A Co-Citation and Cluster Analysis of Scientometrics of Geographic Information Ontology

Yu Liu [1], Lin Li [1,2,*] [ID], Hang Shen [1], Hui Yang [1] and Feng Luo [3]

[1] School of Resource and Environmental Sciences, Wuhan University, 129 Luoyu Road, Wuhan 430079, China; liuyu0201@whu.edu.cn (Y.L.); shenhang@whu.edu.cn (H.S.); yanghui@whu.edu.cn (H.Y.)

[2] Collaborative Innovation Center of Geo Spatial Technology, Wuhan University, 129 Luoyu Road, Wuhan 430079, China

[3] CCCC Second Highway Consultants Co., Ltd. No.18, Chuangye Road, Wuhan Economic and Technologic Development Zone, Wuhan 430056, China; lfeng@whu.edu.cn

* Correspondence: lilin@whu.edu.cn; Tel.: +86-138-7150-4963

Abstract: Geographic information ontology represents an effective means of expressing geographic concepts and relationships between them. As an emerging field of study, it has drawn the attention of increasing numbers of scholars worldwide. In this study, both co-citation and cluster analysis methods of scientometrics are used to perform a comprehensive analysis of the papers on the topic of geographic information ontology indexed by the Web of Science (WoS) and published between 2001 and 2016. The results show that the history of the study of geographic information ontology can be divided roughly into three periods. Computer science and mathematics play important roles in this field of study. The International Journal of Geographical Information Science is an important periodical that provides knowledge resources for the study of geographic information ontology. The papers of Gruber TR and Guarino N are referenced most frequently, as well as that of Smith B., who formally introduced information ontology to the field of geographic information science. Providing personalized and intelligent geographic information services for users is an important focus of geographic information ontology.

1. Introduction

Ontology is a branch of philosophy that aims to provide systematic explanations or descriptions of an objective existence. It concerns the abstract nature of objects [1]. In the 1960s, ontology was introduced into the fields of information science and computer science to assist in the reconstruction and sharing of knowledge, information, and data, which are integrated with the help of computers [2]. The concept of geographic information ontology first appeared in the 1990s, when ontology was introduced into the field of geographic information science [3]. In geographic information science, the concept of "geographic information ontology" now usually contains the dual meaning of information ontology and philosophical ontology [4]. On one hand, philosophical ontology pays sufficient attention to geographic concepts, categories, relations, and the geographical domain itself, especially spatio-temporal ontology, uncertainty ontology, scale ontology, and so on. Related studies can be recognized as pure theoretical research, and they rarely consider the realization of geographic information ontology on computers. On the other hand, by explicitly defining the concept of shared geographic concepts, information ontology is applied to the sharing and interoperation of geographic information and geographic information retrieval based on semantics, integration, and geographic information services [5]. Moreover, such studies can be regarded as applications of ontology that focus entirely on the integration and interoperation of

geographic information that is based on semantics, but they do not consider the theory of geographic information ontology to a great degree.

With the rapid development of network technology, there has been an increasing demand for online retrieving and analysis of geospatial data, which requires sharing and interoperability of the data among different geographic information systems. Nevertheless, existing problems such as multi-source data, multi-scale data, and heterogeneity restrict the exchange of geographic information and the discovery of knowledge [6,7]. The introduction of ontology into geographic information system (GIS) can standardize the process of storage and management of spatial data, improve the precision of spatial query and spatial analysis in applications, and promote the sharing of geography information in the network environment [3]. Currently, geographic information ontology is a highly interdisciplinary field of study that is closely related to geographical cognition, geographic information expression, and interoperability of geographic information [8–10].

Over the past 20 years, many sub-areas of study have formed as the study of geographic information ontology has progressed. These sub-areas include geographic information ontology integration [11,12], geographic information ontology semantic interoperability [13], and ontology-driven geographic information systems (ODGISs) [14]. In the course of studying geographic information ontology, scholars have published thousands of papers on this topic. Due to time and energy constraints, and considering the diversity of sub-areas, it would be difficult to conduct a comprehensive and objective review of geographic information ontology without quantitative analysis methods [15]. Co-citation and cluster analyses based on scientometrics provide objective and quantitative methods for identifying the core knowledge and academic communities of particular fields [16]. Through the co-citation and cluster analysis of subjects, journals, and authors, this study attempts to identify some of the characteristics of geographic information ontology, reveal the potential academic structure and knowledge flow, and explain the evolution of the study of geographic information ontology to help researchers obtain continuous updates on knowledge relating to the development of the field of geographic information ontology.

At present, there are several commonly used academic search platforms, such as Google Scholar, arXiv.org, CNKI, and Web of Science (WoS). Google Scholar is a free academic search engine that is run by Google. It can help users search academic information conveniently including academic works published by publishers and professional communities, preprints, and peer-reviewed papers from universities and other academic organizations [17]. arXiv.org is an online database that collects preprints of science literature with a scale of more than 700,000. The most important characteristic of arXiv.org is that it is "open access", which means that everyone has access to full-text documents for free [18]. CNKI is presently the largest full-text database of Chinese journals all over the world, and mainly focuses on science, technology, and education. WoS is a large comprehensive citation index database of core journals from multiple subjects, which was developed by Thomson Reuters (New York, NY, USA) based on a web developing platform. It includes the Science Citation Index Expanded, Social Sciences Citation Index, Arts & Humanities Citation Index, Conference Proceedings Citation Index, etc. It covers approximately 12,000 leading journals worldwide with citation references across 256 disciplines, and provides powerful web-based access to bibliographic and citation information (including titles, authors, abstract, keywords, dates, author addresses, subject categories, reference lists, etc.). Types of documents in WoS are Article, Proceedings Paper, Letter and Review, etc., while newspapers, publicly published ontology data, and ontology repositories are not included. WoS bibliographic and citation information can be downloaded, which enables tracing of the intellectual roots of the field under study [19]. Based on these reasons, WoS is often considered to be the best suited data source for current scientometrics investigations, and a large number of scientometrics articles choose WoS database as data source [20].

In this study, a co-citation and cluster analysis was conducted on the papers that are indexed by the WoS database on the subject of geographic information ontology, published from 2001 to 2016. The goals of this analysis were (1) to show the temporal evolution of geographic information ontology; (2) to identify the key disciplines and reveal potential relationships between disciplines in the study

of geographic information ontology; and (3) to analyze frequently cited journals, authors, and their study results and to explore the main fields of study and academic communities that utilize geographic information ontology. The results of this study will help readers, especially newcomers, understand the meaning, origin, current state, and outcomes of geographic information ontology as a field of study and provide a reference for future research.

2. Data Sources and Methods

First, the keywords "geo* AND ontolog*" were applied to the WoS core database and were used to retrieve and download relevant papers on the subject of "geographic information ontology". The range of publication dates was set as 2001–2016. In this way, we download all the retrieved data.

Next, there are five researchers in our team, all of whom are engaged in the study of geographic information ontology. Four of the people were divided into two groups, and each group was respectively assigned to check the detail of records by verifying article titles, author names, abstracts, keywords, affiliated institution names, and publication dates excluded. The decision to exclude a record was discussed by the two researchers in the group. If they could not reach an agreement, the fifth researcher joined the discussion and voted to make the decision. Finally, 1682 articles remained. To be consistent with other scientometrics research [20,21], we chose 1533 articles in the "article" form for our analysis.

In this study, the CiteSpace was employed for the cluster analysis with Spectral Clustering Algorithm. Spectral clustering is an algorithm based on graph theory, and compared to other clustering methods, such as K-means clustering and EM, it has an evident advantage in co-citation network clustering that is based on connected relations rather than node properties. The K-means and EM clustering algorithms are based on a spherical sample space. When the sample space is not convex, the K-means and EM algorithms will converge to a local optimum [19]. However, derived from the spectral graph theory, the spectral clustering algorithm can work on any spatial dataset without considering its type and converge to the global optimal value [22,23]. Therefore, the spectral clustering algorithm has become a widely used method in the research of scientometrics and information science. Besides this, Excel was deployed to make the histogram.

3. Results

3.1. Article Outputs

Geographic information ontology was developed in the 1990s when ontology was introduced into the field of geographic information science [3]. However, the total the number of articles published before 2000 was very small. The situation changed fundamentally after 2001, and persistent explosive growth in the number of papers related to geographic information ontology occurred from 2001 to 2009. Compared to 2001–2009, in 2010–2016 the trend of growth declined slightly, but the number of papers was more and relatively stable (Figure 1). According to the growth curve shown in Figure 1 and the characteristics and development of geographic ontology research, the research can be divided into the three following stages.

Embryonic stage (before 2000): During this time, some computer scientists and geographers had already used ontology technology in geographic information science research, such as Smith and Mark [24–27] and Fonseca and Egenhofer [25]. The growth in the number of papers was slow, and most of these papers were published in comprehensive journals.

Rapid development stage (2001–2009): During this period, 692 papers were published, and 54 countries and regions were involved in the research. The US contributed most of the published papers. The 238 papers from the US accounted for 35.69% of contemporary documents, followed by China (119 papers, 17.19%), the UK (117 papers, 16.91%), and Germany (81 papers, 11.71%). In this period, geographic information science developed rapidly, and the means of obtaining geospatial data increased in diversity. The emergence of geographic information systems, volunteered geographic information, and big data with different systems and structures hindered the communication and

sharing of geographic information and made it quite difficult for users to process massive amounts of geographic data [2]. Moreover, the expressions of these data focused on structure and ignored their semantics. How to provide knowledge-based geographic information services was an urgent issue [12]. Geographic information ontology can remove the obstacles to the sharing of multi-source, multi-data structure, and multi-classification systems, and can also produce a clear understanding of user requirements, thus providing knowledge-level smart geographic information services [3,28]. Therefore, studies related to geographic information ontology displayed considerable development in this stage.

Relatively stable stage (2010–2016): In this stage, 841 papers were published, and a total of 58 countries and regions participated in the research. Although the growth of papers somewhat declined compared to the papers published between 2001 and 2009, the number of papers is increasing, and the annual number fluctuated only slightly. In general, the number of published papers in this period tended to be stable.
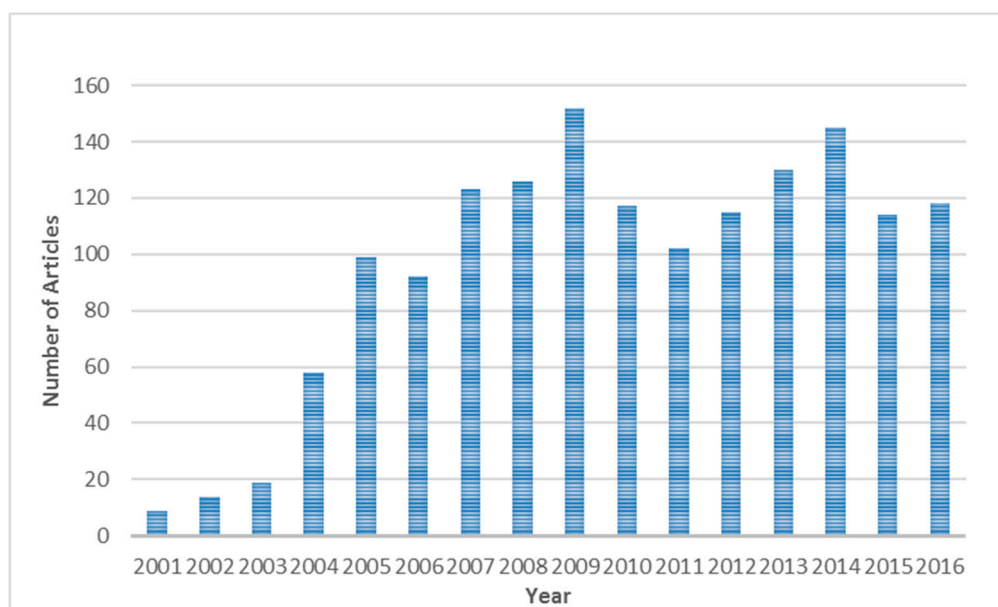
Figure 1. The number of articles published on geographic information ontology from 2001 to 2016.

3.2. Subject Categories

The disciplinary constitution of a specific research field reveals the degree of merging of disciplines involved in that research field and, to some degree, their impact [29]. Each paper in the WoS is classified into one or more subject categories. In this article, we use statistical analyses of subject categories to reveal the focus of geographic information ontology research. Moreover, the relationships among these subject categories are also revealed through cluster relationships based on the co-occurrence of geographic information ontology disciplines.

The 52 subject categories involved in geographic information ontology research occur 2256 times altogether. Note that some papers are classified into more than one category; when the statistical analyses are performed, these different categories are counted separately. For example, if a paper is classified as representing both computer science and physical geography, then we add 1 to the numbers of both computer science and physical geography papers. Computer science (868, 38.47%), engineering (285, 12.63%), and geography (267, 11.83%) are the top three categories in terms of number of publications, and these categories make up 70.61% of the papers published from 2001 to 2016. In addition, remote sensing (128, 5.67%) and mathematics (45, 1.99%) display the fastest growth among all of the subject categories.

To study the relationships among subjects regarding the study of geographic information ontology and their evolution, the CiteSpace software was used to perform a co-occurrence network cluster analysis of these subjects [19]. The time span from 2001 to 2016 was divided into four equal sections, and 20 subject categories with the highest frequency of reference are selected for analysis from each section to draw the subject co-occurrence network cluster graph, which is shown in Figure 2 and includes 28 nodes and 44 co-occurrence links. In this figure, the colors of the bars at the top represent the year limits, and each node represents a subject category. The colors at the centers of the nodes represent the dates of the earliest appearance of the subjects, and the circular structures that encircle the nodes represent the history since each subject's appearance. The colors of the circles correspond to those on the time color bar, and the thicknesses of the circles vary in direct proportion to the number of papers published within the time frame. The nodes with purplish-red apertures have high betweenness centrality (The betweenness centrality refers to the number of intermediate nodes that a node acts as the shortest path between the other two nodes. It is an index to measure the importance of a point in the network [19]), the links between nodes represent the co-occurrence relations between pairs of subjects, the thicknesses of the connecting lines are proportional to the number of co-occurrences of pairs of jointly appearing subjects, and the colors of the links correspond to the color of the time bar when both subjects occur together for the first time.

The figure shows that there are universal relationships and coordination between the subjects. The diversity of these subjects suggests that the study of geographic information ontology is not limited to geographical science and computer-related specialties; instead, it extends to other related subjects. Figure 2 is divided into approximately four clusters, of which Cluster #3 is the largest. This cluster consists of six main subjects, the published articles of which are mainly modeling and data processing-related studies. The published articles in Cluster #1 are mainly related to studies of cyber-infrastructure. The published articles in Cluster #2 have economics and energy as their topics. Finally, the main areas of study in Cluster #4 are geography and computer science. The subjects in these four clusters are interconnected, implying that the subjects become connected to one another during development, followed by continuous progress through learning from other subjects.

In Figure 2, computer science (842) is the subject with the largest number of published papers. The position of computer science at the center of the graph suggests that it is closely related to other branches of study in geographic information ontology; thus, it is one of the most important subjects in this field. Establishing geographic information ontologies is intended to permit mutual understanding between computers and humans or between computer systems; intelligent man–machine interaction; interoperation between computer systems; and knowledge expression, sharing and reusing in computer systems [14,28]. In recent years, geographic information ontology has been introduced into artificial intelligence and used in studies of deep machine learning to enable computers to serve mankind intelligently [9]. Mathematics (27) has a small number of published papers but is encircled by a very thick purple ring, indicating that this subject has high centrality. Moreover, the mathematics node is the critical node connecting Clusters #2 and #3 (geoscience, economics, geology, and remote sensing), suggesting that the study of geographic information ontology is no exception to the rule that any natural science must become pragmatic using mathematical methods [30]. As part of our study of geographic information ontology, we perform a semantic analysis first in order to permit mathematical explanations to formalize language in the database. The present content of ontology studies shows insufficient detail and a low degree of formalization, and a majority of these studies only provide a natural language in the terminology or definition databases of quasi-natural language instead of realizing interactions between computers and humans or between computers. Additional future mathematical studies of geographic information ontology are required [2,16].
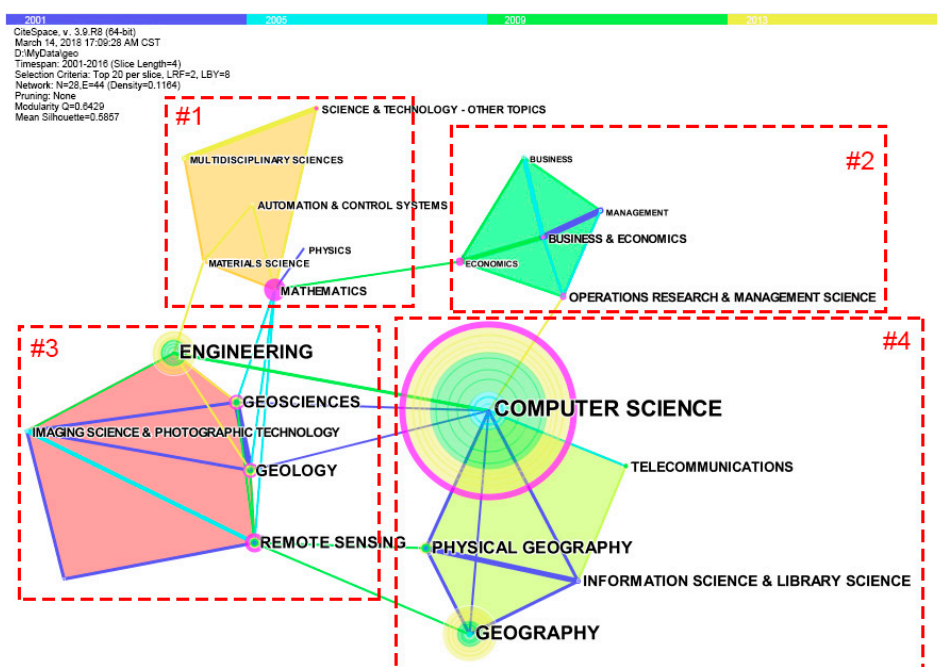
Figure 2. Cluster network of the co-occurrence of subject categories with 28 categories and 44 links.

*3.3. Characteristics of Journal Co-Citation*

As mentioned above, co-occurrence clustering analysis of subject categories depends on the subject categories of the journals in which the papers were published. The same statement holds for journal co-citation analysis, which depends on the journals that published the cited papers. Journal co-citation refers to the phenomenon in which two journals are cited in one paper, reflecting the connectivity between journals and disciplines. The knowledge base distribution of a field can be revealed through journal co-citation [31,32].

We divide the time from 2001 to 2016 into four periods. In each period, the top 20 journals cited with the greatest frequency by papers about geographic information ontology are selected for analysis, and the Pathfinder algorithm is used for pruning. Moreover, we have checked the corresponding cited documents for the cited journals in the network. If the citation rates of a journal in the network are high in only one or two articles, we will delete this journal information. After pruning, the network includes forty nodes that represent journals with the highest citation frequencies and fifty-four links representing co-citations. Next, on the basis of the topic, we perform a clustering analysis of the citing papers from the cited journals. The results are shown in Figure 3. This figure contains three obvious clusters. (a) In Cluster #1, the citing articles are mainly papers concerned with geographic information ontology and computers and include frontier terms such as "spatial", "systems", "modeling", and "semantic". (b) In Cluster #2, the citing articles primarily describe research into geographic information ontology and engineering and include frontier terms such as "knowledge acquisition", "artificial intelligence", "fuzzy", and "flexible". (c) Cluster #3 is the largest cluster and contains eleven journals, in which the citing articles focus on geographic information ontology, geography, philosophy, and other related fields of study. Here, the frontier terms are "geography", "difference", and "post-Marxism". The top three journals with the highest referring frequency are the International Journal of Geographical Information Science, Transactions in GIS, and Knowledge Acquisition (Table 1). These journals are the main sources of the cited documents and represent the potential journals that may accept contributions related to geographic information ontology.

Table 1. Frequency distribution of the top 15 cited journals with their between centrality values.

Frequency	Centrality	Journal	Impact Factor (IF) in 2016	Country	SC	Cluster
342	0.91	International Journal of Geographical Information Science	2.502	UK	CS;G1;PG;IS&LS	1
176	0.09	Transactions In GIS	2.252	USA	G1	1
162	0.15	Knowledge Acquisition		UK	CS;IS&LS	2
156	0	International Journal of Human-Computer Studies	2.863	UK	CS;E;P	2
156	0.13	Communications of the ACM	4.027	USA	CS	2
154	0.71	Annals of The Association of American Geographers	2.799	UK	G1	3
146	0.09	Geoinformatica	2.392	The Netherlands	CS;PG	1
145	0.04	Environment and Planning D-Society & Space	2.031	UK	ES&E;G1	3
144	0	Progress in Human Geography	5.776	UK	G1	3
144	0.04	Transactions of The Institute of British Geographers	2.727	USA	G1	3
139	0.47	Environment and Planning A	1.389	USA	ES&E;G	3
135	0	Computers & Geosciences	2.533	UK	CS;G2	1
125	0.02	Data & Knowledge Engineering	1.694	The Netherlands	CS	2
119	0.13	Antipode	2.413	USA	G1	3
112	0	Geoforum	2.067	UK	G1	3

Source: Own elaboration based on data from the WoS and Journal Citation Reports 2016. SC: subject category; G1: geography; PG: physical geography; IS&LS: information science and library science; G&L: government & law; ES&E: environmental sciences and ecology; E: engineering; P: psychology; G2: geology; CS: computer science.
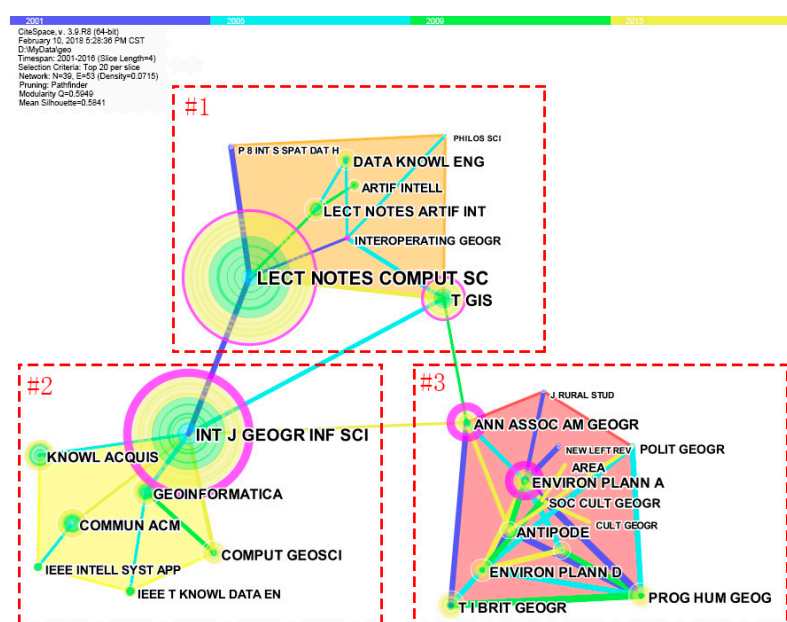
Figure 3. Cluster network of journal co-citations for 39 documents and 53 links.

More importantly, some of the nodes in Figure 3 are surrounded by purple rings, indicating that the cited publications have high betweenness centrality. Moreover, the degree of centrality is proportional to the width of the purple ring. These journals have strong co-citation relationships with others in the co-citation network (Table 1). The International Journal of Geographical Information Science has the highest centrality (0.91), followed by the Annals of the Association of American Geographers (0.71), Environment and Planning A (0.47), Knowledge Acquisition (0.15), Communications of the ACM (0.13), and Antipode (0.13).

3.4. Characteristics of Author Co-Citation

Author co-citation refers to the phenomenon in which two authors are co-cited in other documents. By computing author co-citation relationships, the interconnections between academic communities and authors within a research field can be revealed [33–36]. Author co-citation analysis has now become a potentially prolific analysis method that can be used to uncover the current status of scientific structures as well as their changes, and it can also be used to carry out frontier analyses, domain analyses, and scientific assessments. Figure 4 is a network clustering mapping that was pruned using the Pathfinder algorithm. This network contains the twenty most-cited authors in each time slice (the time from 2001 to 2016 is divided into four equal time slices) and consists of 53 co-citation nodes and 60 links.

In Figure 4, each node corresponds to an author. The sizes of the nodes represent the total numbers of citations received by the authors. Attention should be paid on the color of node center which tells us the earliest time this author was cited. The ring around the node reflects the history of his/her citation, and the color of the ring is consistent with its time partition spectrum. For example, in Figure 4, the earliest time that GUARINO N was cited by articles of geographic information ontology in WoS between 2001 and 2016 is 2001 (the color of node center is blue). Since then, there has been a stable growth in his citations (the color of ring enclosing the node from sky-blue, green, to yellow). We not only observe the citation history of earlier published articles but also can see those of recently published articles by observing the colors. The outermost purple rings enclosing the nodes indicate betweenness centrality, which means that the scholar is key in connecting different academic clusters. Hence, frequently cited authors may not have high betweenness centrality. If an author with betweenness centrality has been cited many times, he (or she) may have had a fundamental influence on the development of geographic information ontology research. In addition, the red inner rings surrounding some nodes indicate that the number of citations of by those authors' papers changed (increased or decreased suddenly) over a short period of time. These changes are usually associated with fundamental changes in the study of geographic information ontology.

With respect to the content of the citing articles shown in Figure 4, Cluster #1 focuses on applications of geographic information ontology in urban studies; Cluster #2 studies the theory of geographic information ontology in terms of philosophy; Cluster #3 is primarily composed of semantic normalization studies; Cluster #4 includes several different topics, such as the recognition of geographic information ontologies and artificial intelligence (AI) ontology; and Cluster #5 concentrates on spatial analysis, semantic analysis, etc.

Figure 4 shows that Gruber TR and Guarino have the highest numbers of citations as well as high centrality (Table 2). According to [37], in some sense, an ontology is a detailed software specification [37]. The first widely accepted definition of an ontology, which is due to [38], states that it is an explicit declaration of a conceptualization. On this basis, in 1995, Gruber gave a further definition, namely, that an ontology is a formal and explicit declaration of a conceptual model shared within a domain [13]. He also listed five principles of ontology design: clarity and objectivity, uniformity, extendibility, minimum code deviation, and minimum ontology commitment. Subsequently, in 2008, Gruber attempted to comprehensively solve applied problems in social information and semantic information using collective knowledge systems [39].

In 1995, Guarino added that an ontology is a logical language model. In addition, in order to differentiate this theory from others, Guarino usually uses the word "ontology" to mean the "ontology of philosophy" [40]. Contrary to epistemology, which studies the origin and essence of human knowledge, his ontology theory focuses on existence [41,42] proposed that the concept classification of ontology should be based on the level of specification and domain dependence. To better classify ontology [43] studied the concept classification of ontology thoroughly and carefully. Through analysis of the nature of concepts, the properties of concepts, and relations among concepts, he provided a theory guiding concept classification [43]. In the opinion of Guarino, the differences among ontological concepts exist in the definition of concepts as well as their properties. Therefore, he provided further explanation of the consistency and unity of ontological concepts and concentrated on the dependence concept of ontology, which is applied to attributes [44,45].

Cluster #4 occurs in the center of Figure 4 and stands for studies by scholars with significant impacts on geographic information ontology. Apart from Guarino, Smith B is also surrounded by a thick purple ring (the highest centrality from Table 2) with red dots in the center, suggesting that his studies have made important contributions to the development of geographic information ontology. Smith B collaborated with Mark DM in 1998, and they published a paper titled "Ontology and geographic kinds". Specific reasons for carrying out geographic information ontology-related research are proposed in that paper, thus marking the official position of information ontology in geographic information science [24]. They also co-led a project titled "Geographic classification: ontology investigation" in 1999. In this project, they attempted to clarify geographic objects and ontologies associated with cognitive classification. They therefore carried out a systematic study into different categories of geographic information ontology using a questionnaire, and they performed stress tests on human cognition of some humanistic concepts in different language environments [26]. In 2004, Smith defined the SNAP and SPAN ontologies to distinguish points and periods of time and applied this concept to the representation of processes and spatio-temporal reasoning in spatio-temporal ontologies [46].

Kuhn is also a researcher who investigates who investigates geographic information ontology. As early as 1993, he proposed that spatial information theory should be oriented to GIS users [47]. In regard to the method of construction of geographic information ontology, Kuhn designed and implemented a construction method based on natural language. By analyzing keywords and their semantic implications in German transport norms and standards, he then extracted an ontology that could be applied to automobile navigation from traffic data [4,48]. Moreover, he also proposed the concept of semantic reference systems [49,50]. He then continued his work to apply semantic reference systems to semantic transformations among instance objects in different land management systems using the theory and methods of Conceptual Space [51]. In 2005, he suggested that semantic interactions between geographic information services could be divided into two categories, specifically (1) defining the semantic connotations of metadata corresponding to every service interface for querying and searching semantic information, and (2) defining a mathematical model of semantic interoperation for the expression of semantic knowledge and reasoning using that knowledge [52].

According to [53], ontology can be used in information systems to avoid problems including potential inconsistencies among GISs. Geographic information systems contain information about objects and their attributes at specific locations, and these attributes range from surface elevation to natural land cover. Both natural and artificial elements are involved, which suggests a need for different consistency rules. As a consequence, ontology is necessary for element classification to achieve consistency. Based on the above points, Frank proposed a five-layer classification structure that applies to both GISs and other systems related to large-scale geographic entities. In this classification structure, the top-level ontology is a type of natural entity that is independent of the human mind, whereas the others are associated with epistemology, and different rules correspond to different layers [53].

According to [14], ontology is a theory that describes entities, concepts and the relationship between their features and relative functions using specific words from a specific point of view. The authors of [14,54,55] performed many studies in ontology-based geographic information integration, and they

proposed the idea of ontology-driven geographic information systems (ODGISs) using the theory of ontology-driven information systems from Guarino. According to a paper by Fonseca, in different information processing periods, real things processed using information ontology can be divided into four spatial paradigms. In addition, with respect to interoperation, that paper mainly discusses ontology applications in the design and realization of information systems. Subsequently, in 2002, on the basis of the four proposed spatial paradigms described above and the features and reality of geographic information science, Fonseca proposed four new spatial paradigms (physical space, logical space, expression space, and recognition space) to analyze the relationships among the human mind, information systems, and the real world on different levels and from different perspectives [54,55].

In addition, Rovert G and his colleagues have made great contributions to the research of ontology. In the project aiming at construction of semantic network data about earth environment ontology conducted by the JPL (Jet Propulsion Lab, in Pasadena, CA, USA) of NASA, according to related concepts in the GCMD (Global Change Master Directory), JPL developed SWEET, which contains thousands of notions related to Earth and space science and definitions of relationships of these notions under the guidance of ISO1911X, and they published SWEET on the internet using OWL (Ontology Web Language) [56]. Pier Luigi Buttigieg and his colleagues have achieved a lot in environmental ontology research; they have summarized the content and structure of environmental ontology [57]. Their research results also contribute to the research of the whole geospatial ontology.

With respect to the topic of citing articles, the citing articles of Cluster #4 mostly focus on subjects such as ontology recognition and AI as well as ontological theory. In addition, this cluster contains many citing articles that show great concern for the semantic-based integration and interoperation of geographic information from a real-world perspective. At present, the rapid growth of AI and the demands of individual users open a new door for geographic information ontology research. Organizing and expressing data on the basis of geographic information ontology to satisfy individual demand will be an important direction of geographic information ontology research [3].

Table 2. The profile of the 15 most co-cited authors.

Frequency	Centrality	Author	ClusterID
242	0.11	Gruber TR	3
194	0.38	Guarino N	4
152	0.65	Smith B	4
129	0.29	Fonseca F	4
102	0.15	Kuhn W	4
98	0.08	Egenhofer MJ	4
83	0	Berners-lee T	3
76	0	Baader F	3
77	0	Bishr Y	4
70	0	Uschold M	3
76	0	Mark DM	4
62	0.06	Horrocks I	3
62	0.13	Frank AU	4
61	0.13	Lutz M	3
60	0.06	Klien E	3

From the above analysis, geographic information ontology studies ontological cognition and ontology concept primitively. With the development of Internet, more and more information systems appear on the Internet. Data integration and interoperation between multi-source heterogeneous information systems become an important question to be solved. However, both data integration and interoperation are not simply doing connection among databases, any research related to information sharing and interoperation could be traced to semantic heterogeneity [5,52]. Specifically, in Geographic Information Science, researchers often use ontology to solve this problem. The semantic web based on ontology makes internet information more specific and semantic-rich, enabling computer to do analysis

and reasoning for intellectualization of network data and network service processing. Then, our Internet will be a universal and powerful information integration and exchange platform [2,3,28].
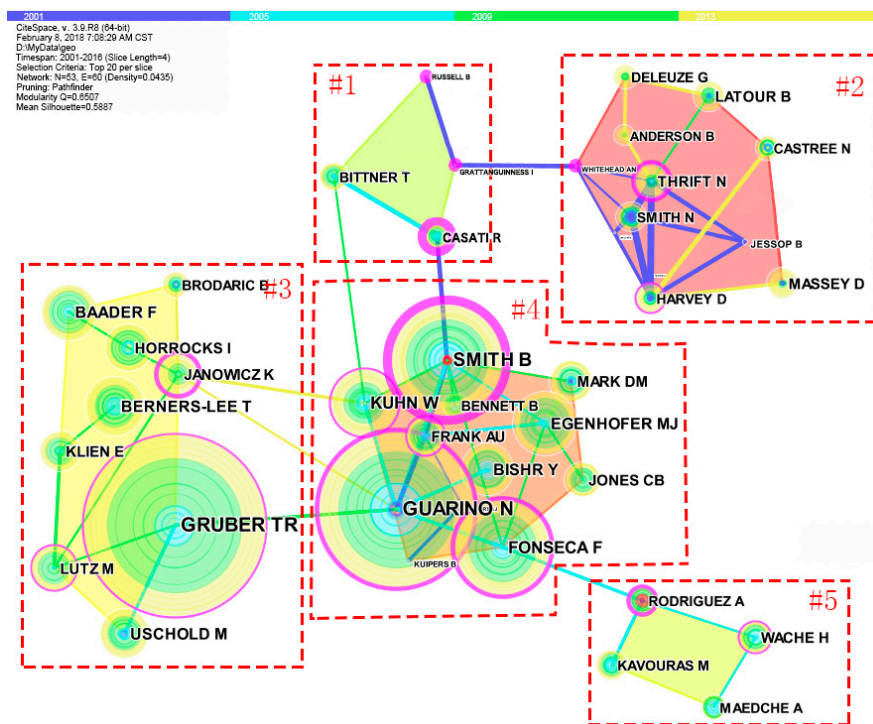

Figure 4. Cluster network of author co-citations for 53 documents and 60 links.

4. Discussion

Ontology had dominated western philosophy for many years, and it is still heatedly debated by philosophers today. However, ontology is no longer considered only by philosophers, as it has been extensively applied to other disciplines such as natural language processing, biological sciences, and medicine, particularly in the field of information science, such as artificial intelligence, knowledge engineering, and geographic information science.

Introducing ontology to the field of geographic information science mainly reflects the shift in focus from the previous formalization of calculation models to the space target domain. Due to the complexity of space information, ontology research in geographic information science is a highly interdisciplinary research. As stated in Section 3.1, ontology was studied as a prominent research subject in the field of geographic information science around 2000. Afterward, from 2001–2009, attention was paid to geographic information ontology, which was obviously a growing trend. Based on the analysis of cited authors in Section 3.4, we observe that ontology in geographic information science primarily includes three relatively large clusters that respectively correspond to the contents at three levels in geographic information ontology research: (i) the study of concepts involved in establishing sound and elaborate geographic information ontology [24,27,43]; (ii) the study of methods and instruments that can be used in geographic information ontology formalization to make ontology conveniently shared and utilized [4,11,48]; and (iii) the study of transition among different ontologies, integration of different ontologies, ontology-driven geographic information system development, ontology-based geographic information integration and interoperability, and so on [5,54,55]. Around 2009, as diverse strategies such as Smart Earth emerged, geographic information was demanded to be shared across regions, countries, and even the whole world. However, the conceptual understanding of geography varies by country or region, which obstructs information sharing and leads to substantial repetitive work and resource waste. The introduction of ontology could potentially solve these problems [12].

At present, along with the development of mass social media and artificial intelligence, geographic information ontology could play a role in transforming data into knowledge. The transformation of data into knowledge contributes to spatial data mining and emergency response analysis of isomers data [2]. In the future, it is feasible to apply geographic information ontology technology into the network environment, establish an ontology-driven and intelligent geographic information system, and provide intelligent and personalized geographic information services for users [3,10].

5. Conclusions

Based on the 1533 articles collected from the WoS database from 2001 to 2016, this paper performs co-citation and cluster analysis with respect to research area, discipline, journal, author, and so on. We attempt to identify the key features of global geographic information ontology research, and aim to help researchers, especially newcomers, provide a reference for future research.

Research on geographic information ontology dates back to the 1990s. After 2001, it showed a continuation in its upward trend before peaking in 2009. The period from 2010 to 2016 was a relatively stable stage in the development of this field of study.

Many disciplines have been involved in geographic information ontology research. Among these disciplines, computer science has contributed the largest number of published papers, whereas mathematics has provided theoretical support.

The most-cited journals are from the US and Europe. Moreover, among those journals, the International Journal of Geographical Information Science and the Annals of the Association of American Geographers are the main knowledge sources for geographic information ontology research.

Gruber TR, Guarino N, and Smith B are the top scholars in terms of citation frequency. Of these, Smith B has the highest centrality, and both Gruber TR and Guarino N proposed widely accepted definitions of ontology. Smith B provided specific reasons for carrying out geographic information ontology research and formally introduced information ontology to geographic information science. The construction of an intelligent geographic information system in a network environment to achieve geographic information sharing and to provide users with intelligent and personalized geographic information service would be an important future direction of geographic information ontology research.

Author Contributions: This article was mainly conducted by Yu Liu and Lin Li. Lin Li supervised and guided this study, Yu Liu designed the proposal of the experiments and conducted the analyses, and Hang Shen joined in the design of study and conducted the experiments and analyses. Hui Yang and Feng Luo collected and analyzed the data; Yu Liu wrote the paper. All authors have read and approved the final manuscript.

References

1.   Deng, Z.; Tang, S.; Zhang, M.; Yang, D.; Chen, J. Overview of ontology. *Acta Sci. Nat. Univ. Pekin.* **2002**, *38*, 730–738.
2.   Li, B.; Liu, B.; Liu, J.; Dong, C. The research progress review and prospect on geo-ontology. *Sci. Surv. Map.* **2015**, *40*, 53–57.
3.   Miao, J.; Wang, J.; Cheng, Y.; Ma, M.; Lu, J. Overview and trend analysis of geo-ontology. *J. Geom. Sci. Technol.* **2014**, *31*, 653–658.
4.   Werner, K. Ontologies in support of activities in geographical space. *Int. J. Geogr. Infor. Sci.* **2001**, *15*, 613–631.
5.   Huang, M. Study on Formal Representation of Geographic Ontology and Its Application in Map Services. Doctor's Dissertation, Wuhan University, Wuhan, China, 2005.
6.   Hillen, F.; Hofle, B. Geo-recaptcha: Crowdsourcing large amounts of geographic information from earth observation data. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *40*, 29–38. [CrossRef]

7. Lukinbeal, C.; Monk, J.J. Master's in geographic information systems programs in the united states: Professional education in gis and geography. *Prof. Geogr.* **2015**, *67*, 482–489. [CrossRef]
8. Mark, D.M.; Skupin, A.; Smith, B. *Features, Objects, and Other Things: Ontological Distinctions in the Geographic Domain*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 489–502.
9. Li, L.; Liu, Y.; Zhu, H.; Ying, S.; Luo, Q.; Luo, H.; Xi, K.; Xia, H.; Shen, H. A bibliometric and visual analysis of global geo-ontology research. *Comput. Geosci.* **2017**, *99*, 1–8. [CrossRef]
10. Wang, J.; Worboys, M. Ontologies and representation spaces for sketch map interpretation. *Int. J. Geogr. Infor. Sci.* **2017**, *31*, 1697–1721. [CrossRef]
11. Kavouras, M.; Kokla, M. A method for the formalization and integration of geographical categorizations. *Int. J. Geogr. Infor. Sci.* **2002**, *16*, 439–453. [CrossRef]
12. Hu, L.; Li, L.; Wang, H. Methodology of semantic integration of gis based on category of ontology. *Sci. Surv. Map.* **2011**, *36*, 52–55.
13. Gruber, T.R. Toward principles for the design of ontologies used for knowledge sharing. *Int. J. Hum.-Comput. Stud.* **1995**, *43*, 907–928. [CrossRef]
14. Fonseca, F.T. Ontology-driven geographic information systems. *ACM SAGIS* **2001**, *71*, 14–19.
15. Bornmann, L. How are excellent (highly cited) papers defined in bibliometrics? A quantitative analysis of the literature. *Res. Eval.* **2014**, *23*, 166–173. [CrossRef]
16. Madani, F. *'Technology Mining' Bibliometrics Analysis: Applying Network Analysis and Cluster Analysis*; Springer: New York, NY, USA, 2015; pp. 1–13.
17. Harzing, A.W.; Alakangas, S. Google scholar, scopus and the web of science: A longitudinal and cross-disciplinary comparison. *Scientometrics* **2016**, *106*, 787–804. [CrossRef]
18. Li, J.; Chen, C. *Citespace: Text Mining and Visualization in Scientific Literature*; Capital University of Economics and Business Press: Beijing, China, 2016.
19. Chen, Y.; Chen, C.; Hu, Z.; Wang, X. *Principles and Applications of Analyzing a Citation Space*; Science Press: Beijing, China, 2014.
20. Liu, Z.; Yin, Y.; Liu, W.; Dunford, M. Visualizing the intellectual structure and evolution of innovation systems research: A bibliometric analysis. *Scientometrics* **2015**, *103*, 135–158. [CrossRef]
21. Wu, X.; Chen, X.; Zhan, F.B.; Hong, S. Global research trends in landslides during 1991–2014: A bibliometric analysis. *Landslides* **2015**, *12*, 1215–1226. [CrossRef]
22. Mall, R.; Langone, R.; Suykens, J.A.K. Kernel spectral clustering for big data networks. *Entropy* **2013**, *15*, 1567–1586. [CrossRef]
23. Frias-Martinez, V.; Frias-Martinez, E. Spectral clustering for sensing urban land use using twitter activity. *Eng. Appl. Artif. Intell.* **2014**, *35*, 237–245. [CrossRef]
24. Smith, B.; Mark, D.M. Ontology and Geographic Kinds. In Proceedings of the International Symposium on Spatial Data Handling, Vancouver, BC, Canada, 12–15 July 1998.
25. Fonseca, F.T.; Egenhofer, M.J. Ontology-driven geographic information systems. In Proceedings of the International Symposium on Advances in Geographic Information Systems (Acm-Gis '99), Kansas City, MO, USA, 2–6 November 1999; pp. 14–19.
26. Smith, B.; Mark, D.M. Ontology with human subjects testing: An empirical investigation of geographic categories. *Am. J. Econ. Soc.* **1999**, *58*, 245–272. [CrossRef]
27. Mark, D.M.; Smith, B.; Tversky, B. *Ontology and Geographic Objects: An Empirical Study of Cognitive Categorization*; Springer: Berlin/Heidelberg, Germany, 1999; pp. 283–298.
28. Janowicz, K. Observation-driven geo-ontology engineering. *Trans. GIS* **2012**, *16*, 351–374. [CrossRef]
29. Palmer, A.L.; Sesé, A.; Montaño, J.J. Tourism and statistics: Bibliometric study 1998–2002. *Ann. Tour. Res.* **2005**, *32*, 167–178. [CrossRef]
30. Si, S.; Sun, Z.; Sun, X. *Mathematical Modeling Algorithms and Applications*; National Defend Industry Press: Beijing, China, 2015.
31. McCain, K.W. Mapping economics through the journal literature: An experiment in journal cocitation analysis. *J. Am. Soc. Inf. Sci.* **1991**, *42*, 290–296. [CrossRef]
32. Leydesdorff, L. Can scientific journals be classified in terms of aggregated journal-journal citation relations using the journal citation reports? *J. Am. Soc. Inf. Sci. Technol.* **2006**, *57*, 601–613. [CrossRef]
33. Small, H. Co-citation in the scientific literature: A new measure of the relationship between two documents. *J. Am. Soc. Inf. Sci.* **1973**, *24*, 265–269. [CrossRef]

34. McCain, K.W. Longitudinal author cocitation mapping—The changing structure of macroeconomics. *J. Am. Soc. Inf. Sci.* **1984**, *35*, 351–359. [CrossRef]

35. He, Y.L.; Hui, S.C. Mining a web citation database for author co-citation analysis. *Inf. Process. Manag.* **2002**, *38*, 491–508. [CrossRef]

36. Hsiao, C.H.; Yang, C. The intellectual development of the technology acceptance model: A co-citation analysis. *Int. J. Inf. Manag.* **2011**, *31*, 128–136. [CrossRef]

37. Gruber, T.R. The role of common ontology in achieving sharable, reusable knowledge bases. In *International Conference on Principles of Knowledge Representation and Reasoning*; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1991; pp. 601–602.

38. Gruber, T.R. A translation approach to portable ontology specifications. *Knowl. Acquis.* **1993**, *5*, 199–220. [CrossRef]

39. Gruber, T. *Collective Knowledge Systems: Where the Social Web Meets the Semantic Web*; Elsevier Science Publishers B.V.: Amsterdam, The Netherlands, 2008; pp. 4–13.

40. Guarino, N. Formal ontology, conceptual analysis and knowledge representation. *Int. J. Hum.-Comput. Stud.* **1995**, *43*, 625–640. [CrossRef]

41. Guarino, N. Semantic matching: Formal ontological distinctions for information organization, extraction, and integration. In *International Summer School on Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology*; Springer: Berlin/Heidelberg, Germany, 1997; pp. 139–170.

42. Guarino, N. Some organizing principles for a unified top-level ontology. In Proceedings of the AAAI Spring Symposium on Ontological Engineering, Stanford, CA, USA, 24–26 March 1997; pp. 57–63.

43. Guarino, N. *Formal Ontology in Information Systems: Proceedings of the 1st International Conference, Trento, Italy, 6–8 June 1998*; Lightning Source Inc.: La Vergne, TN, USA, 1998; pp. 3–15.

44. Guarino, N.; Welty, C. Identity, unity, and individuality: Towards a formal toolkit for ontological analysis. In Proceedings of the 14th European Conference on Artificial Intelligence, Berlin, Germany, 20–25 August 2000.

45. Guarino, N.; Welty, C. A formal ontology of properties. In Proceedings of the European Workshop on Knowledge Acquisition, Modeling and Management, London, UK, 2–6 October 2000; pp. 97–112.

46. Grenon, P.; Smith, B. Snap and span: Towards dynamic spatial ontology. *Spat. Cogn. Comput.* **2004**, *4*, 69–104. [CrossRef]

47. Kuhn, W. *Metaphors Create Theories for Users*; Springer: Berlin/Heidelberg, Germany, 1993; pp. 366–376.

48. Werner, K. Modeling the semantics of geographic categories through conceptual integration. In Proceedings of the International Conference on Geographic Information Science, Boulder, CO, USA, 25–28 September 2002; pp. 108–118.

49. Kuhn, W. Semantic reference systems. *Int. J. Geogr. Inf. Sci.* **2003**, *17*, 405–409. [CrossRef]

50. Kuhn, W.; Raubal, M. Implementing semantic reference systems. In Proceedings of the 6th AGILE, Lyon, France, 24–26 April 2003.

51. Baglatzi, A.; Kuhn, W. *On the Formulation of Conceptual Spaces for Land Cover Classification Systems*; Springer International Publishing: Cham, Switzerland, 2013; pp. 173–188.

52. Kuhn, W. *Geospatial Semantics: Why, of What, and How?* Springer: Berlin/Heidelberg, Germany, 2005; Volume 3, pp. 1–24.

53. Frank, A.U. *Spatial Ontology: A Geographical Point of View*; Springer: Amsterdam, The Netherlands, 1997; pp. 135–153.

54. Fonseca, F.T.; Egenhofer, M.J.; Agouris, P.; Câmara, G. Using ontologies for integrated geographic information systems. *Trans. GIS* **2002**, *6*, 231–257. [CrossRef]

55. Fonseca, F.; Davis, C.; Davis, C. Semantic granularity in ontology-driven geographic information systems. *Ann. Math. Artif. Intell.* **2002**, *36*, 121–151. [CrossRef]

56. Raskin, R.G.; Pan, M.J. Knowledge representation in the semantic web for earth and environmental terminology (sweet). *Comput. Geosci.* **2005**, *31*, 1119–1125. [CrossRef]

57. Buttigieg, P.L.; Morrison, N.; Smith, B.; Mungall, C.J.; Lewis, S.E.; Consortium, T.E. The environment ontology: Contextualising biological and biomedical entities. *J. Biomed. Semant.* **2013**, *4*, 1–9. [CrossRef] [PubMed]