minerals

MDPI

Article

A Review on Global Cooperation Network in the Interdisciplinary Research of Geochemistry Combined with Artificial Intelligence

Qianlong Zhang [1,2,3], Yongzhang Zhou [1,2,3,*], Juxiang He [1,2,3], Biaobiao Zhu [1,2,3], Feng Han [1,2,3] and Shiyao Long [1,2,3]

[1] School of Earth Sciences & Engineering, Sun Yat-sen University, Zhuhai 519000, China; zhangqlong3@mail2.sysu.edu.cn (Q.Z.); hejx65@mail2.sysu.edu.cn (J.H.); zhubb8@mail2.sysu.edu.cn (B.Z.); hfhanfeng@hotmail.com (F.H.); longshy6@mail2.sysu.edu.cn (S.L.)
[2] Guangdong Provincial Key Lab of Geological Process and Mineral Resources Survey, Sun Yat-sen University, Zhuhai 519000, China
[3] Center for Earth Environment & Resources, Sun Yat-sen University, Zhuhai 519000, China
* Correspondence: zhouyz@mail.sysu.edu.cn

Abstract: With the rapid development of modern geochemical analysis techniques, massive volumes of data are being generated from various sources and forms, and geochemical data acquisition and analysis have become important tools for studying geochemical processes and environmental changes. However, geochemical data have high-dimensional, nonlinear characteristics, and traditional geochemical data analysis methods have struggled to meet the demands of modern science. Nowadays, the development of big data and artificial intelligence technologies has provided new ideas and methods for geochemical data analysis. However, geochemical research involves numerous fields such as petrology, ore deposit, mineralogy, and others, each with its specific research methods and objectives, making it difficult to strike a balance between depth and breadth of investigation. Additionally, due to limitations in data sources and collection methods, existing studies often focus on a specific discipline or issue, lacking a comprehensive understanding of the bigger picture and foresight for the future. To assist geochemists in identifying research hotspots in the field and exploring solutions to the aforementioned issues, this article comprehensively reviews related studies in recent years, elaborates on the necessity and challenges of combining geochemistry and artificial intelligence, and analyzes the characteristics and research hotspots of the global collaboration network in this field. The study reveals that the investigation into artificial intelligence techniques to address geochemical issues is progressing swiftly. Joint research papers serve as the primary means of contact within a worldwide collaborative network. The primary areas of focus in the ongoing research on the integration of geochemistry and artificial intelligence include methodologies for analyzing geochemical data, environmental modifications, and mineral prospectivity mapping. Geochemical data analysis is currently a significant focus of research, encompassing a range of methods including machine learning and deep learning. Predicting mineral resources for deep space, deep Earth, and deep sea is also a pressing topic in contemporary research. This paper explores the factors driving research interest and future trends, identifies current research challenges, and considers opportunities for future research.

Keywords: geochemistry; big data; artificial intelligence; knowledge graph; analysis of cooperation network; CiteSpace

check for updates

Citation: Zhang, Q.; Zhou, Y.; He, J.; Zhu, B.; Han, F.; Long, S. A Review on Global Cooperation Network in the Interdisciplinary Research of Geochemistry Combined with Artificial Intelligence. *Minerals* **2023**, *13*, 1332. https://doi.org/10.3390/min13101332

Academic Editor: Paul Alexandre

Received: 27 September 2023
Revised: 9 October 2023
Accepted: 13 October 2023
Published: 15 October 2023

1. Introduction

Geochemistry is a field that examines the chemical makeup and transformations occurring in the Earth's interior and surface. It has a crucial function in addressing numerous fundamental Earth science challenges. Geochemical data analysis holds immense significance in comprehending the Earth system, identifying mineral reserves, monitoring the environment, and predicting earthquakes and volcanic activity, amongst others, thereby offering a data-supported foundation for Earth's scientific research and associated applications. Since the 1990s, scientists have endeavored to combine geochemical data with computer technology, augmenting the efficiency and precision of geochemical data analysis. Turing Award winner Jim Gary postulates that scientific inquiry has traversed four paradigms: the empirical paradigm, the theoretical paradigm, the computational paradigm, and the data-driven paradigm. Over the past decade, scientific research has shifted from being problem-driven to data-driven, leading to the emergence of the fourth paradigm of scientific discovery: data-intensive scientific discovery [1]. In recent times, the utilization of big data and artificial intelligence algorithms has transformed geology [2–4]. The integration of big data and artificial intelligence techniques with geochemistry has become a prominent research topic in the geochemistry field due to technological advances. This research includes an array of data types, such as geochemical and literature data, and encompasses various fields, including Earth science, environmental science, and information science. Through the application of big data analytics and methods, scientists have achieved significant advances in geochemical data analysis research. One such achievement includes the accurate classification of lithology [5], as well as distinguishing the rock's tectonic environment [6–8], mineral classification [9–11], the genesis of ore deposits [12–14], and mineral prospectivity mapping [15,16].

In previous studies, researchers obtained a large volume of geochemical data on deposits, minerals, rocks, soil, etc., through sampling analysis. Currently, there are various methods for the analysis of geochemical data to identify the frequency and spatial characteristics of these data, including classical statistics such as mean $\pm$ k [17], probability plot [18], exploratory data analysis [19], multivariate statistics [20], geostatistics [21], and fractal and multifractal models [22–24]. However, these data possess high-dimensionality and nonlinear characteristics, and traditional methods have limited ability to handle large volumes of multidimensional data. The information contained in multidimensional space has largely remained unexplored, posing a challenge for geochemical data analysis. With the development of big data and artificial intelligence technologies, machine learning algorithms have increasingly been applied to the field of geochemistry, providing a new perspective for the interpretation of geochemical data. Previous studies have highlighted various methods such as random forests [25,26], support vector machines [27,28], one-class support vector machines [29–31], neural networks [32,33], metric learning [34], and deep learning [35–41]. These studies have demonstrated the tremendous potential of artificial intelligence technologies, accelerating new discoveries and innovative solutions in the field of geochemistry.

The field of geochemistry encompasses a wide range of research areas, ranging from the microscale of atoms and molecules to the macroscale interactions within various spheres of the Earth system. The scale of this research makes geochemistry a challenging field that requires extensive data analysis and computation. The application of big data technology can assist geochemists in collecting, processing, and analyzing data more efficiently. For example, the real-time monitoring and data collection of chemical substances in global soil, water, and atmosphere can facilitate more accurate predictions and control of their impacts on the environment and human health [42–44]. Additionally, big data can be utilized in the development and optimization of geochemical models, enabling researchers to simulate and predict chemical processes within the Earth system more accurately [45–47]. At the same time, the application of artificial intelligence can further enhance the efficiency and precision of geochemical research. Techniques such as machine learning and deep learning can automatically identify and predict patterns and trends in geochemical processes. For instance, artificial intelligence can assist researchers in automatically classifying and identifying outliers in geochemical data, reducing human errors and improving research precision [48–52].

With the emergence of an increasing number of big data methods, the integration of geochemistry with big data and artificial intelligence has become increasingly prominent, demonstrating the vitality of intelligent research in geochemistry. Simultaneously, as globalization continues to advance, international collaboration in the field of research has become more prevalent. By analyzing the structure and dynamic changes of global collaboration networks, insights can be gained into the development trends and frontier hotspots within the field of geochemistry. However, geochemical research encompasses numerous subfields, such as ore deposits, petrology, and mineralogy, each with its specific research methods and objects. This diversity results in challenges in attaining a comprehensive understanding of the depth and breadth of research. Furthermore, existing research often focuses on specific areas or issues due to limitations in data sources and collection methods, lacking a global perspective and future foresight. To gain in-depth insights into the current development of the integration of geochemistry with big data and artificial intelligence, this paper provides a comprehensive review of recent relevant studies. Through in-depth analysis and mining of the related literature and data, it summarizes the research progress of big data and artificial intelligence methods in geochemistry (including mineralogy, petrology, and ore deposits). It also reveals the characteristics of the global collaboration network in the integration of geochemistry with big data and artificial intelligence, as well as the frontier and hotspots in this field. The purpose of this paper is to provide valuable reference information for related research fields, offer new perspectives and ideas to researchers, and promote the deep integration and development of geochemistry with big data and artificial intelligence. The rest of this paper is organized as follows: Section 2 presents the selection of data and research methods. Section 3 analyzes the characteristics of cooperation networks in research. Section 4 discusses the analysis of research hotspots. Section 5 concludes the paper and provides future prospects.

2. Data and Methods

2.1. Data Resources

This study aims to investigate the latest research trends in the field of geochemistry–big data–artificial intelligence (GC-AI) over the past decade. To achieve this, we conducted a literature search using the Web of Science database with the following keywords: "geochemi*," "mineralogy," "petrology," "Ore deposit," "data science," "data-driven," "big data," "artificial intelligence," "machine learning," "deep learning," "neural network," "big data analysis," and "big data method." The search was limited to articles published between 1 January 2013 and 1 August 2023. Initially, we retrieved 2707 articles. After removing duplicates, we ended up with a final dataset of 2421 articles.

2.2. Methods

Bibliometric analysis enables the visualization of knowledge map characteristics in specific fields, providing insights into the development and trends of scientific research. The most commonly used bibliometric visualization software programs include CiteSpace (version 5.4), VOSviewer, and HistCite, which can effectively illustrate the evolving trends in a particular field over time. VOSviewer, while lacking cluster labeling and supporting only one clustering algorithm, has its own strengths. Similarly, HistCite offers fewer visualization analysis methods. Meanwhile, CiteSpace stands out with its clear developmental background, efficient relationship visualization, and comprehensive analysis and display of bibliometric data.

CiteSpace, as an information visualization analysis program, is used for exploring latent knowledge within scientific research [53,54]. It includes cooperative network analysis and keyword analysis. The cooperative network analysis reveals the collaborative relationships between different countries/regions and institutions in the field, while keyword analysis involves co-occurrence, cluster, and burst analysis. Co-occurrence and cluster analysis unveil the key topics, keywords, and frontier knowledge, while burst analysis identifies research hotspots and potential future trends [55,56]. In this study, CiteSpace version 5.4 was utilized for bibliometric analysis in the GC-AI field, and the research workflow is depicted in Figure 1.

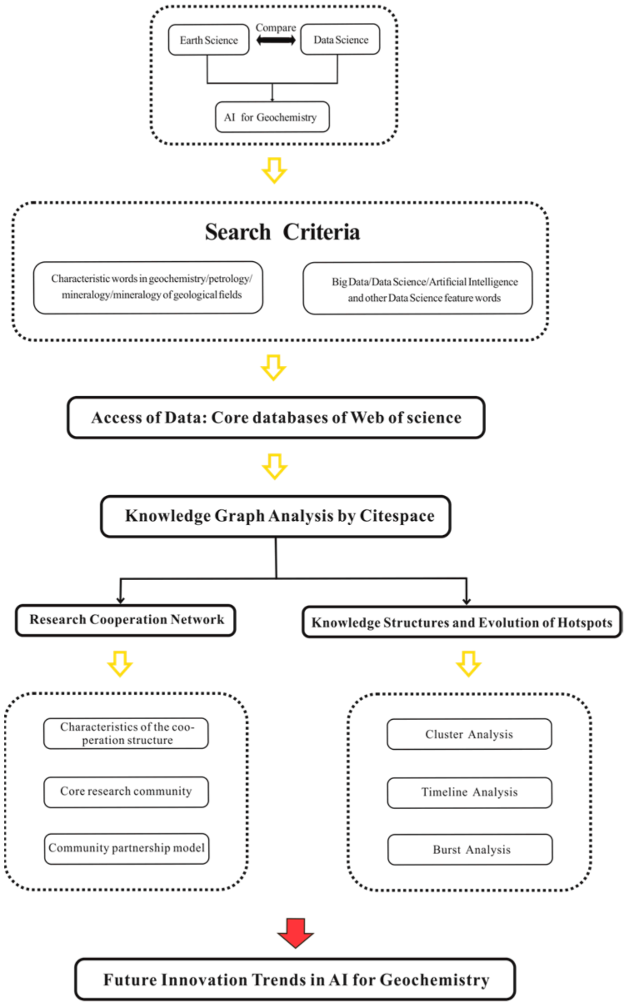

Figure 1. Bibliometric analysis process of GC-AI.

3. Analysis of Research Network Cooperation

3.1. Parameter Settings

CiteSpace has several important parameters and terminology, and adjusting these parameters plays an important role in the analysis of the knowledge graph.

3.1.1. G-Index

The G-index is an index based on citations, which sorts all papers by an author in descending order according to citations, and the largest g value is determined such that the total citations of the top g papers are not less than g^2. Typically, a higher G-index indicates a greater influence of the author in the field. The formula is as follows:

$$g^2 \leq k\sum_{i\leq g} c_i, k \in Z^+ \quad (1)$$

The value of k is positively correlated with the number of nodes.

3.1.2. Top N and Top N%

Top N refers to the selection of the top N citations based on their citation counts, which can be used for analyzing important studies. Top N% refers to the percentage of citations selected.

In this study, the analysis of collaboration networks resulted in a g-value of 15, while the k-value for the literature citation network was 5. The top N and top N% values were set to 50 and 10%, respectively.

3.1.3. Centrality

Centrality is a metric defined for each node within a network, measuring the likelihood of the node being on any shortest path within the network. It is used to identify and quantify the importance of nodes, and nodes with high centrality are often key hubs connecting two different domains.

3.1.4. Pruning

Pruning is a data processing method used to filter out unimportant nodes and edges. This includes using path-finding algorithms such as Pathfinder, minimum spanning tree algorithms such as Minimum Spanning Tree, and network slicing algorithms such as Pruning Sliced Networks to trim nodes for better display of the core structure and main features of the research network. This study is centered around a small-scale network structure. It is worth noting that certain algorithms, such as MST, may not yield a singular solution, resulting in multiple potential pruning scenarios, which could introduce ambiguity in the analysis findings. In contrast, Pathfinder offers a distinct pruning solution that enhances the stability and replicability of the analysis. Moreover, Pathfinder selectively retains nodes of higher significance within the knowledge network, facilitating the improved identification of key nodes. Consequently, Pathfinder was selected for node pruning in this study.

3.2. Analysis of Publications

A statistical analysis of the annual distribution of literature data exported from WOS was conducted to reflect the development trend of research regarding big data and artificial intelligence in the field of geochemistry. As shown in Figure 2, research in this field before 2017 was in a stable and ascending stage, with a small number of average publications ranging between 100 and 160 papers. Since 2017, the number of publications has shown a clear upward trend, with a rapid increase every year. In 2022, the number of publications reached 370, which is 3.7 times that of 2013, indicating that research on this topic has progressed rapidly and has attracted increasing attention and emphasis from researchers. Figure 3 presents the data distribution pattern in the field of GC-AI, with a total of 583 countries/regions, 7318 institutions, and 544 journals with 2421 publications.

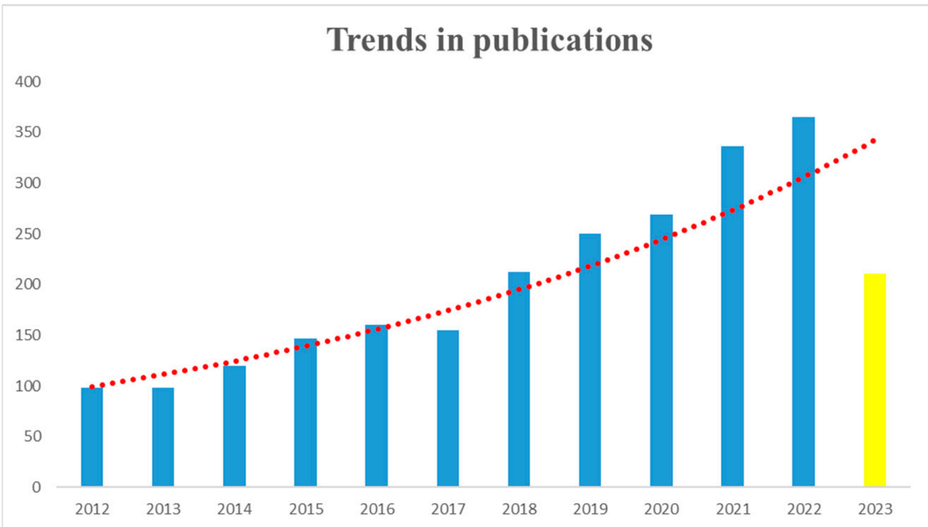

Figure 2. Trends in the number of publications of GC-AI.

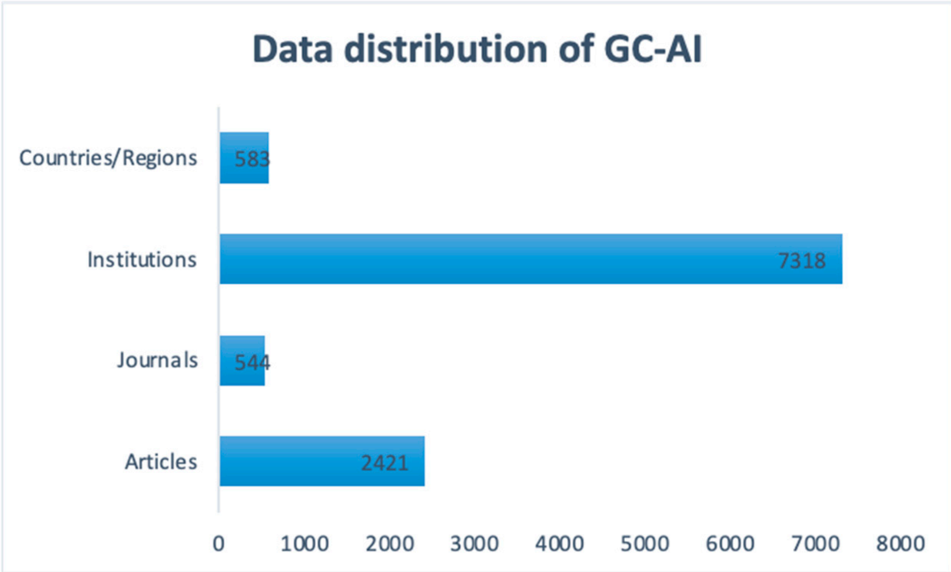

Figure 3. Data distribution in publications of GC-AI.

3.3. Cooperative Network of Countries/Regions

Analyzing the cooperative networks among countries/regions using CiteSpace allows us to comprehend the cooperative relationships in a specific research field across different nations. This analysis helps us identify countries with a high level of cooperative awareness and cooperation, thus inferring the international collaborative characteristics of the field. Figure 4 depicts the cooperative network of country/region-based research in the GC-AI field, with Table 1 presenting the top 10 countries/regions based on publication output.

From Figure 4 and Table 1, it is evident that the main research countries are the United States, China, Germany, the United Kingdom, Australia, and Canada. The United States ranks first with 729 papers, followed by China (620 papers), and Germany ranks third with 290 papers. The United Kingdom (270 papers) ranks fourth, with Australia (254 papers) rounding out the top five. The international cooperation among these countries is closely interwoven and exhibits a substantial level of cooperation. Furthermore, Table 1 reveals that Germany (0.23), the United Kingdom (0.16), and Canada (0.15) show the highest centrality, indicating their relative prominence within the cooperative network of countries/regions.

The findings demonstrate that the field of GC-AI exhibits a high level of research proficiency and internationalization, with the cooperative network of countries for research in this field characterized by close and diverse ties.

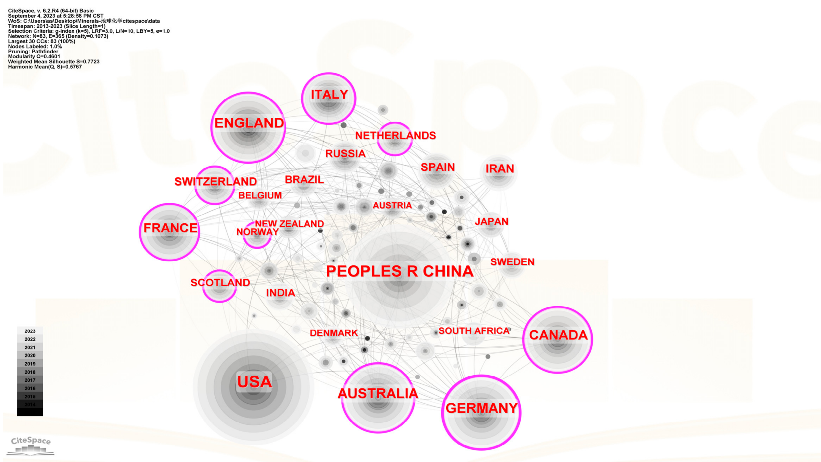

Figure 4. Cooperative network of countries/regions.

Table 1. Top 10 countries/regions in publications.

No.	Counts	Counts/Total	Countries	Centrality
1	729	30.1%	USA	0.1
2	620	25.6%	China	0.02
3	290	12.0%	Germany	0.23
4	270	11.2%	England	0.16
5	254	10.5%	Australia	0.13
6	226	9.3%	Canada	0.15
7	176	7.3%	France	0.12
8	150	6.2%	Italy	0.12
9	92	3.8%	Iran	0.03
10	83	3.4%	Spain	0.07

3.4. Cooperative Network of Research Institution

In conducting an analysis of the institutional cooperation network in the GC-AI field, it is observed that there is extensive cooperation between universities, as well as between universities and research institutions (Figure 5). This cooperative network is primarily composed of academic institutions such as universities and research institutes, with three main cooperative groups emerging: China, Europe, and the United States. However, these groups exhibit distinct geographical divisions, with the most cooperation within institutions being limited to the same country, and international cooperation across borders being relatively scarce. The top 10 institutions based on publication output are listed in Table 2, with China University of Geosciences (193), Chinese Academy of Sciences (185), and Centre National de la Recherche Scientifique (122) occupying the top three positions. Among them, the Helmholtz Association demonstrates the highest centrality (0.32), followed by the National Aeronautics and Space Administration and UDICE French Research Universities, with centrality values of 0.27 and 0.24, respectively. These centrality values signify their gross influence, rich research expertise, and available resources in the field.

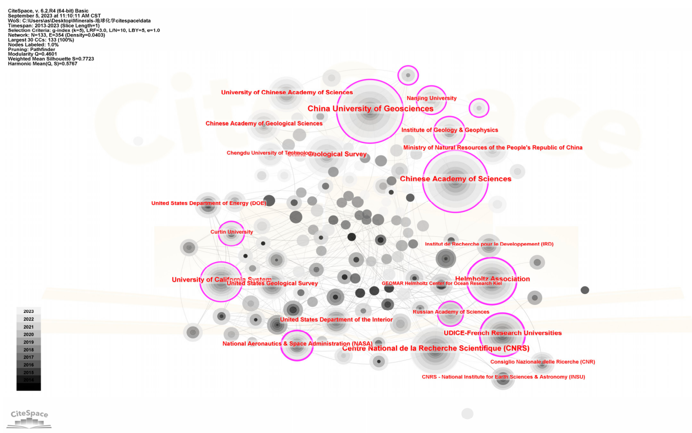

Figure 5. Cooperative network of research institutions.

Table 2. Top 10 institutions in publications.

No.	Count	Institutions	Centrality
1	193	China University of Geosciences	0.11
2	185	Chinese Academy of Sciences	0.13
3	122	Centre National de la Recherche Scientifique (CNRS)	0.07
4	113	Helmholtz Association	0.32
5	89	China Geological Survey	0.04
6	87	UDICE French Research Universities	0.24
7	81	University of California System	0.17
8	61	University of Chinese Academy of Sciences	0.01
9	53	United States Department of the Interior	0.1
10	53	United States Geological Survey	0.1

4. Discussion of Research Hotspots

4.1. Keywords Clustering

High-frequency and highly central keywords illustrate the focal points of most authors within a given time period, indicating the research hotspots and frontiers. The co-occurrence network of keywords is shown in Figure 6, where nodes represent keywords and the size of each node represents the frequency of co-occurrence. The color of the lines between nodes reflects the chronological order of their appearance. It can be observed that the research hotspots in the GC-AI field include climate change, water, geochemical models, geochemical evolution, sediments, machine learning, and mineral prospectivity mapping.

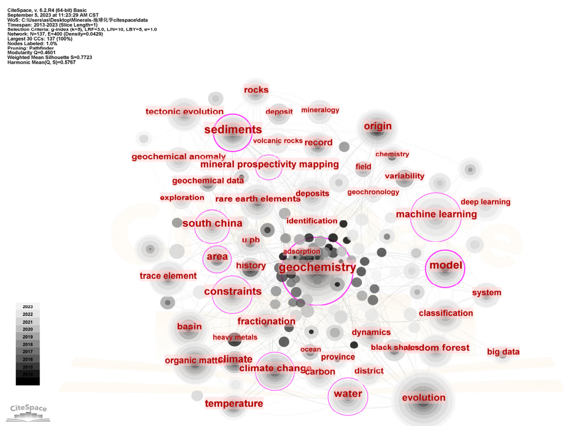

Figure 6. Keywords concurrence networks of GC-AI.

Based on the analysis of the graph data, the top 10 keywords with centrality in GC-AI research literature are presented in Table 3. It can be observed that there are three nodes with centrality values above 0.2 and six nodes with centrality values above 0.15. These nodes can be considered core nodes within the GC-AI field, reflecting the significance and influence of these keywords in the research domain. These keywords govern the development trends in this field.

Table 3. Top 10 keywords for centrality.

No.	Count	Centrality	Keywords
1	165	0.27	Model
2	109	0.24	Geochemistry
3	50	0.22	Sediments
4	69	0.18	Area
5	97	0.16	Water
6	90	0.15	South China
7	42	0.14	Constraints
8	36	0.14	Climate change
9	37	0.1	Machine learning
10	57	0.1	Mineral prospectivity mapping

Meanwhile, in this study, K-means clustering and LSI algorithms were employed to conduct network analysis on the keywords. The silhouette coefficient was utilized to evaluate the similarity between elements within the clusters. Ranging from 0 to 1, the silhouette coefficient measures the compactness of samples within the cluster and the separation between clusters. A value greater than 0.5 indicates a reasonable clustering, while a value exceeding 0.7 suggests a well-defined and convincing cluster. Figure 7 illustrates seven clusters in the order of 0 to 6, namely “machine learning” (0.8) for cluster #0, “climate change” (0.794) for cluster #1, “rare earth elements” (0.705) for cluster #2, “geochemistry” (0.673) for cluster #3, “mineral prospectivity mapping” (0.883) for cluster #4, “uranyl” (0.864) for cluster #5, and “deep sea” (0.909) for cluster #6. The size of clusters and their corresponding silhouette coefficients are summarized in Table 4. It can be observed that clusters 0, 1, and 2 contain the largest number of keywords, resulting in compact samples within the clusters and well-separated samples between clusters. All seven clusters have silhouette coefficients greater than 0.5, with six of them exceeding 0.7, indicating that this method successfully achieved the clustering of the GC-AI field with favorable performance.

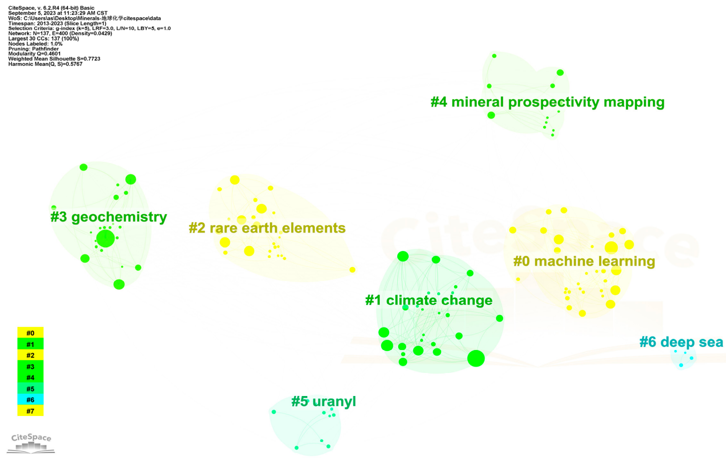

Figure 7. Keyword clustering of GC-AI.

Table 4. Silhouette ranking for keyword clustering.

Clusters	Count	Silhouette
Machine learning	62	0.8
Climate change	55	0.794
Rare earth elements	54	0.705
Geochemistry	30	0.673
Mineral prospectivity mapping	29	0.883
Uranyl	20	0.864
Deep sea	14	0.909

4.2. Keyword Timeline

To elucidate the developmental trajectory of this field, we conducted a timeline analysis of keyword progression. Figure 8 illustrates the progression of seven clusters during the past decade. Based on the law of timeline development, it can be inferred that the field of GC-AI has expanded into two crucial directions, from the advancement of research techniques to the breakthrough of exploration areas. This indicates the incessant strides made by the field in terms of theorization and application.

1. In the initial phase of GC-AI, researchers concentrated on utilizing rule-based techniques to recognize and categorize patterns and characteristics in geochemical data applications, including Bayesian networks and decision trees, among others. The field of geochemical artificial intelligence started to shift from rule-based methods to machine learning with the emergence of artificial intelligence (AI) technologies. Machine learning techniques enable automatic learning of data features and patterns without the need for manual feature design and selection, thus saving time. Some examples of these techniques include random forest, artificial neural networks, deep neural networks (i.e., deep learning), and others.
2. In recent years, the study and implementation of geochemical artificial intelligence have broadened to encompass various domains. These involve deploying machine learning algorithms for ore formation prophecy, the automated detection and prediction of groundwater contamination, the quantitative examination of rare earth elements' distribution patterns, climate change forecasting and prevention, and more.

Currently, the application domain is expanding continuously, and investigations on the deep Earth, deep space, and deep sea are the primary focus of upcoming research.

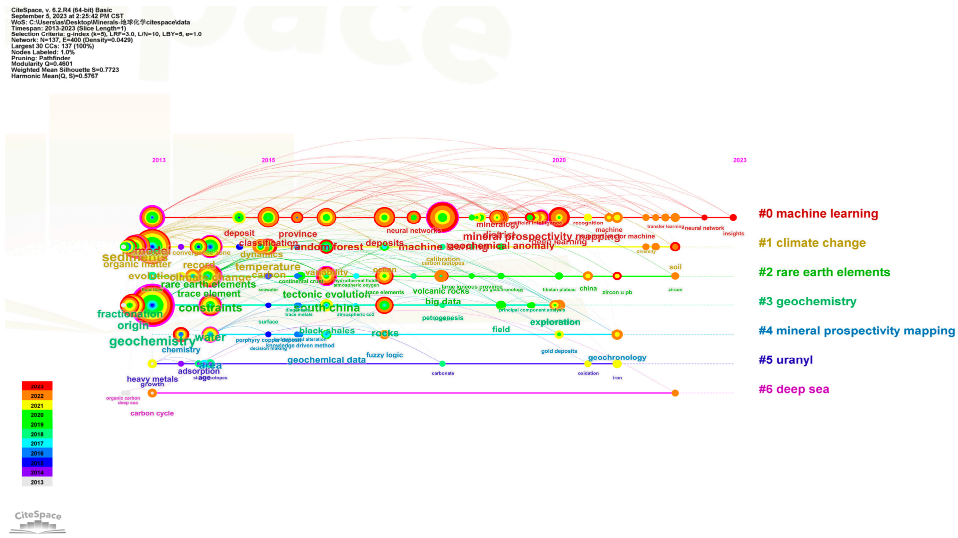

Figure 8. Keywords clustering timeline of GC-AI.

4.3. Keyword Burst Analysis

The field of GC-AI has undergone research priorities and changes at different time periods. To ascertain research hotspots in this field, a keyword burst analysis was undertaken. The burst word detection function was implemented to scrutinize keywords that had significantly changed in frequency in short-term research articles, demonstrating research topic hotspots at various time periods and exploring future research development trends. This study identified 37 keywords that had experienced bursts (Figure 9). “Strength” denotes the burst strength of the burst term, while “Begin” and “End” signify the start and end years of the burst term, respectively. The blue section denotes the timeline, while the red section reflects the significance and focus of keywords in this field of research. A longer burst length implies that the popularity of the keyword is sustained for a longer duration, thus indicating a stronger research front. From the diagram, it can be observed that the greatest surge in GC-AI research was in the areas of climate, fluid flow, and sediments. However, their research popularity has waned, suggesting that these were the most popular research subjects in the early stages of this field.

From the perspective of evolutionary trends, burst words emerged during three intensive periods, namely 2013–2014, 2016–2018, and 2019–2021, indicating the development of research areas and methods. In the field of GC-AI, specific keywords such as “model” and “evolution” burst relatively early. The appearance of these keywords suggests that researchers in the field have been interested in using data analysis techniques to address geochemical problems since the early stages. Over time, novel keywords, including “artificial neural networks” and “support vector machine”, have gradually surfaced, highlighting the field’s emphasis on artificial intelligence methods. Initially, research efforts were directed towards studying climate, crust, rocks, and other related areas. Nevertheless, with the advent of breakthroughs in research technology and the demands of social development, there is now an emerging focus on mineral resource research, which has become a hot topic in the field.

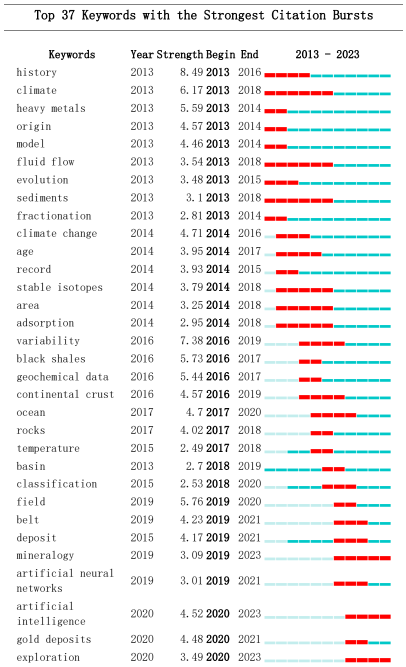

Figure 9. Keywords burst of GC-AI.

4.4. Discussion

Scientific research is currently experiencing its fourth paradigm [1], and data science has brought about significant advancements in various industries. Big data mining and artificial intelligence are necessary research topics for advances in geosciences [2,3,45]. Geochemical analysis based on big data has been gradually acknowledged by geochemical researchers in recent years due to the intersection of geochemistry and data science [57–60]. The use of a global cooperative network has emerged as the primary model for cooperation in the field of geochemistry and artificial intelligence. An examination of pertinent research and data reveals that this model displays characteristics of multidisciplinary intersection, spanning diverse data types such as geochemical and bibliographic data, and covering varied fields, including Earth science, environmental science, and information science. Simultaneously, the worldwide cooperative network is increasingly adopting co-authored papers as the primary means of contact, leading to the creation of a network based on regions, institutions, and themes. Consequently, these cooperative networks serve as a platform for scientific researchers to connect and cooperate, endorsing the sharing and refinement of scientific research resources and advancing innovation and the development of scientific research.

Although the field of geochemical artificial intelligence has made many advances, there are still many challenges and problems that need further research and resolution. For example, how to improve the generalization and robustness of models, how to handle imbalanced data distribution, and how to apply advanced algorithms to practical scenarios are issues that need further research and resolution in this field [4,45,61–64]. In addition, compared with other fields, research achievements and applications in geochemical artificial intelligence are relatively limited, requiring more innovation and breakthroughs.

In the future, new geochemical artificial intelligence technologies and applications can be further explored and developed. For example, generative models such as generative adversarial networks can be used for geochemical data generation and processing, and transfer learning technology can be applied to transfer models from other fields into geochemistry. Strengthening international cooperation and communication in the field of geochemical artificial intelligence to promote the development and application of this field will make greater contributions to solving global environmental issues.

5. Conclusions and Prospect

This article presents an all-encompassing evaluation of the status and research hotpots concerning the global cooperative network that merges geochemistry with artificial intelligence. Through knowledge graph analysis of geochemistry and artificial intelligence publications published within the past 10 years, we gained the following insights:

1. Research trend: Based on the analysis of publications in the literature, it was observed that the research in this field witnessed a gradual increase before 2017, followed by a notable surge in the number of publications thereafter, with a steep annual rise. By 2022, the volume of publications had surged to 3.7 times that of 2013, highlighting the increasing significance of artificial intelligence techniques in the research toolbox of geochemists. The study of artificial intelligence techniques for geochemistry is quickly progressing within the GC-AI field.
2. Characteristics of cooperative network: A network analysis was conducted on the national/regional and institutional cooperative network to uncover research collaboration patterns in this area. The results indicate a high research standard and internationalization level. The cooperative network has the characteristics of being close-knit and diverse. Research institutions from China, the United States, and Europe are prominent in this field, with many cooperative relationships and research outcomes. Notably, the National Aeronautics and Space Administration and Université libre de Bruxelles exhibit a high level of research and impact.
3. Research hotspots: Through analysis of keyword co-occurrence networks and clustering, timeline, and burst word detection, it was found that some clusters have high relevance and similarity in the field of geochemical artificial intelligence. The application of artificial intelligence technology for mineral resource prediction is currently a research hotspot, and deep space, deep Earth, and deep sea mineral resource exploration will become future research trends.

Author Contributions: Conceptualization, Q.Z. and Y.Z.; methodology, Q.Z.; software, Q.Z.; validation, Q.Z., J.H. and F.H.; formal analysis, Q.Z.; investigation, Q.Z.; resources, Y.Z.; data curation, B.Z.; writing—original draft preparation, Q.Z.; writing—review and editing, Q.Z.; visualization, Q.Z. and S.L.; supervision, Y.Z.; project administration, Y.Z.; funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was financially supported by the National Key Research and Development Program of China (Grant No. 2022YFF0801201), the National Natural Science Foundation of China (Grant No. U1911202), and the Science and Technology Project of Guangdong Provincial Department of Science and Technology (Grant No. GDKTP2020053500).

Data Availability Statement: Data are available upon request from the authors.

Acknowledgments: We express our utmost gratitude to Fan Xiao and Pengpeng Yu from Sun Yat-sen University. Furthermore, we extend our sincere appreciation to Renguang Zuo, the esteemed Guest Editor of this Special Issue, for his exceptional organization and guidance.

Conflicts of Interest: The authors declare no conflict of interest. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

1. Tolle, K.M.; Tansley, D.S.W.; Hey, A.J.G. The fourth paradigm: Data-intensive scientific discovery. *Proc. IEEE* **2011**, *99*, 1334–1337. [CrossRef]
2. Zhou, Y.; Zuo, R.; Liu, G.; Yuan, F.; Mao, X.; Guo, Y.; Xiao, F.; Liao, J.; Liu, Y. The great-leap-forward development of mathematical geoscience during 2010–2019: Big data and artificial intelligence algorithm are changing mathematical geoscience. *Bull. Miner. Pet. Geochem.* **2021**, *40*, 556–573. [CrossRef]
3. Zhang, W.G.; Ching, J.Y.; Goh, A.T.C.; Leung, A.Y.F. Big data and machine learning in geoscience and geoengineering: Introduction. *Geosci. Front.* **2020**, *12*, 327–329. [CrossRef]
4. Zhang, W.; Gu, X.; Tang, L.; Yin, Y.; Liu, D.; Zhang, Y. Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: Comprehensive review and future challenge. *Gondwana Res.* **2022**, *109*, 1–17. [CrossRef]

5. Petrelli, M.; Bizzarri, R.; Morgavi, D.; Baldanza, A.; Perugini, D. Combining machine learning techniques, microanalyses and large geochemical datasets for tephrochronological studies in complex volcanic areas: New age constraints for the Pleistocene magmatism of central Italy. *Quat. Geochronol.* **2017**, *40*, 33–44. [CrossRef]
6. Jiao, S.; Zhang, Q.; Zhou, Y.; Chen, W.; Liu, X. Gopalakrishnan. Progress and challenges of big data research on petrology and geochemistry. *Solid Earth Sci.* **2018**, *3*, 105–114. [CrossRef]
7. Han, S.; Li, M.; Zhang, Q.; Song, L. An Automated Method to Generate and Evaluate Geochemical Tectonic Discrimination Diagrams Based on Topological Theory. *Minerals* **2020**, *10*, 62. [CrossRef]
8. Liu, X.; Zhang, Q.; Zhang, C. Identification of the Original Tectonic Setting for Oceanic Andesite Using Discrimination Diagrams: An Approach Based on Global Geochemical Data Synthesis. *J. Earth Sci.* **2022**, *33*, 696–705. [CrossRef]
9. O'Sullivan, G.; Chew, D.; Kenny, G.; Henrichs, I.; Mulligan, D. The trace element composition of apatite and its application to detrital provenance studies. *Earth Sci. Rev.* **2020**, *201*, 103044. [CrossRef]
10. Saha, R.; Upadhyay, D.; Mishra, B. Discriminating tectonic setting of igneous rocks using biotite major element chemistry—A machine learning approach. *Geochem. Geophys. Geosyst.* **2021**, *22*, e2021GC010053. [CrossRef]
11. Zheng, D.; Wu, S.; Ma, C.; Xiang, L.; Hou, L.; Chen, A.; Hou, M. Zircon classification from cathodoluminescence images using deep learning. *Geosci. Front.* **2022**, *13*, 101436. [CrossRef]
12. Wang, C.-L.; Zhu, W.-N.; Wang, Y.-T. Genesis of the Kiruna-type Nixintage iron deposit, Chinese Western Tianshan, NW China: Constrains of ore geology, geochemistry and geochronology. *J. Geochem. Explor.* **2022**, *243*, 107094. [CrossRef]
13. Chen, N.; Mao, J.; Zhang, Z.; Duan, Z.; Santos, A.; Li, H. Arc magmatic evolution and porphyry copper deposit formation under compressional regime: A geochemical perspective from the Toquepala arc in Southern Peru. *Earth-Sci. Rev.* **2023**, *240*, 104383. [CrossRef]
14. Zhang, P.; Zhang, Z.; Yang, J.; Cheng, Q. Machine Learning Prediction of Ore Deposit Genetic Type Using Magnetite Geochemistry. *Nat. Resour. Res.* **2023**, *32*, 99–116. [CrossRef]
15. Zuo, R.; Xiong, Y. Big data analytics of identifying geochemical anomalies supported by machine learning methods. *Nat. Resour. Res.* **2018**, *27*, 5–13. [CrossRef]
16. Chen, Y.; Zhao, Q. Mineral exploration targeting by combination of recursive indicator elimination with the ℓ2-regularization logistic regression based on geochemical data. *Ore Geol. Rev.* **2021**, *135*, 104213. [CrossRef]
17. Hawkes, H.E.; Webb, J.S. *Geochemistry in Mineral Exploration*; Harper and Row: New York, NY, USA, 1962.
18. Sinclair, A.J. Selection of threshold values in geochemical data using probability graphs. *J. Geochem. Explor.* **1974**, *3*, 129–149. [CrossRef]
19. Tukey, J.W. *Exploratory Data Analysis*; Addison-Wesley: Reading, PA, USA, 1977.
20. Reimann, C.; Filzmoser, P.; Garrett, R.; Dutter, R. *Statistical Data Analysis Explained: Applied Environmental Statistics with R*; John Wiley & Sons: Hoboken, NJ, USA, 2008.
21. Matheron, G. *Traité de Géostatistique Appliquée*; Editions Technip: Paris, France, 1962.
22. Cheng, Q.; Agterberg, F.P.; Ballantyne, S.B. The separation of geochemical anomalies from background by fractal methods. *J. Geochem. Explor.* **1994**, *51*, 109–130. [CrossRef]
23. Cheng, Q.; Xu, Y.; Grunsky, E. Integrated spatial and spectrum method for geochemical anomaly separation. *Nat. Resour. Res.* **2000**, *9*, 43–52. [CrossRef]
24. Cheng, Q. Mapping singularities with stream sediment geochemical data for prediction of undiscovered mineral deposits in Gejiu, Yunnan Province, China. *Ore Geol. Rev.* **2007**, *32*, 314–324. [CrossRef]
25. Gonbadi, A.M.; Tabatabaei, S.H.; Carranza, E.J.M. Supervised geochemical anomaly detection by pattern recognition. *J. Geochem. Explor.* **2015**, *157*, 81–91. [CrossRef]
26. Parsa, M.; Maghsoudi, A. Assessing the effects of mineral systems-derived exploration targeting criteria for random Forests-based predictive mapping of mineral prospectivity in Ahar-Arasbaran area, Iran. *Ore Geol. Rev.* **2021**, *138*, 104399. [CrossRef]
27. Zuo, R.; Carranza, E.J.M. Support vector machine: A tool for mapping mineral prospectivity. *Comput. Geosci.* **2011**, *37*, 1967–1975. [CrossRef]
28. Maepa, F.; Smith, R.S.; Tessema, A. Support vector machine and artificial neural network modelling of orogenic gold prospectivity mapping in the Swayze greenstone belt, Ontario, Canada. *Ore Geol. Rev.* **2020**, *130*, 103968. [CrossRef]
29. Chen, Y.; Wu, W. Mapping mineral prospectivity by using one-class support vector machine to identify multivariate geological anomalies from digital geological survey data. *Aust. J. Earth Sci.* **2017**, *64*, 639–651. [CrossRef]
30. Chen, Y.; Wu, W.; Zhao, Q. A bat-optimized one-class support vector machine for mineral prospectivity mapping. *Minerals* **2019**, *9*, 317. [CrossRef]
31. Xiong, Y.; Zuo, R. A positive and unlabeled learning algorithm for mineral prospectivity mapping. *Comput. Geosci.* **2020**, *147*, 104667. [CrossRef]
32. Ziaii, M.; Pouyan, A.A.; Ziaei, M. Neuro-fuzzy modelling in mining geochemistry: Identification of geochemical anomalies. *J. Geochem. Explor.* **2009**, *100*, 25–36. [CrossRef]
33. Yu, X.; Xiao, F.; Zhou, Y.; Wang, Y.; Wang, K. Application of hierarchical clustering, singularity mapping, and Kohonen neural network to identify Ag-Au-Pb-Zn polymetallic mineralization associated geochemical anomaly in Pangxidong district. *J. Geochem. Explor.* **2019**, *203*, 87–95. [CrossRef]

34. Wang, Z.; Dong, Y.; Zuo, R. Mapping geochemical anomalies related to Fe–polymetallic mineralization using the maximum margin metric learning method. *Ore Geol. Rev.* **2019**, *107*, 258–265. [CrossRef]
35. Xiong, Y.; Zuo, R. Recognition of geochemical anomalies using a deep autoencoder network. *Comput. Geosci.* **2016**, *86*, 75–82. [CrossRef]
36. Luo, Z.; Xiong, Y.; Zuo, R. Recognition of geochemical anomalies using a deep variational autoencoder network. *Appl. Geochem.* **2020**, *122*, 104710. [CrossRef]
37. Luo, Z.; Zuo, R.; Xiong, Y.; Wang, X. Detection of geochemical anomalies related to mineralization using the GANomaly network. *Appl. Geochem.* **2021**, *131*, 105043. [CrossRef]
38. Yang, N.; Zhang, Z.; Yang, J.; Hong, Z.; Shi, J. A convolutional neural network of GoogLeNet applied in mineral prospectivity prediction based on multi-source geoinformation. *Nat. Resour. Res.* **2021**, *30*, 3905–3923. [CrossRef]
39. Yang, N.; Zhang, Z.; Yang, J.; Hong, Z. Mineral Prospectivity Prediction by Integration of Convolutional Autoencoder Network and Random Forest. *Nat. Resour. Res.* **2022**, *31*, 1103–1119. [CrossRef]
40. Wang, Z.; Zuo, R. Mineral prospectivity mapping using a joint singularity-based weighting method and long short-term memory network. *Comput. Geosci.* **2022**, *158*, 104974. [CrossRef]
41. Li, T.; Zuo, R.; Zhao, X.; Zhao, K. Mapping prospectivity for regolith-hosted REE deposits via convolutional neural network with generative adversarial network augmented data. *Ore Geol. Rev.* **2022**, *142*, 104693. [CrossRef]
42. Lyons, K.J.; Ikonen, J.; Hokajärvi, A.-M.; Räsänen, T.; Pitkänen, T.; Kauppinen, A.; Kujala, K.; Rossi, P.M.; Miettinen, I.T. Monitoring Groundwater Quality with Real-Time Data, Stable Water Isotopes, and Microbial Community Analysis: A Comparison with Conventional Methods. *Sci. Total. Environ.* **2023**, *864*, 161199. [CrossRef]
43. Huang, J.; Wang, D.; Zhu, Y.; Yang, Z.; Yao, M.; Shi, X.; An, T.; Zhang, Q.; Huang, C.; Bi, X.; et al. An overview for monitoring and prediction of pathogenic microorganisms in the atmosphere. *Fundam. Res.* **2023**. [CrossRef]
44. Siddiqi, S.A.; Al-Mulla, Y. Wireless Sensor Network System for Precision Irrigation using Soil and Plant Based Near-Real Time Monitoring Sensors. *Procedia Comput. Sci.* **2022**, *203*, 407–412. [CrossRef]
45. Zuo, R.; Wang, J.; Xiong, Y.; Wang, Z. The processing methods of geochemical exploration data: Past, present, and future. *Appl. Geochem.* **2021**, *132*, 105072. [CrossRef]
46. Parsa, M.; Sadeghi, M.; Grunsky, E. Innovative methods applied to processing and interpreting geochemical data. *J. Geochem. Explor.* **2022**, *237*, 106983. [CrossRef]
47. Gelencsér, O.; Árvai, C.; Mika, L.; Breitner, D.; LeClair, D.; Szabó, C.; Falus, G.; Szabó-Krausz, Z. Effect of hydrogen on calcite reactivity in sandstone reservoirs: Experimental results compared to geochemical modeling predictions. *J. Energy Storage* **2023**, *61*, 106737. [CrossRef]
48. Wang, J.; Zhou, Y.; Xiao, F. Identification of multi-element geochemical anomalies using unsupervised machine learning algorithms: A case study from Ag–Pb–Zn deposits in north-western Zhejiang, China. *Appl. Geochem.* **2020**, *120*, 104679. [CrossRef]
49. Zhang, S.; Bourdeau, J.; Nwaila, G.; Ghorbani, Y. Advanced geochemical exploration knowledge using machine learning: Prediction of unknown elemental concentrations and operational prioritization of re-analysis campaigns. *Artif. Intell. Geosci.* **2022**, *3*, 86–100. [CrossRef]
50. Zuo, R.; Wang, J.; Yin, B. Visualization and interpretation of geochemical exploration data using GIS and machine learning methods. *Appl. Geochem.* **2021**, *134*, 105111. [CrossRef]
51. Zuo, R.; Xiong, Y.; Wang, J.; Carranza, E.J.M. Deep learning and its application in geochemical mapping. *Earth-Science Rev.* **2019**, *192*, 1–14. [CrossRef]
52. Mkono, C.N.; Chuanbo, S.; Mulashani, A.K.; Mwakipunda, G.C. Deep learning integrated approach for hydrocarbon source rock evaluation and geochemical indicators prediction in the Jurassic–Paleogene of the Mandawa basin, SE Tanzania. *Energy* **2023**, *2023*, 129232. [CrossRef]
53. Chen, C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *J. Am. Soc. Inf. Sci. Technol.* **2006**, *57*, 359–377. [CrossRef]
54. Chen, C.; Ibekwe-SanJuan, F.; Hou, J. The structure and dynamics of co-citation clusters: A multiple-perspective co-citation analysis. *J. Am. Soc. Inf. Sci. Technol.* **2010**, *61*, 1386–1409. [CrossRef]
55. Dan, Z.; Jiapeng, X.; Yizhu, Z. Study on sustainable urbanization literature based on Web of Science, scopus, and China national knowledge infrastructure: A scientometric analysis in CiteSpace. *J. Clean. Prod.* **2020**, *264*, 121537.
56. Zuo, Z.; Cheng, J.; Guo, H.; Li, Y. Knowledge mapping of research on strategic mineral resource security: A visual analysis using CiteSpace. *Resour. Policy* **2021**, *74*, 102372. [CrossRef]
57. Yu, S.; Deng, H.; Liu, Z.; Chen, J.; Gu, X.; Li, J.; Xiao, K.; Mao, X. Identifying multivariate geochemical anomalies via tensor dictionary learning over spatial-elemental dimensionalities. *Comput. Geosci.* **2022**, *165*, 105153. [CrossRef]
58. Doucet, L.S.; Tetley, M.G.; Li, Z.-X.; Liu, Y.; Gamaleldien, H. Geochemical fingerprinting of continental and oceanic basalts: A machine learning approach. *Earth-Science Rev.* **2022**, *233*, 104192. [CrossRef]
59. Salgado, L.; López-Sánchez, C.; Colina, A.; Baragaño, D.; Forján, R.; Gallego, J. Hg and As pollution in the soil-plant system evaluated by combining multispectral UAV-RS, geochemical survey and machine learning. *Environ. Pollut.* **2023**, *333*, 122066. [CrossRef]

60. Lindsay, J.J.; Hughes, H.S.; Yeomans, C.M.; Andersen, J.C.; McDonald, I. A machine learning approach for regional geochemical data: Platinum-group element geochemistry vs geodynamic settings of the North Atlantic Igneous Province. *Geosci. Front.* **2021**, *12*, 101098. [CrossRef]
61. Sadeghi, B.; Cohen, D.R. Decision-making within geochemical exploration data based on spatial uncertainty–A new insight and a futuristic review. *Ore Geol. Rev.* **2023**, *161*, 105660. [CrossRef]
62. Engle, M.A.; Chaput, J. Visualizing high dimensional structures in geochemical datasets using a combined compositional data analysis and Databionic swarm approach. *Int. J. Coal Geol.* **2023**, *275*, 104303. [CrossRef]
63. Parsa, M.; Shirazy, A.; Shirazi, A.; Pour, A.B. 4–Processing and interpretation of geochemical data for mineral exploration. In *Geospatial Analysis Applied to Mineral Exploration*; Elsevier: Amsterdam, The Netherlands, 2023; pp. 171–188, ISBN 9780323956086. [CrossRef]
64. Wang, J.; Zuo, R. Model averaging for identification of geochemical anomalies linked to mineralization. *Ore Geol. Rev.* **2022**, *146*, 104955. [CrossRef]