Management and Relief	19
Emergency Management and Response Systems	17	Smart Parking Solutions	14
Satellite Communication and Network Security	18	Insurance and Risk Management	13
		Sustainable Textile Manufacturing	16
		IoT and Environmental Monitoring	31



(**a**) Top 5 of most cited topics obtained with sciBERT

(**b**) Top 5 most cited topics obtained with climateBERT

Figure 8. Top 5 most cited topics obtained by the two topic extraction processes.

3.3. Practical Project Analysis

Based on the issue and repository data extracted from GitHub, we performed a topic analysis based on the title and body of each issue, and we analysed the description and topic for each repository. Based on the idea of Vaccargiu et al. [27], for the analysis of the issues, we applied *BERTopic* using the bge embedding model; specifically, in this case, we considered the one called *BAAI/bge-reranker-base* because it is trained both in English and in Chinese, and some issues in our dataset are written in Chinese. At the end of the process, we obtained 20 topics and a c_v score measure of 0.6174. As carried out previously for scientific articles, wanting to classify all issues, we reduced the outliers by distributing them into the other classes, thus obtaining 19 topics and a c_v score measure of 0.6176. The results obtained are shown in Table 4.

Table 4. Issue topics extracted with BAAI/bge-reranker-bas.

Issue Topic Name	Count
GitHub Repository Management	209
Arxiv Paper Discussions	85
Academic Research and Documentation	84
Scholarly Communication	78
Coding and Development	71
Version Control and Collaboration	56
Software Tools and Configurations	51
Social Media and Personalities	48
Digital Assets and Web Content	43
Development and Issue Tracking	36
Software Documentation	37
Social Media Analysis	35
Web Development	28
Research Publications	27
Technical Configuration	27
Community and Resources	24
Scholarly Research	24
Machine Learning and Research	18
Repository Contributions	19

Then, we moved on to the repository data, and we performed another topic analysis. Specifically, in this case, having not obtained satisfactory results from the direct application of BERTopic to the description of repositories, probably because the lengths of many descriptions were too short, we extracted keywords from them using *KeyBERT*, which is a BERT-based package for keyword extraction. By utilising these models, KeyBERT can efficiently identify and extract the most relevant and contextually significant keywords and keyphrases from a body of text. Specifically, in our case, we extracted keywords from the repository descriptions and, combined with those provided by some developers, interpreted them by using Chat-GPT 4 and human checking to understand the topic of each repository. Once we obtained the topics, using a similar approach, we clustered them based on the topics covered. The results obtained are shown in Table 5.

Table 5. Clusters of repository topics.

Cluster	Topics
Blockchain Technology Integration	Polkadot Earth Networks, Ethereum for Earth in Asia, Blockchain Planet Projects, Blockchain Messaging Protocols, Blockchain for Earth Preservation, Blockchain Earth Projects, Blockchain Document Management, MES Protocol Ethereum Solana, Blockchain Metaverse Tutorials, Geospatial Blockchain Applications, Decentralized AI Blockchain Token, ERC20 Ethereum Blockchain Token, Blockchain Security and Hacking, Metaverse and Blockchain Essence, IoT and Blockchain Integration, Decentralized Insurance Platform, Global Blockchain Networks, IoT Blockchain Simulations, Mars Currency Blockchain Exchange, Pinball Protocol Blockchain Exchange, Decentralized Blockchain Ledger Oracles
Environmental and Sustainable Projects	Earth-Focused Content Media, Earth-Centric Bitcoin Projects, Hyperledger Earth Projects, Blockchain EarthDAO Ownership, Blockchain for Functional Earths, Blockchain Agriculture Applications, Real Estate Blockchain Fundraising, Blockchain-Powered Christmas Tree, Global Earthcoin Blockchain Village, Smart Security Blockchain Technology, Blockchain Agricultural Sustainability, Environmental Projects and Investments
Cryptocurrency and Financial Transactions	Earth-Centric Bitcoin Projects, Crypto and Ethereum Exchange, Mars Cryptocurrency Blockchain, Kaseicoin Cryptocurrency Platform
Media and Content Creation	Earth-Focused Content Media, Angular Blockchain Explorer, Blockchain Earthcam Image Encryption
Digital Governance and Smart Contracts	Digital DAOs and Jurisdiction, DAO Proposal Management, Blockchain in Government Trust, Blockchain Investment Funds Ledger
Community and Social Impact	Moralis Web3 Metaverse, NFT Assets Blockchain Management, Blockchain and Community Democracy, NFTs Teenagers Platform, Fractal Databases on Facebook
NFT and Digital Assets	Blockchain NFT Earth Projects, NFT Assets Blockchain Management, DAO Proposal Management, NFTs Teenagers Platform

4. Results

This section presents the results of the bibliometric analysis and the interpretation of the resulting topics obtained in the different NLP processes.

Performing the analysis with VOSviewer, we set the minimum number of occurrences of a term to 10 and considered only the top 60% most relevant term. Using full counting, we selected 532 words, and when using binary counting, we considered 512 items. With the first method, we obtained 7 clusters, 40,563 links, and a total link strength of 183,361, and the results are shown in Figure 4. In contrast, for the second one, we have 3 clusters, 37,357 links, and a total link strength of 83,631. The results obtained are plotted in Figure 5.

We performed a topic analysis with BERT, and we report the results obtained in Table 3. Using SciBERT as an embedding model, we obtained 20 topics; in contrast, using ClimateBERT, we obtained 22 topics. The barplot of topics obtained by the two methods is shown in Figure 6. The plot of the evolution over time of the five topics most commonly found in the literature is shown in Figure 7. For this plot, we considered the time interval from 2017 to 2024 because, as shown in Figure 8, this is the period with the most scientific publications on these topics. Finally, the plot of the top five most cited topics is shown in Figure 1. For all these three graphs, the results obtained by applying the two different embedding models were compared.

Moving on to the analysis of practical projects extracted from GitHub, in Table 4, we report the list of topics obtained from the issue analysis by applying the BERTopic model with the BAAI/bge-reranker-bas embedding model. In contrast, in Table 5, the topics extracted from the repositories and their clasterisation are reported.

5. Discussion

The interest in blockchain applications in Earth science, climate, energy, health, etc., has experienced significant growth in recent years, as confirmed by the number of articles in Figure 1. The fact that many of these articles have few citations is due to the fact that many of them are recent, as evidenced by the graph. Nevertheless, the fact that an article related to the energy sector has been so successful testifies to the many benefits provided by blockchain for energy production and buying and selling. One of the main problems found in blockchain application fields is that projects often remain theoretical ideas. Despite this, in this case, the presence of 59 repositories and numerous issues and comments testify to how even developers are working to bring these theoretical ideas to life.

The bibliometric analysis provides us with the first overview of the main topics of discussion in the literature. From the full counting analysis, we can observe in Figure 4a that the prevalent topics of interest are related to sustainability, smart city, service, network, energy, emission, waste, product, and market. This is also confirmed by Figure 4c,d, which also show the breakdown into clusters, such as agribusiness, green finance and market, waste products, service providers, air pollution and IoT networks, and finally, p2p energy trading and electric vehicles. Finally, from Figure 4b, it can be seen that topics such as sustainability overlap across multiple clusters, while others, such as waste, are more specific to their group. On the other hand, with the binary counting analysis, we can analyse the breadth of topics in the dataset and minimise the influence of common terms. Figure 5a shows that we have three big clusters: one related to network, smart city, vehicles, and traffic; another one related to sustainability, innovation, and food and agriculture supply chain; and finally, the last one related to energy, carbon emissions, greenhouse gases, and the p2p energy trading market. This result can also be observed in Figure 5c,d. Also, in this analysis, as can be seen in Figure 5b, topics such as sustainability, supply chain, artificial intelligence, and carbon emissions cross-cut across several clusters; in contrast, others, such as peer, grid, and driver, are more class-specific.

Going into more detail about the topics of interest in scientific articles, in Table 3, we can observe the topic modelling results obtained by applying the two BERT models. The distribution of these is also plotted in Figure 6, providing an even clearer idea of the main interests of the scientific community. Among the topics obtained with the SciBERT model, it is noticeable that the most discussed topics concern general topics related to the environment (Blockchain in Environmental Systems), agriculture and food (Food and Agricultural Sustainability), energy (Renewable Energy and Grid Technology), and urban development (Urban Development and Smart Cities). Other smaller topics discuss recycling, emissions, air quality, space and planets, and, last but not least and no less interesting, aspects of cryptocurrency and mining. These results are also confirmed by the topics obtained with the ClimateBERT model, which, in addition, extrapolates on topics such as "Sustainable Textile Manufacturing", which is very important nowadays due to the emissions caused by the mass production of clothes and the difficulty of recycling them once they are discarded. The evolution of these main topics in recent years, specifically the top five, can be seen in Figure 7. It is interesting to note the steady growth in topics related to agriculture and food, underscoring how even sectors historically characterised by manual labour are facing technological innovations. Also interesting in Figure 8a is the growth in themes related to the environment in general over the last year. This highlights how the cross-cutting use of blockchain for environmental monitoring and prevention is increasingly impactful. One difference between the two models, as noted in Figure 8b, is related to the cryptocurrency topic, which is not present in the top five of the SciBERT model. While the latter is the most extreme among the top topics, it is one related to urban development, which is not highlighted by ClimateBERT. This result is interesting in that one would expect more technical topics (cryptocurrencies) to be closer to a model trained on scientific data, while urban development topics are more related to an "environmental" model such as ClimateBERT. Evidence of how cryptocurrencies are impacting the environment, an aspect of interest to researchers, is shown in Figure 8, where these topics are the most frequently mentioned. In both cases, this is followed by topics related to energy, a sector that offers great application possibilities, such as renewable energy certification, energy trading, and energy management. Finally, the other most cited topics concern the environment, agriculture, food, and finally, climate change.

Now, discussing the results of practical projects on GitHub, let us first observe that Figures 2 and 3 are connected. In fact, we can see that in 2018 and 2022, there were peaks of open repositories, and in 2019 and 2023, there were increases in open issues, probably related to work on new projects by developers.

The topics extracted from the issues and reported in Table 4 show academic and technical topics. Many topics are related to scientific research, Arxiv, etc., highlighting how many scientific researchers probably propose preprints of their solutions to the community to receive feedback and make improvements. Other technical topics concern software tools, digital assets and web content, GitHub repository management, and social media analysis, highlighting the aspects that blockchain developers in the Earth sector most discuss and collaborate on. The same results are also confirmed by the analysis of the repositories in Table 5, which probably explains, in even more detail, some technical aspects, such as the use of particular blockchains, like Ethereum, Hyperledger, or Polkadot, and tokens like ERC-20, NFT, DAO, Crypto, and Bitcoin applications. Despite this, environmental and sustainable project applications are also presented, such as Earth-focused projects, blockchain agriculture applications, and real estate blockchain fundraising.

To summarise our results, the first question, **RQ1**, asked the following: **What are the main discussion topics within the scientific literature regarding the use of blockchain in critical areas such as Earth sciences, climate change, and environmental health? How have these discussions evolved over time?** Our findings reveal that *food, agriculture, energy, cryptocurrency, carbon emissions, and waste* are the areas of greatest applicability of blockchain, showing steady growth from 2017 to the current time. In these fields, blockchain technology improves *data transparency, security, and collaboration* between the various stakeholders.

The second question, **RQ2**, asked the following: **What themes emerge from practical blockchain projects in these fields, and how do they develop over time?** The analysis shows that *initial themes revolved around experimental and pilot projects aimed at testing the feasibility of blockchain technologies in real-world contexts, such as the transparent tracking of carbon offsets, supply chain management for sustainable resources, and decentralised energy trading platforms*. Another aspect of interest is the use of GitHub to *propose new research works, receive feedback, and make improvements*. The evolution of repositories and issues has *two peaks, one between 2018 and 2019 and one between 2022 and 2023*, demonstrating a growth in interest in recent years also linked to technological innovations.

Finally, the last question, **RQ3**, asked the following: **Is there an overlap between the themes explored in academic research and those implemented in practical projects?** Considering both analyses, we observe that topics related to *energy, agriculture, and environmental management* are widely discussed in the literature and have practical projects implemented on GitHub. Another aspect in common is the *spike in growth in recent years* towards the use of blockchain in Earth sciences, climate change, and environmental health.

6. Threats to Validity

In this study, we recognise a range of potential threats to the validity of the findings, covering external, internal, and construct validity concerns:

- *Sampling Bias:* Our analysis primarily draws from data collected from the Scopus database and GitHub. This sample may not reflect the full diversity of blockchain's scientific articles and applications in environmental management. This limitation could influence the extent to which our results can be generalised.
- *Technological Evolution:* The swift advancement in blockchain technology may render our findings less relevant as new platforms or methodologies emerge that were not included in our initial analysis.
- *Analytical Limitations:* The use of bibliometric analysis and BERTopic for topic modelling introduces inherent biases. These methodologies might impose constraints on the data that could overlook subtle or emerging themes. The interpretation of the topics through Chat-GPT 4 and the subsequent check by the authors may also introduce a potential bias due to their background and expertise.

7. Conclusions and Future Works

This study highlights the potential of blockchain technology when aligned with environmental sustainability efforts. The results obtained, therefore, show the importance of combining theoretical ideas with practical implementations. Using platforms such as GitHub helps researchers receive feedback on their proposals and make improvements to the draft projects they implement.

The topics of energy, climate, sustainable mobility, smart cities, and food and sustainable agriculture are becoming increasingly important nowadays. Blockchain fits perfectly with these applications, providing secure, transparent, and immutable records for transactions and data management.

The first research question revealed that topics like food security, sustainable agriculture, and energy management are consistently highlighted within the scientific literature on blockchain. This shows a steady focus on these areas since 2017, improving transparency, security, and collaboration among stakeholders.

The second research question showed that practical blockchain projects initially focused on small-scale trials to test the viability of applications in real-world settings. Projects often address carbon trading, supply chain improvements, and energy solutions, with notable development peaks in 2018–2019 and again in 2022–2023.

Finally, the last research question confirmed an overlap between academic research and practical implementations, especially in the energy and agricultural sectors. This reflects a robust link between theoretical discussions and real-world applications, particularly with increased activity in recent years.

It follows that it is essential to enhance understanding and knowledge among technicians, researchers, and policymakers about the benefits and applications of blockchain technology in environmental sectors. Educating these key stakeholders on how blockchain can be effectively utilised in areas such as energy sustainability, waste management, and conservation efforts is crucial for fostering widespread adoption and maximising the technology's positive impact on environmental practices.

Future research should consider other academic databases, such as Web of Science and IEEE Xplore, and explore the regulatory, ethical, and security challenges associated with deploying blockchain technology in sensitive environmental and social contexts. This includes a detailed examination of global regulatory frameworks, the ethical implications of data usage and privacy in blockchain systems, and the robustness of these technologies against evolving security threats. Additionally, integrating interdisciplinary approaches that combine environmental science, information technology, and policy studies will provide a more comprehensive understanding of blockchain's potential to address complex environmental challenges.

Author Contributions: Conceptualization, M.V. and R.T.; methodology, M.V.; software, M.V.; validation, M.V. and R.T.; formal analysis, M.V.; investigation, M.V.; resources, M.V.; data curation, M.V.; writing—original draft preparation, M.V.; writing—review and editing, M.V.; visualization, M.V.; supervision, R.T.; project administration, R.T.; funding acquisition, R.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by Ministero dell'Università e della Ricerca (MUR), issue D.M. 351/2022 "Borse di Dottorato"—Dottorato di Ricerca di Interesse Nazionale in "Blockchain e Distributed Ledger Technology", under the National Recovery and Resilience Plan (NRRP). By the project HALO (Hazard-Analysis and Anti-Counterfeiting Ledger Oriented Protection) CUP F23C23000310008 in the context of Development Regional Program 2020–2024 Strategy 2-Economics Identity Project 2.1—Technological Research and Innovation by the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.5-Call for tender No. 3277 published on 30 December 2021 by the Italian Ministry of University and Research (MUR) funded by the European Union NextGenerationEU. Project Code ECS0000038—Project Title eINS Ecosystem of Innovation for Next Generation Sardinia—CUP F53C22000430001-Grant Assignment Decree No. 1056 adopted on 23 June 2022 by the Italian Ministry of University and Research (MUR). By the project "Network 4 Energy Sustainable Transition—NEST", code PE0000021, Concession Decree No. 1561 of 11.10.2022 adopted by Ministero dell'Università e della Ricerca (MUR).

Data Availability Statement: The data used can be downloaded free of charge from the Scopus site and GitHub using the same search queries used in this work.

Acknowledgments: This work was partially funded by Ministero dell'Università e della Ricerca (MUR), issue D.M. 351/2022 "Borse di Dottorato"—Dottorato di Ricerca di Interesse Nazionale in "Blockchain e Distributed Ledger Technology", under the National Recovery and Resilience Plan (NRRP). We acknowledge financial support under the project HALO (Hazard-Analysis and Anti-Counterfeiting Ledger Oriented Protection) CUP F23C23000310008 in the context of Development Regional Program 2020–2024 Strategy 2-Economics Identity Project 2.1—Technological Research and Innovation. We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.5-Call for tender No. 3277 published on 30 December 2021 by the Italian Ministry of University and Research (MUR) funded by the European Union NextGenerationEU. Project Code ECS0000038—Project Title eINS Ecosystem of Innovation for Next Generation Sardinia—CUP F53C22000430001-Grant Assignment Decree No. 1056 adopted on 23 June 2022 by the Italian Ministry of University and Research (MUR). We acknowledge financial support under the project code PE0000021, Concession Decree No. 1561 of 11.10.2022 adopted by Ministero dell'Università e della Ricerca (MUR), CUP—to be indicated by each Beneficiary, according to attachment E of Decree No. 1561/2022, Project title "Network 4 Energy Sustainable Transition—NEST".

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

References

1. Bhutta, M.N.M.; Khwaja, A.A.; Nadeem, A.; Ahmad, H.F.; Khan, M.K.; Hanif, M.A.; Song, H.; Alshamari, M.; Cao, Y. A Survey on Blockchain Technology: Evolution, Architecture and Security. *IEEE Access* **2021**, *9*, 61048–61073. [CrossRef]
2. Parmentola, A.; Petrillo, A.; Tutore, I.; De Felice, F. Is blockchain able to enhance environmental sustainability? A systematic review and research agenda from the perspective of Sustainable Development Goals (SDGs). *Bus. Strategy Environ.* **2022**, *31*, 194–217. [CrossRef]
3. Pilkington, M. Blockchain Technology: Principles and Applications. In *Research Handbook on Digital Transformations*; Edward Elgar Publishing: Cheltenham, UK, 2016; pp. 225–253.
4. Zhang, A.; Zhong, R.Y.; Farooque, M.; Kang, K.; Venkatesh, V.G. Blockchain-based life cycle assessment: An implementation framework and system architecture. *Resour. Conserv. Recycl.* **2020**, *152*, 104512. [CrossRef]
5. Wang, Y.; Chen, H. Blockchain: A potential technology to improve the performance of collaborative emergency management with multi-agent participation. *Int. J. Disaster Risk Reduct.* **2022**, *72*, 102867. [CrossRef]

6. Jin, S.; Chang, H. The trends of blockchain in environmental management research: A bibliometric analysis. *Environ. Sci. Pollut. Res.* **2023**, *30*, 81707–81724. [CrossRef]
7. O'Donovan, P.; O'Sullivan, D.T.J. A Systematic Analysis of Real-World Energy Blockchain Initiatives. *Future Internet* **2019**, *11*, 174. [CrossRef]
8. Joshi, P.; Tewari, V.; Kumar, S.; Singh, A. Blockchain technology for sustainable development: A systematic literature review. *J. Glob. Oper. Strateg. Sourc.* **2023**, *16*, 683–717. [CrossRef]
9. Popkova, E.G.; Bogoviz, A.V.; Lobova, S.V.; Vovchenko, N.G.; Sergi, B.S. Blockchain, sustainability and clean energy transition. *Glob. Transitions* **2023**, *5*, 64–78. [CrossRef]
10. Böckel, A.; Nuzum, A.K.; Weissbrod, I. Blockchain for the Circular Economy: Analysis of the Research-Practice Gap. *Sustain. Prod. Consum.* **2021**, *25*, 525–539. [CrossRef]
11. Gawusu, S.; Zhang, X.; Ahmed, A.; Jamatutu, S.A.; Miensah, E.D.; Amadu, A.A.; Osei, F.A.J. Renewable energy sources from the perspective of blockchain integration: From theory to application. *Sustain. Energy Technol. Assessments* **2022**, *52*, 102108. [CrossRef]
12. Wang, Q.; Su, M. Integrating blockchain technology into the energy sector—From theory of blockchain to research and application of energy blockchain. *Comput. Sci. Rev.* **2020**, *37*, 100275. [CrossRef]
13. Dorfleitner, G.; Muck, F.; Scheckenbach, I. Blockchain applications for climate protection: A global empirical investigation. *Renew. Sustain. Energy Rev.* **2021**, *149*, 111378. [CrossRef]
14. Arshad, A.; Shahzad, F.; Ur Rehman, I.; Sergi, B.S. A systematic literature review of blockchain technology and environmental sustainability: Status quo and future research. *Int. Rev. Econ. Financ.* **2023**, *88*, 1602–1622. [CrossRef]
15. Mulyono, M.; Latief, Y.; Rarasati, A. Agricultural Infrastructure Funding Assessment: Research Opportunities and Trends. *J. Optim. Ind. Eng.* **2024**, *17*, 203–212. [CrossRef]
16. El jaouhari, A.; Arif, J.; Samadhiya, A.; Naz, F.; Kumar, A. Exploring the application of ICTs in decarbonizing the agriculture supply chain: A literature review and research agenda. *Heliyon* **2024**, *10*, e29564. [CrossRef] [PubMed]
17. Argumedo-García, M.; Salas-Navarro, K.; Acevedo-Chedid, J.; Ospina-Mateus, H. Bibliometric Analysis of the Potential of Technologies in the Humanitarian Supply Chain. *J. Open Innov. Technol. Mark. Complex.* **2021**, *7*, 232. [CrossRef]
18. Pesqueira, A.; Sousa, M.J.; de Bem Machado, A. Addressing Counterfeiting and Fraud Concerns in Healthcare Packaging and Labeling with Blockchain: Opportunities and Challenges. *WSEAS Trans. Inf. Sci. Appl.* **2024**, *21*, 246–263. [CrossRef]
19. El Ouarrak, Y.; Hmioui, A. Blockchain technology and supply chain resilience: A bibliometric and visualization analysis. *Logforum* **2024**, *20*, 117–132. [CrossRef]
20. González-Mendes, S.; García-Muiña, F.; Alonso-Muñoz, S.; Costa, C.J.; González-Sánchez, R. *The Relationship between Blockchain Technology and Sustainability a Bibliometric Analysis*; Taylor & Francis: Milton Park, UK, 2024; pp. 235–247. [CrossRef]
21. Kirby, A. Exploratory Bibliometrics: Using VOSviewer as a Preliminary Research Tool. *Publications* **2023**, *11*, 10. [CrossRef]
22. Grootendorst, M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv* **2022**, arXiv:2203.05794.
23. Beltagy, I.; Lo, K.; Cohan, A. SciBERT: A Pretrained Language Model for Scientific Text. *arXiv* **2019**, arXiv:1903.10676.
24. Webersinke, N.; Kraus, M.; Bingler, J.A.; Leippold, M. ClimateBert: A Pretrained Language Model for Climate-Related Text. *arXiv* **2022**, arXiv:2110.12010.
25. Röder, M.; Both, A.; Hinneburg, A. Exploring the Space of Topic Coherence Measures. In Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, New York, NY, USA, 2–6 February 2015; pp. 399–408.
26. Colavito, G.; Lanubile, F.; Novielli, N.; Quaranta, L. Leveraging GPT-like LLMs to Automate Issue Labeling. In Proceedings of the 2024 IEEE/ACM 21st International Conference on Mining Software Repositories (MSR), Lisbon, Portugal, 15–16 April 2024.
27. Vaccargiu, M.; Aufiero, S.; Bartolucci, S.; Neykova, R.; Tonelli, R.; Destefanis, G. Sustainability in Blockchain Development: A BERT-Based Analysis of Ethereum Developer Discussions. In Proceedings of the 28th International Conference on Evaluation and Assessment in Software Engineering, Salerno, Italy, 18–21 June 2024; pp. 381–386.