

# CLOUD COMPUTING IN THE PUBLIC SECTOR: MAPPING THE KNOWLEDGE DOMAIN

Mark Theby

Capitol Technology University, Laurel, Maryland, USA

## **ABSTRACT**

*Cloud computing is a key element in many nations' pursuit of fast-tracked digital transformation and the quick implementation of digital tools but is still facing considerable barriers due to the distinct challenges that information technology adoption faces in public sector environments. Using scientometric data from the Web of Science database, this study explores the current state of research and the structure of the public sector cloud computing knowledge domain in a novel way, utilizing the CiteSpace visual analytic software to produce knowledge maps that visualize public sector cloud computing research in terms of publication activity, constituent authors, and publication venues, as well as exploring the intellectual base of the knowledge domain. For public sector cloud computing researchers and practitioners, the study provides visual insights and analyses that support future research, collaboration, and evidence-based cloud computing implementation and utilization.*

## **KEYWORDS**

*Cloud Computing, Government, Public Sector, Knowledge Mapping, CiteSpace.*

## **1. INTRODUCTION**

The concept of cloud computing alludes to the on-demand, convenient, and ubiquitous delivery of a wide range of pooled and configurable computing resources—such as applications, servers, storage, and development tools—via a computer network and has received increasing attention from a variety of organizations [1]. The cloud computing technology model has been widely embraced by the private sector, where it is providing extensive benefits such as reduced IT costs, greater agility and time-to-value, and improved scalability compared to legacy approaches for computing resource delivery [2], but has experienced a less robust adaption in the public sector [3].

For public sector organizations, the migration of IT resources and processes to the cloud can confer distinct benefits. Cloud computing can lower the upfront costs traditionally associated with resource-intensive computing and provides rapid access to hardware resources without requiring extensive capital investments [4]. Cloud computing can also enhance agility by enabling organizations to dynamically up- or down-scale their services dependent on operational requirements, making it feasible to meet sudden peak requirements without having to maintain costly slack resources [5]. By enabling public sector organizations to access state-of-the-art IT services at the same speed as the commercial world, and by supporting the rapid prototyping and development of tools and services, a transition to the cloud can also foster innovation [6]. Finally, as cloud users and cloud applications inherit accreditation controls from their central cloud platforms, the adoption of cloud computing can also increase the security of IT assets [7].

The distinct, potential benefits of cloud computing in the public sector only gained further importance during the COVID-19 pandemic. Experiencing extraordinary challenges, governments and public sector agencies were forced to mount a rapid response to ensure continuity of operations, protect citizens, and address the economic and social fallout of the pandemic. Public sector work arrangements shifted, with employees being thrust into remote work environments in record numbers, government services flexed radically to meet emergency demands, and critical new services were implemented to address essential government operations and policymaking needs. For government organizations that had cloud capabilities available, pivotal cloud advantages such as agility, scalability, and resiliency provided the foundation for a successful crisis response; for public sector organizations that lagged behind in cloud adoption, the crisis has further strengthened the case for expanded cloud computing [8].

Because of the unique characteristics of the public sector, the application of “lessons learned” from the private sector is difficult or unviable. Government organizations pursue matters of public interests that significantly diverge from those of private interests; they are accountable to citizens and voters and not to shareholders or owners; and they operate in a legal and constitutional environment that generally requires openness, transparency, impartiality, and the pursuit of the rule of the law [9], [10]. The public sector’s distinct missions, legal requirements, and culture translate into a range of organizational and operating-model challenges that are unique and affect cloud adoption in ways not observable in the private sector [3].

Despite the importance of the topic and the public sector’s unique characteristics, there have been no attempts to date to systematically and visually explore the state of the art of the public sector cloud computing domain. This study will address this gap by applying knowledge mapping analysis techniques to the domain that will merge quantitative analysis, classification, and visualization approaches to further the discovery of the foundational composition and the current status of government cloud computing research [11].

## 2. RESEARCH METHOD AND TOOLS

### 2.1. Theoretical Underpinnings and Conceptual Definitions

**Knowledge mapping.** Aiming to visualize the organization and dynamics of scientific disciplines, fields of research, or scientific topic areas, knowledge mapping is an interdisciplinary research field that relies on a combination of approaches and methods, including scientometric and visual analytic tools, to explore how scientific knowledge is intellectually, socially, conceptually structured [11], [12]. Knowledge mapping is based on quantitative analysis and has been successfully applied in a wide variety of research arenas [13], more recently also finding use for information technology-related topics such as business digitalization [14], big data [15], e-learning [16], and artificial intelligence [17].

**Research front.** Tracing its intellectual origins back to the ground-breaking work of Price [18] and Crane [19], the notion of a research front focuses on the complex, volatile, and continuously evolving point in the research enterprise where cutting edge knowledge is created and then disseminated [20]. The front represents the up-to-date scientific understanding of the research arena as identified in the scientific literature and takes account of the contemporary research questions the scientific community is focusing on [20].

**Intellectual base.** The enduring nucleus of a body of scientific literature constitutes the intellectual base of a field of scientific inquiry [11]. This intellectual base represents the essential scientific knowledge required to successfully take on research within a particular field and is

properly identified by the citation trails that emerge from the research front and point to the underlying literature; i.e. the research front's citation and co-citation footprint [21]. The conceptual model for *CiteSpace*, a key analytic tool that serves as the application of choice for the present study, is visualized in Figure 1 and captures the relationship between research front and intellectual base.

**Betweenness centrality.** A metric that identifies how often a network node is located on the shortest path between other nodes in the network [22]. Betweenness centrality quantifies the level at which a specific network node can operate as a control point in a network. On a science map, network nodes with higher levels of betweenness centrality possess higher levels of importance vis-à-vis the overall information flow through the network [23].

**Public sector organizations.** For the purposes of this study, public sector organizations are defined as all government-controlled or government-funded agencies, enterprises, and other entities that deliver government programs, public goods, or public services [37].

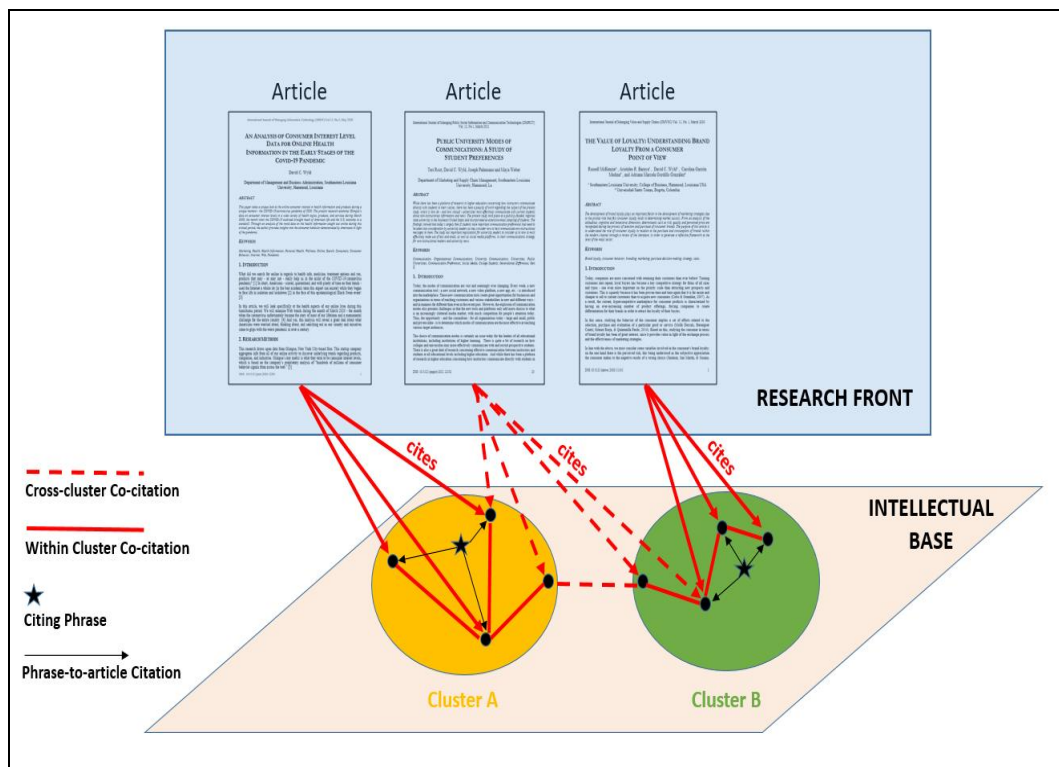


Figure 1. CiteSpace Conceptual Model [1]

## 2.2. Data Collection

The multidisciplinary core collection of the *Web of Science* (WoS) citation database by *Clarivate* served as the data source for this study. Containing the world's leading scientific journals, books, and proceedings, the core collection on the WoS search platform indexes over 21,100 peer-reviewed scholarly journals and provides robust scientometric coverage for over 250 science, social science, and arts & humanities disciplines, including engineering, computing, and technology [24]. *WoS* was selected as a data source because of its robust coverage of computing and technology topics, the confirmed validity and accuracy of its underlying data structure, and the fact that *CiteSpace* is optimized for handling *WoS* data [25].

In October of 2021, following consultation with a research librarian and several preliminary searches to determine relevant keywords, WoS was searched using the following search strategy: TI=(cloud AND (government OR "public sector" OR "public administration" OR "public service") ). The search was not limited in terms of language, publication type, or time range to enable the broadest exploration of available bibliographic records. The search returned 192 records and the titles and abstracts of the search results were individually reviewed to ensure that all records addressed public sector cloud computing. Four articles were removed from the search results as a result of this review because these four articles did not substantively address cloud computing in a public sector environment; the remaining 188 records were exported for further analysis in *CiteSpace*.

### 2.3. Analytic Tool

*CiteSpace* 5.8.R1 software was utilized to conduct a scientometric analysis of the research data obtained in WoS. *CiteSpace* is a Java application designed for the analysis and visualization of knowledge domains, is freely available [21], and has been used and validated extensively for the exploration and mapping of scientific knowledge [25]. With *CiteSpace*, the intellectual base of the public sector cloud research domain will be algorithmically and temporally detected based on emergent research front terminology; and patterns and important features will be visualized.

### 2.4. Research Method

This study is based on scientometric analysis, a quantitative approach to investigating the creation, distribution, and use of scientific information in order to aid in an improved understanding of scientific research endeavours [26]. Relying on mathematical and statistical techniques, scientometrics enables the examination of the formal attributes of scientific knowledge domains [27]. The scientometric analysis approach adopted in this study comprises network and collaboration analysis, cluster analysis for keywords and literature, and co-citation analysis.

**Cluster analysis.** Clustering organizes data objects in a network into groups that share similar features and that are dissimilar from objects in other groups [28]. Clustering reveals the underlying structures of a research field by grouping similar objects with each other and identifying the common themes amongst the objects [29]. *CiteSpace* can characterize the composition of identified clusters by extracting noun phrases from the abstracts, keywords, or titles of the objects in a particular cluster [25].

**Collaboration analysis.** The analysis of collaboration measures focuses on quantitative indicators that identify how researchers and organizations collaborate in the creation of scientific knowledge [30]. Collaboration analysis is a key approach in better understanding the scholarly communication and knowledge diffusion that takes place in a knowledge domain [21].

**Co-citation analysis.** This analysis measures the intellectual relationship between objects by examining the frequency with which these objects were co-cited by other, later objects. Co-citation measures provide valuable information on the structure of knowledge domains by identifying co-citation patterns and broadens the opportunity to identify relevant research and researchers that may not emerge in more traditional approaches to knowledge domain exploration [31].

### 3. RESEARCH RESULTS

#### 3.1. Geographic and Institutional Analysis of Publication Activity

Utilizing the *CiteSpace* node types of *Institution* and *Country*, with a g-index selection criteria scale factor of 40 and node labelling configured to 2 (Threshold), 5 (Font Size), and 200 (Node Size), a knowledge map based on country and institutional publication activity was visualized (Figure 2).

Results were then ranked, with Table 1 displaying the top five countries based on *betweenness centrality* scores. Table 2 reflects the top eight institutions based on the number of published research publications.

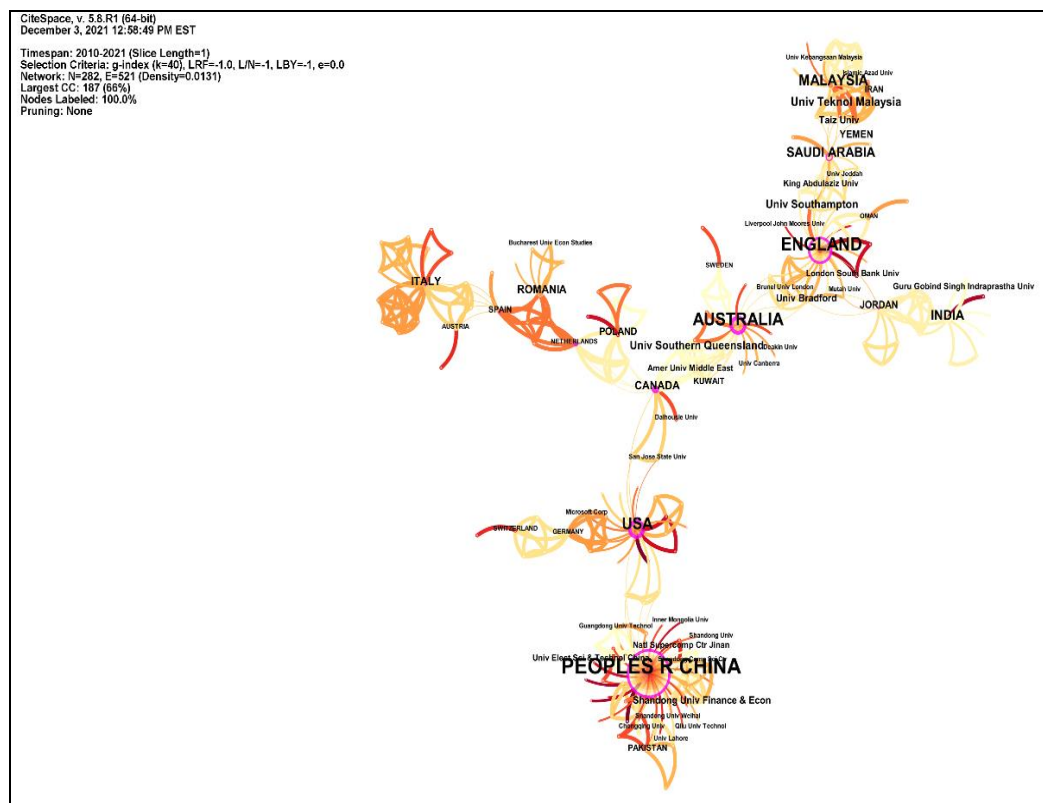


Figure 2. Knowledge Map of Publication Activity by Country and Institution

Table 1. Top 5 Countries based on Betweenness Centrality Scores

Ranking	Country	Betweenness Centrality Score	Publications
1	Canada	0.57	5
2	United States	0.47	13
3	Australia	0.43	16
4	People's Republic of China	0.36	36
5	United Kingdom	0.33	19

Table 2. Top 8 Institutions based on Number of Published Research Publications

Ranking	Institution	Publications	Country
1	Universiti Teknologi Malaysia	7	Malaysia
2	University of Southern Queensland	5	Australia
3	University of Southampton	5	United Kingdom
4	Shandong University of Finance and Economics	4	People's Republic of China
5	American University of the Middle East	4	Kuwait
6	Taiz University	4	Yemen
7	University of Bradford	4	United Kingdom
8	University of Brasilia	4	Brazil

*Ties in the number of publications were resolved based on betweenness centrality scores.*

### 3.2. Author and Author Co-Citation Analysis

Table 3 lists the top nine authors in the knowledge domain, based on the number of publications. Utilizing the *CiteSpace* node type of *Author*, with a g-index selection criteria scale factor of 40 and node labelling configured to 3 (Threshold), 11 (Font Size), and 200 (Node Size), a knowledge map focusing on author activity and collaboration was visualized (Figure 3).

The overall level of author collaboration in the knowledge domain is limited: there are no large and stable research groupings that focus on public sector cloud computing research in a sustained fashion. Nascent research clusters include Omar Ali from the American University of the Middle East and his former colleagues at the University of Southern Queensland, Fathey Mohammed and Othman Ibrahim at universities in Malaysia, Madini Alassafi from King Abdulaziz University and his former colleagues from the University of Southampton, and Gang Li from Qilu University and other colleagues from several universities in the People's Republic of China.

Table 3. Top 9 Authors based on Number of Published Research Publications

Ranking	Author	Publications	Current Affiliation
1	Omar Ali	9	American University of the Middle East
2	Fathey Mohammed	5	Universiti Utara Malaysia
3	Jeffrey Soar	5	University of Southern Queensland
4	Anup Shrestha	4	University of Southern Queensland
5	Gang Li	4	Qilu University of Technology
6	Othman Ibrahim	4	Universiti Teknologi Malaysia
7	Madini Alassafi	4	King Abdulaziz University
8	Yikai Liang	4	Shandong University of Finance and Economics
9	Jianming Yong	4	University of Southern Queensland

*Ties in the number of publications were resolved based on betweenness centrality scores.*

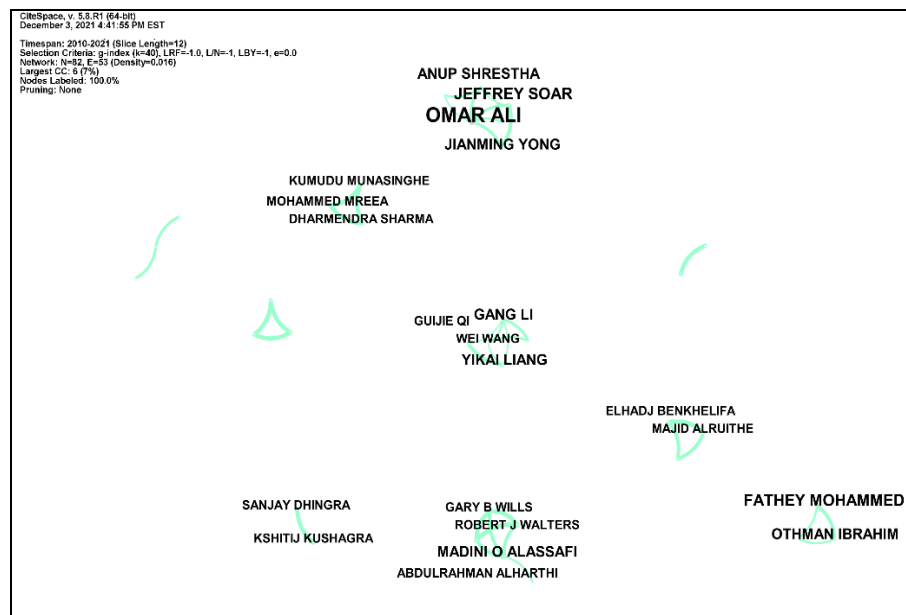


Figure 3. Knowledge Map of Author Collaboration

Next, utilizing the *CiteSpace* node type of *Cited Author*, with a g-index selection criteria scale factor of 40, pathfinder network pruning enabled, and node labelling configured to 17 (Threshold), 11 (Font Size), and 200 (Node Size), a knowledge map focusing on co-cited authors was visualized (Figure 4).

In the network of co-cited authors, Peter Mell from the National Institute of Standards and Technology (USA) evidenced the highest co-citation count (51 citations), followed by Rajkumar Buyya from the University of Melbourne (Australia) with 32 citations, and Scott Paquette from Wilfrid Laurier University (Canada) with 31 citations. Table 4 shows the top 9 co-cited authors.

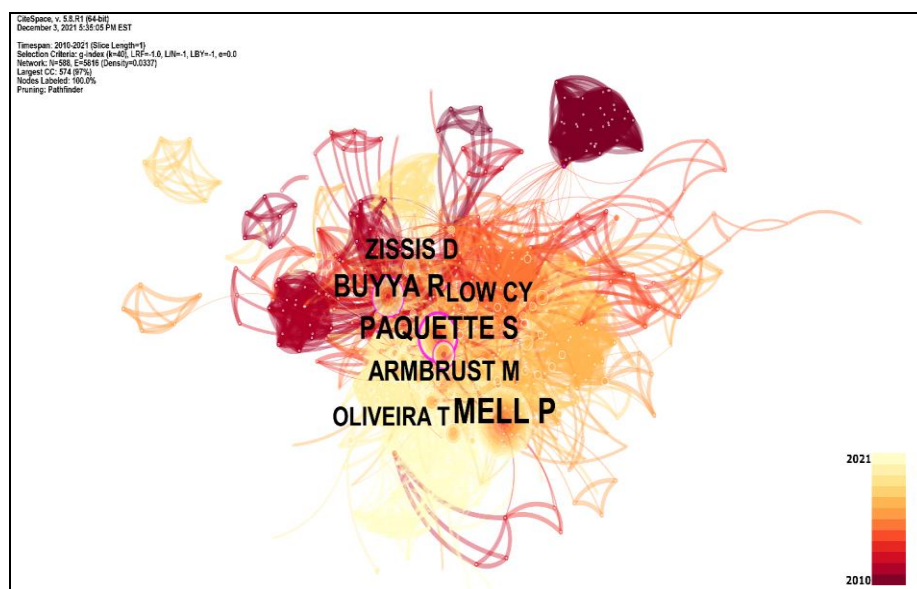


Figure 4. Knowledge Map of Author Co-Citations

Table 4. Top 9 Co-Cited Authors in the Research Domain

Ranking	Author	Co-Citations	Ranking	Author	Co-Citations
1	Peter Mell	51	6	Chinyao Low	19
2	Rajkumar Buyya	32	7	Tiago Oliveira	18
3	Scott Paquette	31	8	Sean Marston	16
4	Dimitris Zissis	22	9	Subashini Sundararajan	16
5	Michael Armbrust	22			

*Ties in the number of co-citations were resolved based on betweenness centrality scores.*

### 3.3. Journal Analysis

With the *CiteSpace* node type of *Cited Journal*, selecting a g-index scale factor of 30 and adjusting node labelling to 25 (Threshold), 4 (Font Size), and 200 (Node Size), a knowledge map for co-citations was created next (Figure 5). The top journals, ranked by their betweenness centrality scores, are listed in Table 5.

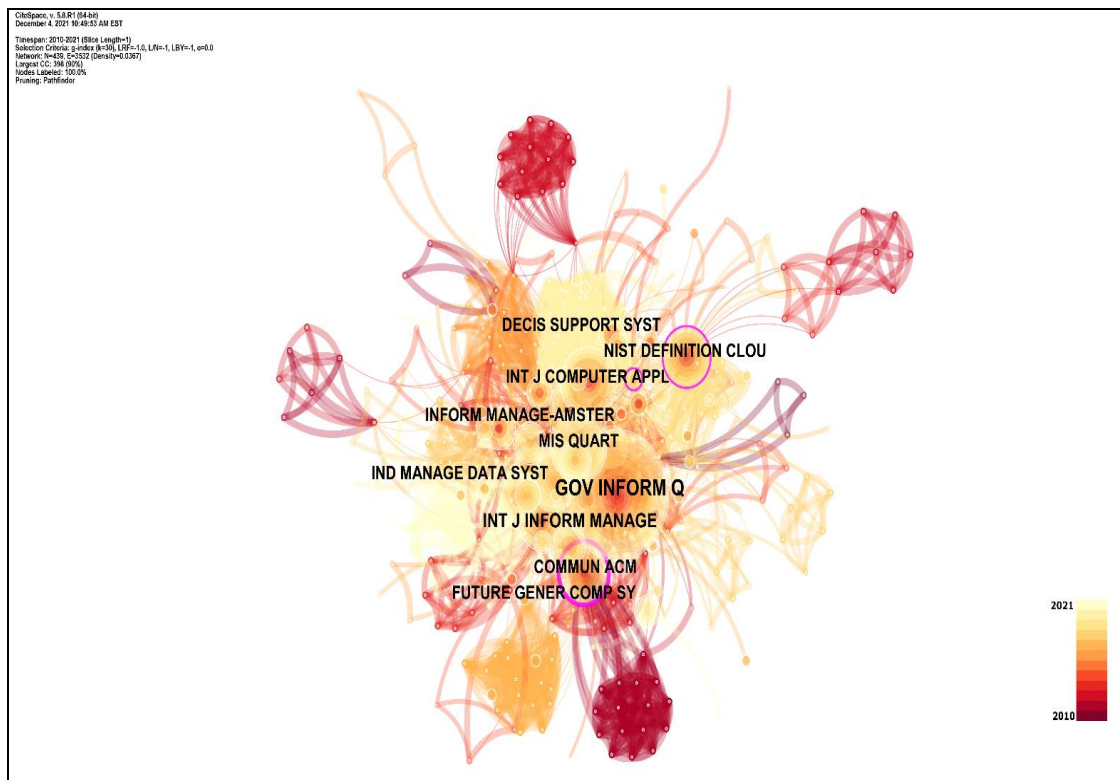


Figure 5. Knowledge Map of Journal Co-Citations

Table 5. Top 7 Co-Cited Journals/Publications Ranked by Betweenness Centrality Scores

Ranking	Co-Citations	Betweenness Centrality	Journal	Impact Factor 2020 [36]
1	33	0.29	<i>Communications of the ACM</i>	4.654
2	30	0.17	<i>NIST Special Publication</i>	N/A
3	18	0.13	<i>Journal of Network and Computer Applications</i>	6.281
4	9	0.10	<i>International Journal of Advanced Research in Computer Engineering &amp; Technology</i>	N/A
5	26	0.08	<i>Decision Support Systems</i>	5.795
6	23	0.08	<i>Lecture Notes in Computer Science</i>	N/A
7	20	0.08	<i>Information Systems Research</i>	5.207

### 3.4. Visualizing the Intellectual Base

A knowledge map based on co-cited references was visualized next (Figure 6), utilizing the *CiteSpace* node type of *Reference* and including pathfinder network pruning, with a g-index selection criteria scale factor of 40 and node labelling configured to 11 (Threshold), 11 (Font Size), and 200 (Node Size). Ranked listings of the leading co-cited references, based on the number of citations and betweenness centrality scores, are reflected in tables 6 and 7.

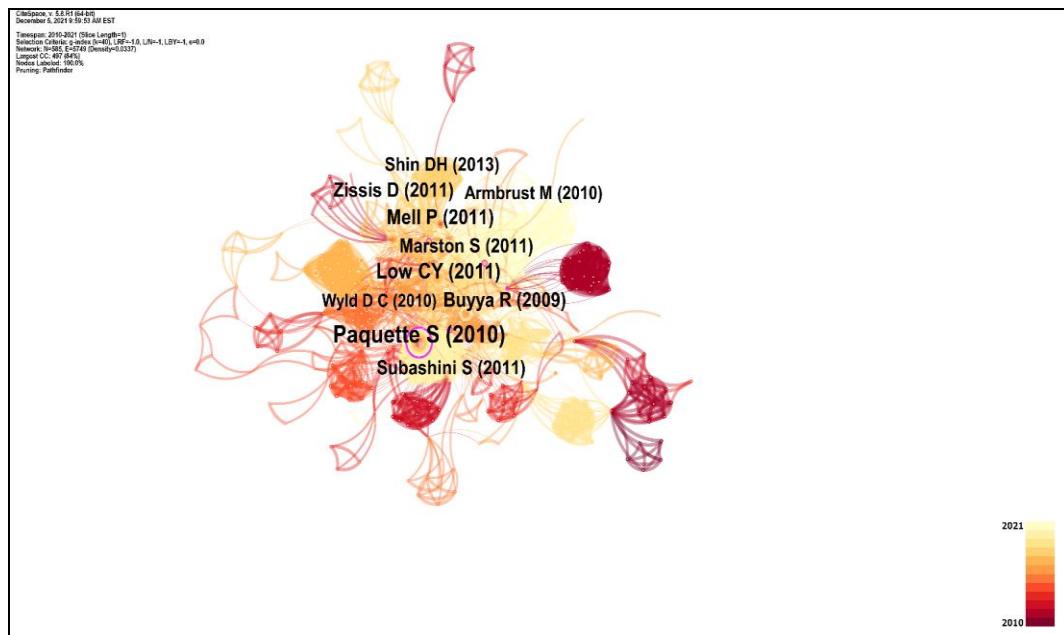


Figure 6. Knowledge Map of Co-Cited References

Table 6. Top 5 Co-Cited References Ranked by Number of Citations

Ranking	Co-Citations	Title	Representative Author
1	29	Identifying the Security Risks associated with Governmental Use of Cloud Computing	Paquette, S (2010)
2	19	Understanding the Determinants of Cloud Computing Adoption	Low, CY (2011)
3	16	The NIST Definition of Cloud Computing: Recommendations of the National Institute of Standards and Technology	Mell, P (2011)
4	15	Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility	Buyya, R (2009)
5	15	Securing e-Government and e-Voting with an Open Cloud Computing Architecture	Zissis, D (2011)

*Ties in the number of co-citations were resolved based on betweenness centrality scores.*

Table 7. Top 5 Co-Cited References Ranked by Betweenness Centrality Scores

Ranking	Betweenness Centrality	Title	Representative Author
1	0.32	Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility	Buyya, R (2009)
2	0.26	Identifying the Security Risks associated with Governmental Use of Cloud Computing	Paquette, S (2010)
3	0.12	A Survey on Security Issues in Service Delivery Models of Cloud Computing	Subashini, S (2011)
4	0.12	A View of Cloud Computing	Armbrust, M (2010)
5	0.12	The Cloudy Future Of Government IT: Cloud Computing and The Public Sector Around The World	Wyld, D C (2010)

*Ties in betweenness centrality were resolved based on number of co-citations.*

The co-cited references were then clustered based on *CiteSpace*'s default clustering method, spectral clustering. Spectral clustering automatically identifies clusters based on the eigenvectors of Laplacian matrices that are constructed from the network [25], [38]. Cluster labels were extracted based on log-likelihood ratio (LLR) analysis to extract representative noun phrases from the titles of papers that cited each cluster (Figure 7).

*CiteSpace* identified a total of ten distinct clusters that were characterized by a modularity Q score of 0.705, an indicator that the network was reasonably split into lightly joined clusters [25]. As a further indicator of the quality of the clustered network, the mean silhouette score was 0.9204, signalling that the identified clusters possessed high homogeneity [25]. Out of the ten distinct clusters, the three largest and most impactful clusters were designated as major and are reflected in Table 8. The top terms in these clusters were extracted and ranked based on log-likelihood ratio (LLR) tests derived from [39]. All three top clusters display strong silhouette values that indicate reliable clustering results.

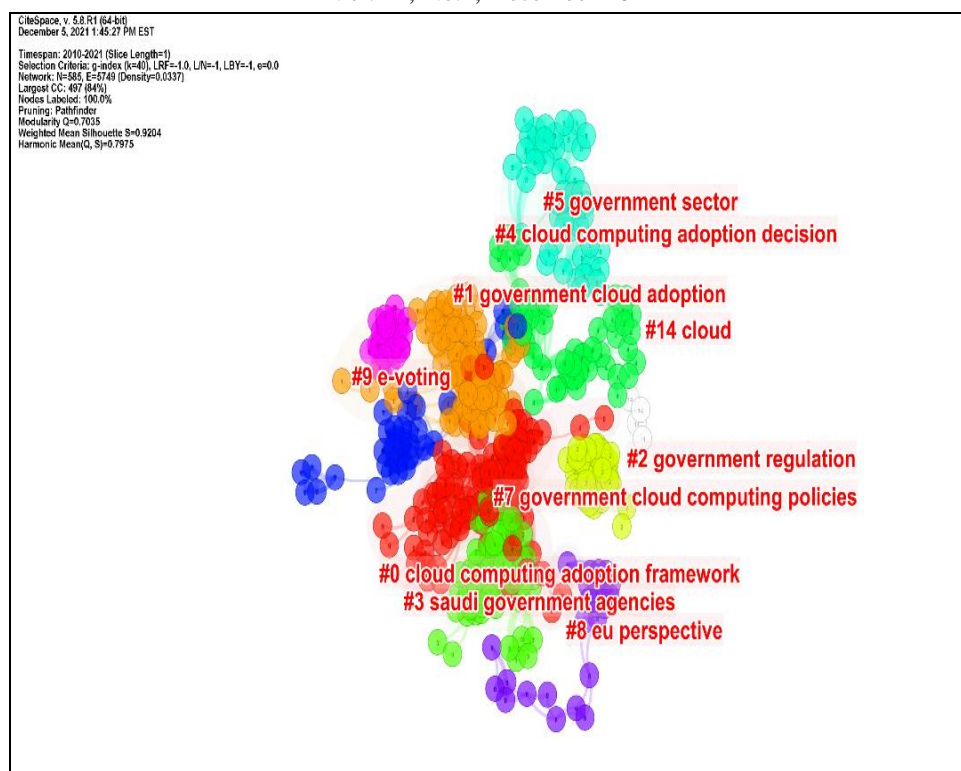


Figure 7. Knowledge Map of Co-Cited Reference Clusters

Table 8. Top 3 Co-Cited Reference Clusters by Size

Cluster ID	Size	Silhouette	Mean Year	Top Terms in Cluster <i>extracted with LLR algorithm</i>
0	98	0.778	2008	<b>cloud computing adoption framework</b> (32.98, 1.0E-4); quantitative study (30.41, 1.0E-4); using extended utaut (30.41, 1.0E-4); g-cloud acceptance (30.41, 1.0E-4)
1	87	0.909	2008	<b>government cloud adoption</b> (21.26, 1.0E-4); empirical analysis (21.26, 1.0E-4); mobile government cloud (18.73, 1.0E-4); empirical study (16.65, 1.0E-4); saudi government agencies (16.55, 1.0E-4)
2	70	0.993	2007	<b>government regulation</b> (16.18, 1.0E-4); case study (16.18, 1.0E-4); assessing information security risk (13.67, 0.001); australian local government authorities (13.67, 0.001); local government (12.81, 0.001)

### 3.5. Keyword Co-Occurrence

A knowledge map based on keyword co-citation was visualized next (Figure 8), using the *CiteSpace* node type of *Keyword* and including pathfinder network pruning, with a g-index selection criteria scale factor of 40 and keyword labelling configured to 23 (Threshold), 13 (Font Size), and 200 (Node Size). Table 9 shows the ten keywords with the most frequent appearance in the knowledge domain. The keyword appearing most commonly in the research domain was

cloud computing (105), followed by e-government (38), adoption (25), determinant (13), and challenge (13).



Figure 8. Knowledge Map of Keyword Co-Occurrence

Table 9. Most Frequent Keywords

Ranking	Keyword	Frequency	Centrality	Ranking	Keyword	Frequency	Centrality
1	cloud	105	0.93	6	government	12	0.17
2	computing			7	model	12	0.12
2	e-	38	0.47	8	security	9	0.12
3	government	25	0.07	9	innovation	9	0.10
4	adoption	13	0.10	10	information	7	0.02
4	determinant	13	0.10		technology		
5	challenge	13	0.06				

*Ties in frequency were resolved based on betweenness centrality scores.*

The co-occurring keywords were then clustered to develop further insight into the dominant research themes in the knowledge domain. 16 keyword clusters emerged, with five clusters being characterized as major (Figure 9).

The biggest cluster (#0) contained 67 members and had a silhouette value of 0.906; it was labelled “cloud adoption” (28.41, 1.0E-4) based on local linear regression (LLR) analysis. This dominant cluster focused on how government entities pursue the transition to cloud computing, including the technology, process, and organizational issues that impact cloud adoption.

Cluster (#1) was the second-biggest collection of members (45) with a silhouette value of 0.854 and was labelled “suggesting cloud computing migration framework” (51.9, 1.0E-4) based on

LLR analysis. This major cluster focused on migration frameworks that can aide public sector organizations in evaluating and structuring cloud migration efforts.

The third-largest cluster (#2) had 34 members and a silhouette value of 0.821; it was labelled “exploratory study” (34.77, 1.0E-4) based on LLR analysis. The cluster focused on exploratory case studies that showed specific focus on cloud adoption in local government organizations. The 4th largest cluster (#3) had 32 members and a silhouette value of 0.874; like cluster #0 it was also labelled “cloud adoption” (45.38, 1.0E-4) by LLR analysis but focused heavily on cloud adoption issues in India and Saudi Arabia.

The 5th largest cluster (#4) contained 29 members and had a silhouette value of 0.922. It was labelled “technological leap” (32.35, 1.0E-4) by LLR analysis and focused on the technological transformation aspects of cloud migration that can enable public sector organizations to escape the shortcomings of legacy information technology.

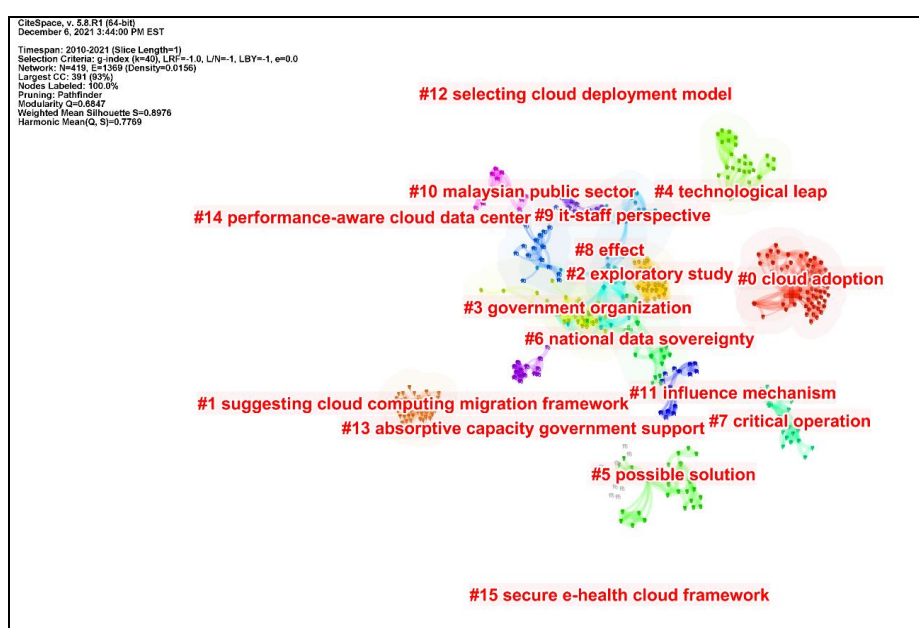


Figure 9. Knowledge Map of Keyword Clusters

## 4. DISCUSSION

### 4.1. Publication Activity

In this study, the People’s Republic of China emerged in the lead regarding the number of publications addressing public sector cloud computing, evidencing a strong level of research interest vis-à-vis the cost, benefit, and application of cloud computing solutions in the government sector. Despite this numerical dominance in the knowledge domain, the analysis of centrality metrics identified Canada, the United States, and Australia as the most important countries in the domain as publications from these three countries held the most influential positions in the knowledge maps that were created. The variety in publication activity and influence also surfaced in the institutional-level review, where the focused research activity of specific authors pushed Universiti Teknologi Malaysia, University of Southern Queensland, and University of Southampton into the top spots in terms of number of publications in the knowledge domain. Based on the academic orientation of the WoS collection and the dominance

of academic journals in the *WoS* dataset, this robust presence by university-connected institutions is not surprising.

Understanding the geographical and institutional distribution of the research in the domain allows researchers to recognize understudied areas and can identify key institutions for collaborative and comparative research on public sector cloud computing.

#### **4.2. Authors and Author Co-Citations**

Finding the productive researchers in a knowledge domain, as well as mapping their collaborative activities, is a useful tool in understanding a domain and its trajectory. Omar Ali emerged as the lead author in the knowledge domain, evidencing the highest number of publications, and was followed by Fathey Mohammed and Jeffrey Soar. In terms of collaboration, the knowledge domain contained small, nascent research clusters that were grouped around Ali, Mohammed, Li, Alassafi and that were responsible for a majority of the output in the domain. The presence of such few distinct author clusters reflects the dearth of exploration and collaboration in the knowledge domain and must serve as a call to action for additional scholars and practitioners to join the important endeavour of investigating the crucial topic of cloud computing in public sector environments.

For co-citations, Peter Mell, the NIST computer scientist behind the initial conceptualization of cloud computing in the U.S. federal government, justifiably dominates the field and is followed by Rajkumar Buyya and Scott Paquette, who have both made significant and innovative contributions to the research domain.

#### **4.3. Journals / Publications**

The exploration of journal/publication co-citations revealed the central publication conduits in the public sector cloud computing knowledge domain. The most highly influential journals/publications, based on their betweenness centrality scores, included Communications of the ACM, Journal of Network and Computer Applications, and International Journal of Advanced Research in Computer Engineering & Technology. NIST Special Publications also emerged as a heavily co-cited publication due to their seminal role in formulating the initial definition for cloud computing that has since then received far-reaching recognition beyond the U.S. federal sector, even though the publication type unambiguously does not present an accessible publication channel for regular scholarly authors.

#### **4.4. Visualizing the Intellectual Base**

In the arena of scientometric analysis, the research front in a knowledge domain consists of the contemporary research articles that address the subject area, while the intellectual base comprises the references contained in the research front articles. This study identifies the intellectual base of the public sector cloud computing knowledge domain by investigating the citation incidence and betweenness centrality scores of the references contained in contemporary scholarly articles focusing on cloud computing in the government sector.

The leading co-cited references in the public sector cloud computing domain—both from a purely numerical and an influential perspective measured by betweenness centrality—revolve around central themes that elucidate the topics that have motivated the scholarly community: cloud computing security, factors that affect the adoption of cloud computing, and empirically-confirmed benefits and drawbacks of cloud computing in the public sector.

For further analytic insight, this study also investigated and visualized the co-cited references in the domain by clustering the intellectual base. The ten distinct clusters that emerged displayed a good mean silhouette score and reasonable modularity Q score that indicated high homogeneity and acceptable cohesion within the clusters. The most dominant clusters were focused on cloud computing adoption frameworks, government cloud adoption, and government regulation.

#### **4.5. Keywords**

This study also utilized a keyword co-citation analysis to provide additional depth for the domain exploration. Cloud computing, e-government, and government adoption emerged as the most common keywords in the knowledge domain, signalling that e-government and adoption present the most dominant areas of concern for the scholarly community from a descriptive, keyword perspective. Other recurring themes that emerged in the keyword analysis revolved around the factors that affect cloud adoption in government (“determinant”), the challenges that government organizations encounter when adopting cloud (“challenge”), and the important security considerations that arise from cloud computing in the public sector (“security”). Finally, the keyword “innovation” emerged as an important theme in the domain, acknowledging the rising importance of cloud computing as an enabler of public sector innovation, expanding the central benefits of the technology from cost savings to digital transformation and adaptation.

Keywords were also clustered and revealed 16 distinct keyword clusters, with five dominant clusters that revolved around central themes of cloud adoption, cloud computing migration frameworks, exploratory investigations of cloud computing, government cloud adoption in distinct geographical contexts, and technological leaps enabled by cloud computing. The technological leap theme again emphasizes the accelerating narrative of cloud computing as an innovation enabler that produces important transformational operational values that significantly exceed the cost benefits that initially created an interest in cloud adoption.

### **5. CONCLUSION**

This study visualized the domain of cloud computing in the public sector and generated insightful, accessible knowledge maps addressing publication activities, author co-citations, publication co-citations, the domain’s intellectual base, and reference and keyword cluster analyses that evidence the breadth of cloud computing research while also revealing the nascent nature of this evolving yet critical research domain. Going beyond traditional surveys of scholarly knowledge, the present study provides a quantitative and replicable investigation of the contemporary research on cloud computing in the public sector and provides insights on research themes and gaps that can guide future research and signal productive topics and venues for research collaboration. Cloud computing is a critical feature of government transformation and innovation and continued, well-aimed research on the subject will benefit governments and citizens alike.

#### **5.1. Limitations**

This study relies on core data from the Web of Science (WoS) database. While WoS is a leading repository for high-quality scientometric data with good coverage of information technology and public sector journals and publications, it is by no means a complete source. As such, the knowledge maps and analysis presented here are subject to limitations in terms of representativeness and further research aiming at other research databases is indicated. Additionally, the analysis and visualization in this study relied exclusively on the widely-utilized CiteSpace software. CiteSpace has a solid track record and has been validated in a variety of

distinct research settings, but for additional exploration of the scientometric data and further triangulation of the results further research utilizing other software tools is advisable. Finally, this present study presents a snapshot of the state-of-the-art of a nascent, constantly evolving knowledge domain and will benefit from updates as time progresses.

## REFERENCES

- [1] P. Mell and T. Grace, "The NIST Definition of Cloud Computing," 2011. [Online]. Available: <https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-145.pdf>.
- [2] D. Insights, "Government Trends 2020 What are the most transformational trends in government today? A REPORT FROM THE DELOITTE CENTER FOR GOVERNMENT INSIGHTS About the Deloitte Center for Government Insights," 2020.
- [3] R. Agarwal, N. Khan, L. Santos, and G. Shenai, "How public-sector tech leaders can speed up the journey to the cloud," 2020. [Online]. Available: <https://www.mckinsey.com/industries/public-and-social-sector/our-insights/how-public-sector-tech-leaders-can-speed-up-the-journey-to-the-cloud>.
- [4] S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, and A. Ghalsasi, "Cloud computing - The business perspective," *Decis. Support Syst.*, vol. 51, no. 1, pp. 176–189, Apr. 2011, doi: 10.1016/j.dss.2010.12.006.
- [5] T. C. Chieu, A. Mohindra, A. A. Karve, and A. Segal, "Dynamic Scaling of Web Applications in a Virtualized Cloud Computing Environment," in *2009 IEEE International Conference on e-Business Engineering*, 2009, pp. 281–286, doi: 10.1109/ICEBE.2009.45.
- [6] "Building a culture of innovation to better serve citizens | AWS Public Sector Blog." <https://aws.amazon.com/blogs/publicsector/building-culture-innovation-better-serve-citizens/> (accessed Apr. 20, 2021).
- [7] AWS, "Amazon Web Services: Risk and Compliance," 2020. [Online]. Available: [https://d1.awsstatic.com/whitepapers/compliance/AWS\\_Risk\\_and\\_Compliance\\_Whitepaper.pdf](https://d1.awsstatic.com/whitepapers/compliance/AWS_Risk_and_Compliance_Whitepaper.pdf).
- [8] B. Shopp, "How the Pandemic Impacted Government's Cloud Migration Plans: The Good, the Bad, and the Ugly," 2021. <https://www.nextgov.com/ideas/2021/06/how-pandemic-impacted-governments-cloud-migration-plans-good-bad-and-ugly/174602/> (accessed Aug. 01, 2021)
- [9] G. Allison, "Public and Private Management: Are They Fundamentally Alike in All the Unimportant Respects?," in *Classics of Public Administration*, J. Shafritz and A. Hyde, Eds. Homewood: Dorsey, 1987.
- [10] T. Christensen, P. Laegreid, and K. A. Rovik, *Organization Theory and the Public Sector: Instrument, Culture and Myth*, 2nd ed. New York, NY: Routledge, 2020.
- [11] C. Chen, "Science Mapping: A Systematic Review of the Literature," *J. Data Inf. Sci.*, vol. 2, pp. 1–40, 2017, doi: 10.1515/jdis-2017-0006.
- [12] C. Chen and M. Song, "Visualizing a field of research: A methodology of systematic scientometric reviews," *PLoS One*, vol. 14, no. 10, 2019, doi: 10.1371/journal.pone.0223994.
- [13] "Science mapping (IEKO)." [https://www.isko.org/cyclo/science\\_mapping](https://www.isko.org/cyclo/science_mapping) (accessed Sep. 07, 2021).
- [14] A. Caputo, S. Pizzi, M. M. Pellegrini, and M. Dabić, "Digitalization and business models: Where are we going? A science map of the field," *J. Bus. Res.*, vol. 123, pp. 489–501, 2021, doi: <https://doi.org/10.1016/j.jbusres.2020.09.053>.
- [15] J. R. López-Robles, M. Rodríguez-Salvador, N. K. Gamboa-Rosales, S. Ramirez-Rosales, and M. J. Cobo, "The last five years of Big Data Research in Economics, Econometrics and Finance: Identification and conceptual analysis," *Procedia Comput. Sci.*, vol. 162, pp. 729–736, 2019, doi: <https://doi.org/10.1016/j.procs.2019.12.044>.
- [16] M. B. Negahban, "Network analysis and scientific mapping of the field of e-learning in Iran," *COLLNET J. Sci. Inf. Manag.*, vol. 15, no. 1, pp. 1–8, 2021, doi: 10.1080/09737766.2020.1853489.
- [17] F. A. Loan, B. Bashir, and N. Nasreen, "Applied artificial intelligence: A bibliometric study of an International Journal," *COLLNET J. Sci. Inf. Manag.*, vol. 15, no. 1, pp. 27–45, 2021, doi: 10.1080/09737766.2021.1938742.
- [18] D. J. de Solla Price, "Networks of scientific papers: The pattern of bibliographic references indicates the nature of the scientific research front," *Science (80-. )*, vol. 149, no. 3683, pp. 510–515, 1965, doi: 10.1126/science.149.3683.510.
- [19] D. Crane, *Invisible colleges: diffusion of knowledge in scientific communities*. Chicago: University of Chicago Press, 1972.

- [20] C. Chen, Mapping Scientific Frontiers: The Quest for Knowledge Visualization, 2nd ed. London, UK: Springer-Verlag, 2013.
- [21] C. Chen, "CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature," *J. Am. Soc. Inf. Sci. Technol.*, vol. 57, no. 3, pp. 359–377, 2006, doi: <https://doi.org/10.1002/asi.20317>.
- [22] L. Leydesdorff, "Betweenness centrality as an indicator of the interdisciplinarity of scientific journals," *J. Am. Soc. Inf. Sci. Technol.*, vol. 58, no. 9, pp. 1303–1319, 2007, doi: <https://doi.org/10.1002/asi.20614>.
- [23] J. Golbeck, "Chapter 21 - Analyzing networks," in *Introduction to Social Media Investigation*, J. Golbeck, Ed. Boston: Syngress, 2015, pp. 221–235.
- [24] "Web of Science Core Collection - Web of Science Group." <https://clarivate.com/webofsciencegroup/solutions/web-of-science-core-collection/> (accessed Sep. 08, 2021).
- [25] C. Chen, *How to use CiteSpace*. Philadelphia, PA: Leanpub, 2020.
- [26] P. Vinkler, *The Evaluation of Research By Scientometric Indicators*, 1st ed. Boston, MA: Chandos Publishing, 2010.
- [27] B. Godin, "On the origins of bibliometrics," *Scientometrics*, vol. 68, no. 1, pp. 109–133, 2006, doi: [10.1007/s11192-006-0086-0](https://doi.org/10.1007/s11192-006-0086-0).
- [28] M. Franceschet, "A cluster analysis of scholar and journal bibliometric indicators," *J. Am. Soc. Inf. Sci. Technol.*, vol. 60, no. 10, pp. 1950–1964, Oct. 2009, doi: <https://doi.org/10.1002/asi.21152>.
- [29] C. Chen, F. Ibekwe-SanJuan, and J. Hou, "The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis," *J. Am. Soc. Inf. Sci. Technol.*, vol. 61, no. 7, pp. 1386–1409, Jul. 2010, doi: <https://doi.org/10.1002/asi.21309>.
- [30] R. Todeschini and A. Baccini, *Handbook of bibliometric indicators: quantitative tools for studying and evaluating research*. Weinheim, Germany: Wiley-VCH, 2016.
- [31] C. M. Trujillo and T. M. Long, "Document co-citation analysis to enhance transdisciplinary research," *Sci. Adv.*, vol. 4, no. 1, p. e1701130, Oct. 2021, doi: [10.1126/sciadv.1701130](https://doi.org/10.1126/sciadv.1701130).
- [32] W. J. Zhang, "E-Open Government Based on Cloud Computing," *EBM 2010: INTERNATIONAL CONFERENCE ON ENGINEERING AND BUSINESS MANAGEMENT, VOLS 1-8*, no. International Conference on Engineering and Business Management. pp. 5442–5446, 2010.
- [33] D. C. Wyld, "Risk in the Clouds?: Security Issues Facing Government Use of Cloud Computing," *INNOVATIONS IN COMPUTING SCIENCES AND SOFTWARE ENGINEERING*, no. International Conference on Systems, Computing Sciences and Software Engineering (SCSS) is part of the International Joint Conferences on Computer, Information, and Systems Sciences, Engineering (CISSE). pp. 7–12, 2010, doi: [10.1007/978-90-481-9112-3\\_2](https://doi.org/10.1007/978-90-481-9112-3_2).
- [34] S. Demir, E. A. Sezer, and H. Sever, "Modifications for the Cluster Content Discovery and the Cluster Label Induction Phases of the Lingo Algorithm," *Int. J. Comput. Theory Eng.*, vol. 6, no. 2, pp. 86–90, Apr. 2014, doi: <http://dx.doi.org/10.7763/IJCTE.2014.V6.842>.
- [35] L. Feng, Y. K. Chiam, E. R. Mohd Faiza, and U. Obaidellah, "Using Suffix Tree Clustering Method To Support The Planning Phase Of Systematic Literature Review," *Malaysian J. Comput. Sci.*, vol. 30, no. 4 SE-Articles, pp. 311–332, Dec. 2017, doi: [10.22452/mjcs.vol30no4.4](https://doi.org/10.22452/mjcs.vol30no4.4).
- [36] Clarivate Analytics, "2020 Journal Citation Reports." 2021.
- [37] Institute of Internal Auditors, *Supplemental Guidance: Public Sector Definition*. Altamonte Springs, FL: Institute of Internal Auditors, 2011.
- [38] C. Chen, F. Ibekwe-SanJuan, and J. Hou, "The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis," *J. Am. Soc. Inf. Sci. Technol.*, vol. 61, no. 7, pp. 1386–1409, Jul. 2010, doi: <https://doi.org/10.1002/asi.21309>.
- [39] T. Dunning, "Accurate Methods for the Statistics of Surprise and Coincidence," *Comput. Linguist*, vol. 19, no. 1, pp. 61–74, 1993.