

Contemporary and future research of digital humanities: a scientometric analysis

Wirapong Chansanam¹, Abdul Rahman Ahmad², Chunqiu Li³

¹Department of Information Science, Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand

²Faculty of Information Management, Universiti Teknologi Mara, Selangor, Malaysia

³School of Government, Beijing Normal University, Beijing, China

Article Info

Article history:

Received Jan 6, 2022

Revised Mar 7, 2022

Accepted Mar 11, 2022

Keywords:

Bibliometric analysis

Biblioshiny

Digital humanities

Scientometric analysis

VOSviewer

ABSTRACT

This article analyses the special publications from 1999 to 2021, focusing on the scope and effect of digital humanities (DH) relevance research. The bibliometric analysis offers information on article publishing patterns, notable authors, cited references, institutions, and nations. Additionally, this paper covers the structure of knowledge of DH, including famous themes, co-citations, and bibliographic networks. Scopus database was used to obtain bibliographic data on September 25th, 2021, from papers published between 1999 and 2021. The bibliographic data for 1,249 publications using open-source analytic tools like the biblioshiny package in RStudio and the VOSviewer software shows it already. It employed the bibliometric package. These programmes visualized bibliography data based on their co-occurrence, co-citation, and coupling. The journal's publishing output and status in the area continue to rise, with 3,270 research papers indexed by Scopus to date. Additionally, the examination provides a thorough grasp of preceding patterns and forecasts a journal's future propensity.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Wirapong Chansanam

Department of Information Science, Faculty of Humanities and Social Sciences, Khon Kaen University

Khon Kaen, Thailand

Email: wirach@kku.ac.th

1. INTRODUCTION

The humanities and social sciences have been transformed due to the globalization of knowledge, tradition, heritage, culture, and expressions. The progress of computational artefacts, social media, and network infrastructure has created a once-in-a-lifetime opportunity for twenty-first-century humanities education and study [1]. As a result, digital humanities (DH) is seen as a novel field with a solid experimental attitude that fosters knowledge convergence through transdisciplinary, multidisciplinary, and internationality [2]. Scientists, academics, and researchers are developing new concepts, roadmaps, and action points for gaining access to and examining this further replication of the humanities [3].

The future path of study will be multidisciplinary. Additionally, this means that DCH's approach to resolving the DCH issue must be interdisciplinary. It is used to develop interdisciplinary models for tackling DCH issues, and it is essential to form an interdisciplinary team of professionals [4]. In addition, the discipline of digital computational humanities must direct computational science, social network theory and analysis, digital literacy, computational pedagogy, computational literacy, machine and deep learning, neuro and cognitive learning, and analysis, as these are all critical components.

Learning to build new modes of enormous knowledge, society, and culture in the networked and digital settings is a roadmap to establishing new forms of tremendous knowledge, society, and culture. It is hoped that the roadmap would aid DCH stakeholders in becoming comprehensive digital humanists and

attain the solution to DCH issues via the application of relevant methodologies. Upadhyay and his associates [5], [6] specific DCH issues relating to social media, digital culture, and big data analytics have been addressed. One of the essential functions of the Roadmap is to offer an adequate framework for conceptualizing DCH activities and developing and implementing strategies that contribute to attaining the overall DCH goal. To sum it up, the Roadmap is divided into two primary scopes of activity: first, the establishment of program-level efforts to achieve a common DCH aim; and second, the development of project scope for attaining DCH competencies and skillsets, together with the necessary environment and infrastructure [6].

The DH research organization has 114 active centres in twenty-four countries, including South Korea, Brazil, South Africa, and Japan. It is one of the largest research organizations in the world. The number of centres in the United States and the United Kingdom is significantly higher, followed by Canada, Germany, and Australia. There is a substantial discrepancy in the number of DH research centres in different countries [7]. Furthermore, academics that attended important DH research conferences tended to be from North America and Europe rather than other parts of the world [8]. According to Zhang and his associates [9], The DH community is geographically unbalanced, with most scholars based in European countries (notably the United Kingdom) and the United States. In contrast, Asia and Oceania have a small number of researchers. It is important to seek solutions through an ongoing process of theorizing to bridge regional and global gaps and reconcile contradictions between local practises and global perceptions in the DH because the subject is naturally interdisciplinary [8]. It is estimated that the number of articles related to DH that were published in Information Science and Technology Abstracts (LISTA), a database for library and information science, approximately tripled between 2005 and 2012, according to Sula [10]. As a result, the importance of DH in academic research is rapidly increasing. Many studies on DH scholarship have been conducted, including research patterns in North America and Europe, research assistance, and scholar views, to name a few topics [7], [11]-[14]. In contrast, no bibliometric studies of research trends have yet been conducted. The results of this study are used to suggest directions for the future development of DH. As a result, this article employs the bibliometrics technique to investigate the current state and perception of DH through systematic analysis and then uses the findings to suggest directions for its future development. This article examines the development and historical movement (past and current) and contemporary (present). Furthermore, it explores current trends and developments in DH study (the future). The research answers two critical questions:

RQ1: The complete information and publishing trend in DH is becoming more popular, includes top-cited articles, top-ranked authors, publishing institutes and countries, and keywords.

RQ2: The conceptual structure of extent and influence of the DH.

The quantitative features of citation information and Scopus publishing structures will be investigated to answer the research issues raised above. There are many characteristics, including a keyword co-occurrence network, a temporal overlay keyword visualization, and co-citation and bibliography information. Furthermore, this study shows how authors, organizations, and nations worked together to structure the knowledge domains of journals in the context of knowledge domain structuring. By identifying existing and emerging research topics that will aid in developing intelligent future studies, this study assesses its contribution to the DH field. The following section provides a breakdown of the paper's organizational structure. The next part covers the research methodologies, followed by the findings and comments; the last section discusses the study's conclusion and its contribution and flaws.

2. METHOD

This article quantitative analyzes the DH using worldwide publishing statistics from the Scopus database. This research was carried out on September 25th, 2021, and included the terms "DH" in the article titles and abstracts, as:

```
TITLE-ABS-KEY ( "digital humanities" ) AND ( LIMIT-TO ( SRCTYPE , "j" ) ) AND ( LIMIT-TO ( DOCTYPE , "ar" ) ) AND ( EXCLUDE ( PUBYEAR , 2022 ) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) )
```

This study examines a wide range of research trends, authors, universities, keywords, publications, citation structures, and the development of research trends through time. The bibliometric approach was used to examine the visual representation of the growth of DH research. Additionally, data from bibliometric analyses were examined with the use of the Bibliometrix R programme [15], [16].

3. RESULTS AND DISCUSSION

Since its first article in 1999, DH has been referenced in 1,249 publications, with an average of 6.55 citations for each item published. The findings of the DH Scopus bibliometric study are described in this part of the paper. Furthermore, the findings of this study are presented in the form of research questions.

3.1. Descriptive analysis

3.1.1. Digital humanities publications during 1999-2021

Figure 1 shows a substantial increase in the number of research papers published over the past twenty years, showing the high level of interest that has grown in the academic community over this period. Compared to 1999, the pace of growth has risen from 2 documents per year to 137 documents per year now (recorded in 2021), and the highest recorded in 2019 with 208 documents. Because just two papers were published in 1999, we can conclude that DH are a relatively new field based on the analysis of publications. The Scopus database has a total of 1,249 papers that have been collected for twenty years. This trend is anticipated to become more pronounced in 2021 and the following years (during this analyzing not the end of the year yet).

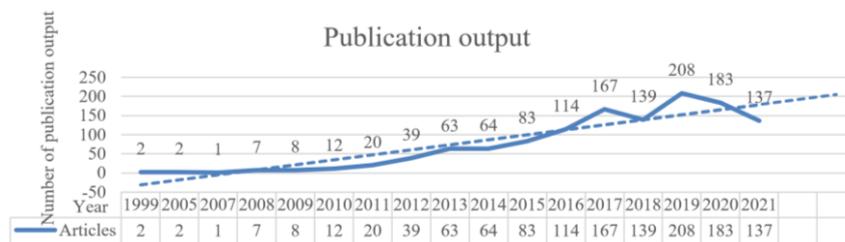


Figure 1. DH publication output

3.1.2. Discipline wise analysis

The research into the DH has been conducted extensively in the area of social sciences (i.e., 39.92%), art and humanities (28.64%), and computer science (19.4%) are the second and third most relevant disciplines in DH, respectively, and additional information can be found in Figure 2.

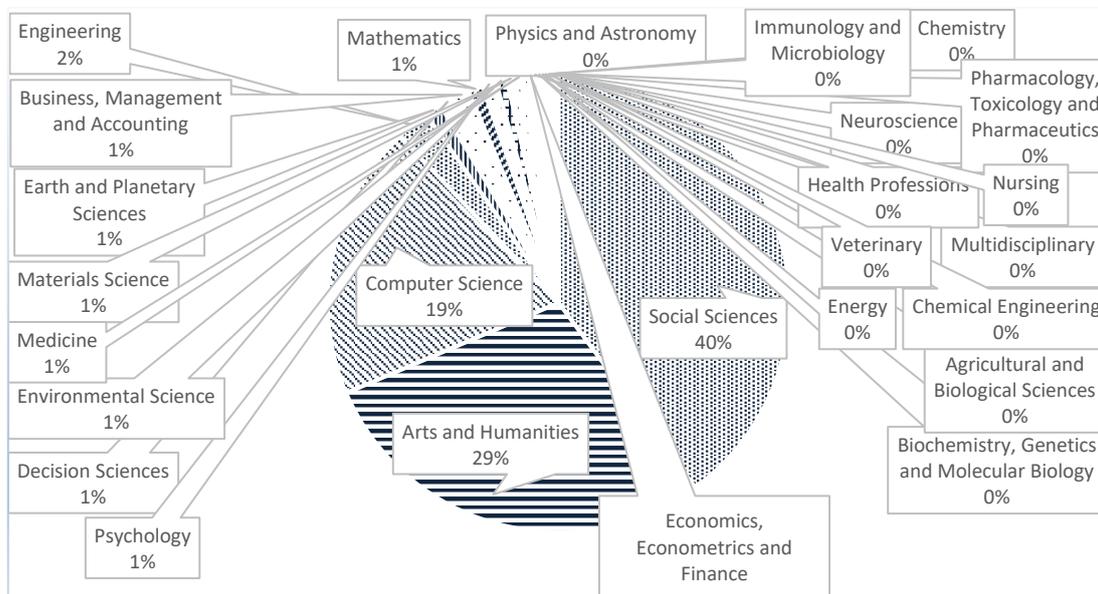


Figure 2. Distribution of documents across subject area

3.2. Bibliometric analysis

3.2.1. Digital humanities articles developments in publishing and general description

To answer the first question, you must first grasp the structure of DH publication descriptions, trends, and citations. Table 1 shows the average citations per year. It was reported in 2014 that the most significant total number of citations per article was 28.76, with an annual average of 4.11, indicating that the

publications in 2014 were the most referenced in the field of DH. Articles with the second most significant number of citations received 18.35 in 2013, with an average of 2.30 each year.

Table 1. Average citations per year

Year	N	MeanTCperArt	MeanTCperYear	CitableYears
1999	2	0.5	0.022727273	22
2000	0	0	0	0
2001	0	0	0	0
2002	0	0	0	0
2003	0	0	0	0
2004	0	0	0	0
2005	2	3	0.1875	16
2006	0	0	0	0
2007	1	0	0	14
2008	7	8.571428571	0.659340659	13
2009	8	5.5	0.458333333	12
2010	12	12.33333333	1.121212121	11
2011	20	22.95	2.295	10
2012	39	10.43589744	1.15954416	9
2013	63	18.34920635	2.293650794	8
2014	64	28.765625	4.109375	7
2015	83	6.975903614	1.162650602	6
2016	114	5.807017544	1.161403509	5
2017	167	4.730538922	1.182634731	4
2018	139	3.007194245	1.002398082	3
Total	721	6.546307251		

3.2.2. Three fields plot

The three-field plot depicts three components, including a list of journal names, authors, and subjects and a list of journal names, authors, and topics. The names of the journals are given first, followed by the names of the writers, and then each author is connected to the topic of the publication to which they contributed. Furthermore, grey connections are used to tie these three components together to demonstrate their interdependence. The number of articles on each element is represented by the rectangles on the lists, with journals being the first element on the left. Scopus has indexed twelve journals, with the most articles appearing in the 'Journal of Documentation.' It is represented as a purple rectangle that is connected to several other journals. The names of the writers are included inside the centre section. The main components are linked with the authors of papers that have been published in well-known publications. This graphic contains a list of the top twenty writers, with the size of each author's rectangle reflecting the number of articles the author has authored. The Miller rectangle is the most significant shape in this graphic. This section contains the most frequently used terms connected to the subject. Each topic is linked with one or more writers who have published the most articles on the subject. There are twenty keyword topics, including DH, digital scholarship, collaboration, digital history, data visualization, digital libraries, pedagogy, cultural heritage, metadata, archives, crowdsourcing, visualization, humanities, digitization, history, mapping, distant reading, gis, academic libraries, and the semantic web as shown in Figure 3.

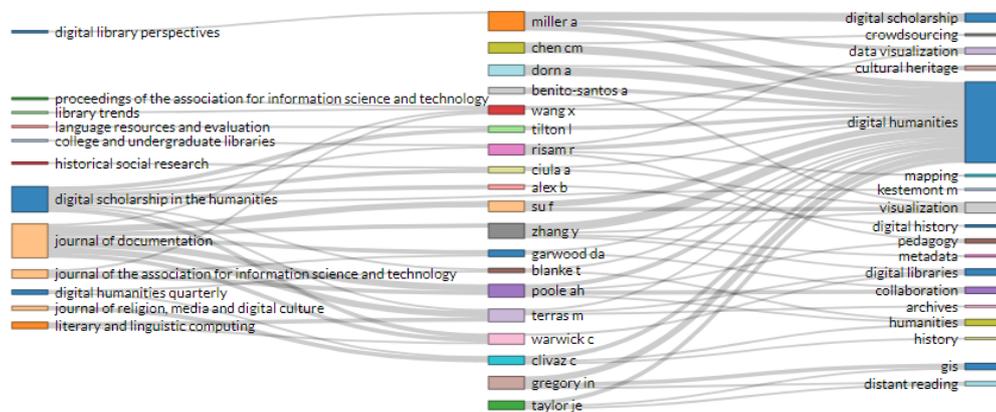


Figure 3. Three fields plot

3.2.3. Most relevant sources

The articles published by each publication are shown in the Figure 4 chart, which ranks them according to their importance to the DH study. Furthermore, the data displays the top journals and papers that have been published, with a ranking ranging from 0 to 55.

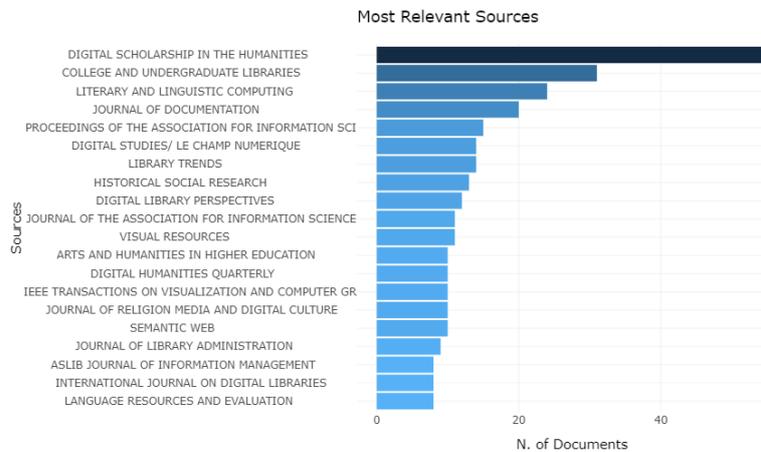


Figure 4. Most relevant sources

3.2.4. Influential authors and their citation structure

According to the bar chart in Figure 5, this research assessed the influence of each journal published in DH by calculating their h-indices based on their number and relevance. Furthermore, the shade of blue in this figure represents the effect of the journals, with more profound blue signifying publications having a more significant impact. With an 864 score, an h-index of 1, and 1 document, Rob Kitchin from Maynooth University in Maynooth, Ireland, is the most cited author, according to Table 2. Rob Kitchin is the most cited author, according to Table 3. Melissa Terras from The University of Edinburgh in Edinburgh, United Kingdom, on the other hand, is the most prolific with six papers to his credit, and he also has the highest h-index score of five. Furthermore, academic writers from the United Kingdom and the United States lead the list of top publishers. Table 3 shows the citation pattern of the top-ranking publishing authors.

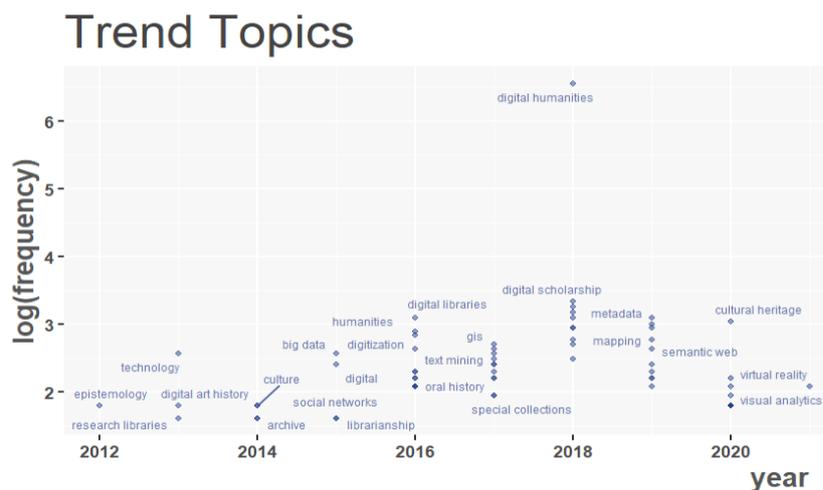


Figure 5. Topic historiography timeline

3.2.5. Trend topics

According to the term frequency found in this MCS study, topics arise as a result. As shown in the accompanying chronology, these are the most frequently mentioned subjects. Keywords often used appear higher on the list, and their occurrence points are represented in the timeline.

Figure 5 depicts the topic's development since 2012 when it saw an accelerated rise. The timeline reveals that epistemology has been the most frequently discussed subject since 2012. Between 2012 and 2014, epistemology, technology, digital art history, research libraries, archive, and culture. Big data, digital, librarianship, and social networks interest between 2014 and 2016. Digitization, digital libraries, gis, text mining, oral history, special collections began to explore DH topics in 2016. Between 2018 and 2020, digital scholarship, metadata, mapping, and semantic web were popular in DH research, especially in the DH topic numerous popular in 2018. From 2020, cultural heritage, virtual reality, and visual analytics became increasingly popular.

Table 2. Most cited article from 1999-2021

R	Article	DOI	TC	TCY
1	Kitchin, R. Big data, new epistemologies and paradigm shifts. <i>Big Data Soc.</i> 2014; 1.	10.1177/2053951714528481	864	108.00
2	Holmberg, K. elwall, M.(2014). Disciplinary differences in Twitter scholarly communication. <i>Scientometrics</i> , 101(2), 1027-1042.	10.1007/s11192-014-1229-3	162	20.25
3	Anderson, C. W. (2013). Towards a sociology of computational and algorithmic journalism. <i>New media & society</i> , 15(7), 1005-1021.	10.1177/1461444812465137	156	17.33
4	Ross, C., Terras, M., Warwick, C., & Welsh, A. (2011). Enabled backchannel: Conference Twitter use by digital humanists. <i>Journal of Documentation</i> .	10.1108/00220411111109449	136	12.36
5	Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. <i>Information, communication & society</i> , 15(5), 662-679.	10.1108/AJIM-09-2013-0094	126	15.75
6	Hu, Y., Boyd-Graber, J., Satinoff, B., & Smith, A. (2014). Interactive topic modeling. <i>Machine learning</i> , 95(3), 423-469.	10.1007/s10994-013-5413-0	116	14.50
7	Eder, M., Rybicki, J., & Kestemont, M. (2016). Stylometry with R: a package for computational text analysis. <i>The R Journal</i> , 8(1).	10.32614/rj-2016-007	94	15.67
8	Armitage, D. (2012). What's the Big Idea? <i>Intellectual History and the Longue Durée. History of European Ideas</i> , 38(4), 493-507.	10.1080/01916599.2012.714635	90	9.00
9	Cooper, D., & Gregory, I. N. (2011). Mapping the English lake district: a literary GIS. <i>Transactions of the Institute of British Geographers</i> , 36(1), 89-108.	10.1111/j.1475-5661.2010.00405.x	81	7.36
10	Posner, M. (2013). No half measures: Overcoming common challenges to doing digital humanities in the library. <i>Journal of Library Administration</i> , 53(1), 43-52.	10.1080/01930826.2013.756694	73	8.11
11	DeLysler, D., & Sui, D. (2014). Crossing the qualitative-quantitative chasm III: Enduring methods, open geography, participatory research, and the fourth paradigm. <i>Progress in Human Geography</i> , 38(2), 294-307.	10.1177/0309132512444063	72	8.00
12	Ronen, S., Gonçalves, B., Hu, K. Z., Vespignani, A., Pinker, S., & Hidalgo, C. A. (2014). Links that speak: The global language network and its association with global fame. <i>Proceedings of the National Academy of Sciences</i> , 111(52), E5616-E5622.	10.1073/pnas.1410931111	71	8.88
13	Liu, A. (2013). The meaning of the digital humanities. <i>PMLA/Publications of the Modern Language Association of America</i> , 128(2), 409-423.	10.1632/pmla.2013.128.2.409	67	7.44
14	Grandjean, M. (2016). A social network analysis of Twitter: Mapping the digital humanities community. <i>Cogent Arts & Humanities</i> , 3(1), 1171458.	10.1080/23311983.2016.1171458	66	11.00
15	Vinopal, J., & McCormick, M. (2013). Supporting digital scholarship in research libraries: Scalability and sustainability. <i>Journal of Library Administration</i> , 53(1), 27-42.	10.1080/01930826.2013.756689	60	6.67
16	Sula, C. A. (2013). digital humanities and libraries: A conceptual model. <i>Journal of Library Administration</i> , 53(1), 10-26.	10.1080/01930826.2013.756680	59	6.56
17	Vandegrift, M., & Varner, S. (2013). Evolving in common: Creating mutually supportive relationships between libraries and the digital humanities. <i>Journal of Library Administration</i> , 53(1), 67-78.	10.1080/01930826.2013.756699	59	6.56
18	Meroño-Peñuela, A., Ashkpour, A., Van Erp, M., Mandemakers, K., Breure, L., Scharnhorst, A., ... & Van Harmelen, F. (2015). Semantic technologies for historical research: A survey. <i>Semantic Web</i> , 6(6), 539-564.	10.3233/SW-140158	58	7.25
19	Hitchcock, T. (2013). Confronting the digital: or how academic history writing lost the plot. <i>Cultural and Social History</i> , 10(1), 9-23.	10.2752/147800413X13515292098070	58	6.44
20	Burdick, A., & Willis, H. (2011). Digital learning, digital scholarship and design thinking. <i>Design Studies</i> , 32(6), 546-556.	10.1016/j.destud.2011.07.005	56	5.09

Source: Scopus database, R=rank, TC=total citations, TCY=total citations per year

Table 3. Most productive authors and the citation structure

R	Author	Affiliation	H-index	G-index	M-index	TC	NP	PY start
1	Kitchin R	Maynooth University, Maynooth, Ireland	1	1	0.125	864	1	2014
2	Warwick C	Durham University, Durham, United Kingdom	4	5	0.308	172	5	2009
3	Terras M	The University of Edinburgh, Edinburgh, United Kingdom	5	6	0.385	166	6	2009
4	Holmberg K	Turun yliopisto, Turku, Finland	1	1	0.125	162	1	2014
5	Thelwall M	University of Wolverhampton, Wolverhampton, United Kingdom	1	1	0.125	162	1	2014
6	Anderson Cw	University of Leeds, Leeds, United Kingdom	1	1	0.111	156	1	2013
7	Welsh A	University College London, London, United Kingdom	2	3	0.182	146	3	2011
8	Ross C	University College London, London, United Kingdom	1	1	0.091	136	1	2011
9	Borra E	Universiteit van Amsterdam, Amsterdam, Netherlands	1	1	0.125	126	1	2014
10	Rieder B	Universiteit van Amsterdam, Amsterdam, Netherlands	1	1	0.125	126	1	2014
11	Boyd-Graber J	University of Maryland, Bethesda, United States	1	1	0.125	116	1	2014
12	Hu Y	University of Maryland, College Park, Maryland, USA	1	1	0.125	116	1	2014
13	Satinoff B	University of Maryland, College Park, College Park, United States	1	1	0.125	116	1	2014
14	Smith A	Arizona State University, Tempe, United States	1	1	0.125	116	1	2014
15	Gregory In	Lancaster University, Lancaster, United Kingdom	4	5	0.364	113	5	2011
16	Liu A	University of California, Santa Barbara, Santa Barbara, United States	2	2	0.200	109	2	2012
17	Kestemont M	Universiteit Antwerpen, Antwerpen, Belgium	3	4	0.429	101	4	2015
18	Rybicki J	Forschungszentrum Jülich (FZJ), Jülich, Germany	1	3	0.167	95	3	2016
19	Eder M	Polish Academy of Sciences, Warszawa, Poland	1	1	0.167	94	1	2016
20	Armitage D	Harvard University, Cambridge, MA, USA	1	1	0.100	90	1	2012

3.2.6. Thematic map

This research examined a thematic map by separating it into four topic quadrants based on density and centrality, as shown in the picture to the right. Due to their great density and centrality, the topics in the upper right quadrant should be explored and researched further. In contrast, a unique, uncommon, but highly developing theme with high density and low centrality is in the top left quadrant. Themes with a downward tendency are located in the lower-left quadrant, while basic themes with a high centrality but a low density are located in the lower-right quadrant. In Figure 6, the timeline demonstrates that the subjects of digital libraries, cultural heritage, digitalization, semantic web, and libraries in the upper-right quadrant of the timeline are the most promising for future study. This is consistent with Qing Wang's research [17], finding that this field of research is closely related to history, literary, cultural heritage, and information and library science. However, the results differ from research by [18], which found that the network of co-authors remained relatively sparse. The collaboration was largely limited by language and geography. Domain-specific practices in the DH have been discussed that may contribute to such fragmentation.

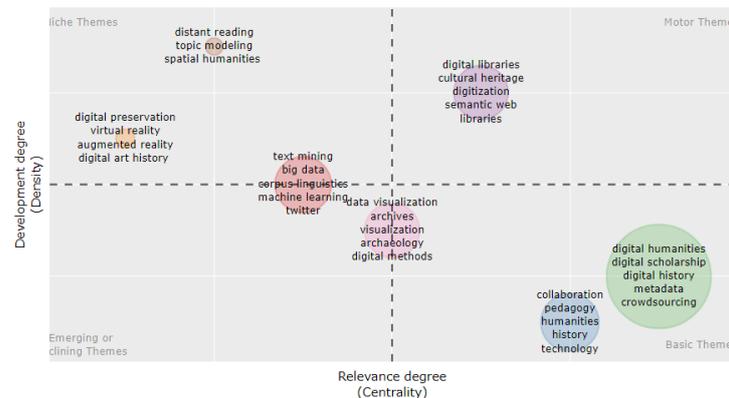


Figure 6. Thematic map

3.3. Structure of digital humanities research

The research indicates that the author and journal co-citation analysis and coupling bibliometric and network techniques are used in this study [15].

3.3.1. Historiography of digital humanities on Scopus

A historiographic map of the most important publications was created to comprehend better and illustrate the development of DH research. Figure 7 depicts three different study areas. The first research stream is primarily concerned with a list of debates surrounding the DH [19]. Four risks are identified based on observations of the current state of the DH and their environment: (1) the emphasis on infrastructures for the DH may conceal the fact that analytical techniques and tools ultimately drive research, not merely data or publishing technologies, (2) information technology may help the Humanities in various ways, depending on the country. The fact that textual analysis is a hot topic right now should not obscure the perspective of a larger academic area, (3) the impending mobile revolution may result in a repeat of the very destructive processes seen during the PC and internet revolutions, (4) the DH may need to play a more significant role in creating technology, not only the reception of it. The second research stream focuses on the DH and digital libraries coming together [9]. First, DH's creative approach to research and teaching techniques creates both possibilities and problems, according to his study results. Second, DH research is a team effort. Third, vital avenues for the DH community are developed. Fourth, various tools and data sets are created to assist various project kinds. Fifth, the geographical and disciplinary dispersion of the DH community is uneven. Sixth, attention, integration, and sustainability are still lacking in DH research output. Finally, in DH initiatives, LIS experts perform a particular role. Overall, the DH and DL groups have similar aims and duties. It offers the following current and future LIS roles: creator and contributor; curator; messenger and liaison; educator; mediator and interpreter; host; partner; innovator; "hybrid scholar"; advocate; consultant. Libraries could improve their efficiency and effectiveness in their services by redesigning their organizational culture or structure to encourage and actualize more profound cross-border dialogues and partnerships. On a broader scale, the DL community should work to become more visible, valuable, and approachable to the DH community; and, better still, to become a part of it. The Scalar was used in the third stream to evaluate DH tools at a Research University [20]. This research librarian is more supportive of digital publishing platforms; yet, they must also be aware of the user experience with these tools. The usage of Scalar, a DH publishing platform for media-rich projects, at the University of Illinois in Urbana–Champaign is examined in this case study. Based on a survey, interviews, and content analysis, the research emphasizes the platform's usability, functionality, and achievements and failures in achieving user expectations. User concerns were influenced by the media upload procedure, picture annotation, and aesthetics. Writing pedagogy was also mentioned as a factor to consider. The findings point to lessons for digital literacy training and how and when Scalar may help patrons with their publishing requirements.

Historical Direct Citation Network

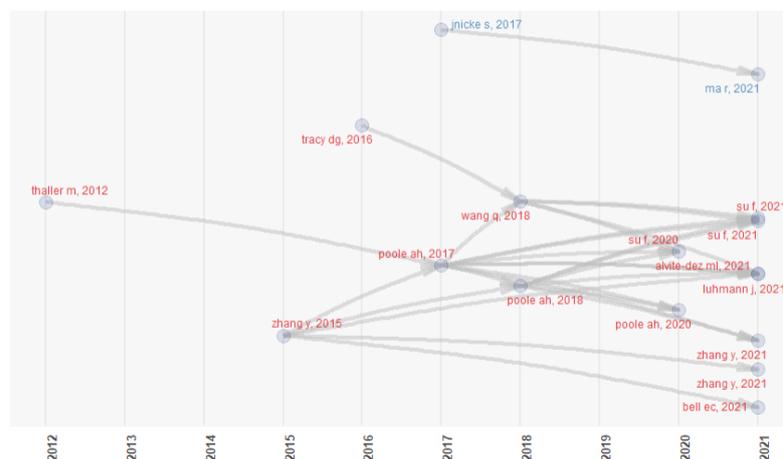


Figure 7. Thematic map

3.3.2. Co-citation analysis of sources and documents

This research advanced a network of co-cited publications, with 101 of the 26,685 sources fulfilling the first criterion of a minimum of 20 citations. Seven clusters emerged from a co-citation study of these 101 sources, including the top-100 sources with the most vital links. The first cluster (37 sources) is the largest and includes A companion to DH, communications of the ACM, Comput. Graph. Forum. The second (25 sources) is the next largest cluster, consisting of studies from American Archivist, Archival science,

The published papers were then bibliographically linked using the top 100 articles with the highest link strength and at least 50 citations from 1,249 publications. Figure 10 shows a four-cluster network with 26 linkages and a total link strength of 34 due to the coupling. The first cluster (3 authors) is the biggest. Kitchin [29] examines how the availability of big data, combined with new data analytics, challenges established epistemologies across the sciences, social sciences, and humanities and assesses the extent to which they are engendering paradigm shifts across multiple disciplines, and denotes this cluster. It critically examines new forms of empiricism that proclaim "the end of theory," the development of data-driven rather than knowledge-driven science, and the development of DH and computational social sciences, all of which propose radically different ways to make sense of culture, history, economy, and society. It is argued that (1) big data and new data analytics are disruptive innovations that are reshaping how research is conducted in many cases; and (2) there is an urgent need for broader critical reflection within the academy on the epistemological implications of the unfolding data revolution, a task that has only just begun to be addressed despite the rapid changes in research practices that have occurred. Following a critical examination of developing epistemological views, it is proposed that creating a situated, reflexive, and contextually nuanced epistemology would be a potentially successful approach. His article aims to introduce the Twitter Capture and Analysis Toolset from the Digital Methods Initiative, which is a suite of tools for gathering and analyzing Twitter data. Rather than just giving a technical article describing the system, the authors argue that the sort of data utilized in computational systems and the techniques embedded in them have epistemological implications for research. As a result, the authors want to place the toolset's development in the context of methodological disputes in the social sciences and humanities. To address the different ways computational systems frame research, the authors evaluate the potential and limits of existing techniques to gather and analyze Twitter data. The authors next go through the open-source toolkit and present a strategy that considers methodological and epistemological variety. When developing tools for computational social research or DH, the authors discover that design decisions and more general methodological thinking may and should go hand in hand. Aside from methodological openness, the program offers reliable and repeatable data gathering and analysis and interoperability with other analytical applications. By considering how Twitter organizes information, allowing for a range of sampling methodologies, permitting a diversity of analytical approaches or paradigms, and supporting work at the micro, meso, and macro levels, epistemic plurality is highlighted. At the same time, Borra and Rieder [30] and Hu and his associates [31] propose a method for providing users with a voice by encoding users' feedback to topic models as word correlations in a topic model. Interactive topic modelling (ITM) is a system that allows untrained users to and repeatedly encode their input into topic models. They create more efficient inference methods for tree-based topic models since latency is critical in interactive systems. The framework is tested with both simulated and actual users.

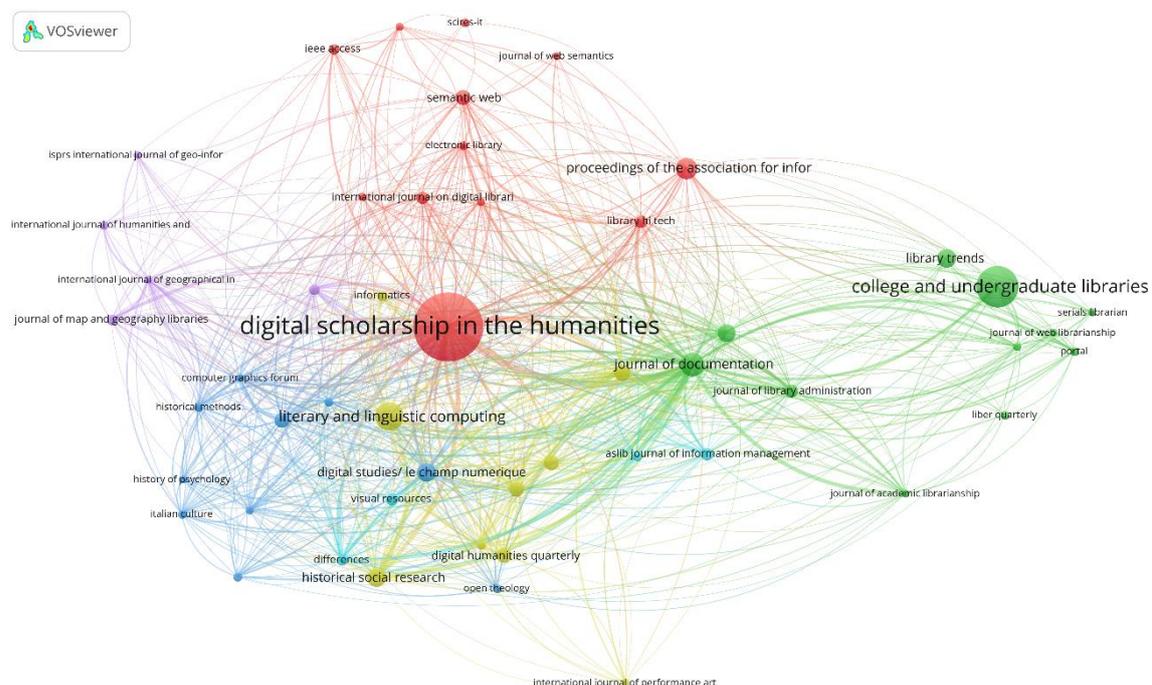


Figure 10. Bibliographic document coupling

Anderson [32], which develops a sociological approach to computational journalism, represents cluster two (4 authors). The article uses the term "computational journalism" to refer to the increasingly common kinds of algorithmic, social scientific, and mathematical networks used by many 21st-century newsrooms and hailed by many educational institutions as "the future of news." The term "sociological approach" is used in the article to describe a study paradigm that, for the time being, ignores many of the current industry concerns about the practical usefulness of the journalistic analysis. Digital learning, digital scholarship, and design thinking [33] is another crucial work in this cluster, identifying potential for design thinking to be integrated into digital learning and digital scholarship efforts. They ended with a proposal for design thinking research to engage with new models for learning and knowledge creation, work that would have long-term epistemic consequences. They proposed a list of highlights Learning models and literary production have both been affected by digital technologies. Design thinking talents are valued in digital learning projects. The design of scholarly platforms and interfaces has an epistemic influence. This necessitates design thinking that is determined by interpretative, situated modes of knowledge. Digital educators and design thinkers can create new learning and research modalities. The outcome is to support consilience, synergy, and a positive acceptance of variety in geographical scholarship, according to DeLyser and Sui [34]. They crossed the qualitative-quantitative divide II: Inventive approaches to big data, mobile techniques, and rhythm analysis. Liu [35] defines the DH.

Eder and his associates [36], which created a program for computational text analysis, represents cluster three (3 authors). They describe 'Stylometry with R' (stylo), a powerful R program for stylometry's high-level examination of writing style. Making DH work at a smaller university by doing a lot with a little [37]. A historical study using semantic technologies: A survey [38] describes open issues and potential research areas for advancing the historical Semantic Web, which is still new yet promising. Cluster four (three authors) focuses on the following big idea: The Longue Durée and Intellectual History [39]. In the library, there are DH [10], [40] and the digital divide [41].

3.3.4. Digital humanities publications structure

This research made use of a previously published author cooperation analysis in DH research. The study was based on the most significant cooperation between 198 core authors, examined for 1,000 authors. Figure 11 shows cluster one is the largest (Links 169 TLS 3237 Citations 172) with 55 authors focusing on digital art history and represented by Johanna Drucker [42]. Next is cluster two, the second largest (Links 8 TLS 1116 Citations 96) with 41 authors. The primary authors in this cluster are Thomas Jacobs and Robin Tschötschel; they focus on topic models that meet discourse analysis by the quantitative tool for a qualitative approach [43]. Cluster three includes 29 authors (Links 120 TLS 1708 Citations 44) and is represented by Jänicke, Franzini, Cheema, Scheuermann, Cooper, Gregory, Visual Text Analysis in DH [44], [45]. The fourth (Links 173 TLS 2499 Citations 192), five (Links 100 TLS 690 Citations 39), six (Links 50 TLS 798 Citations 47) cluster have 21 authors and is represented by Moretti, Sprugnoli, Menini, & Tonelli, Extracting and visualizing content from extensive document collections to support humanities studies [46]. Clusters, seven has five authors (Links 104 TLS 1115 Citations 30) and is represented by Thaller [17]. Cluster eight has three authors (Links 54 TLS 522 Citations 43) and is represented by Burman, Green, and Shanker [47] DH in service of conceptual clarity.

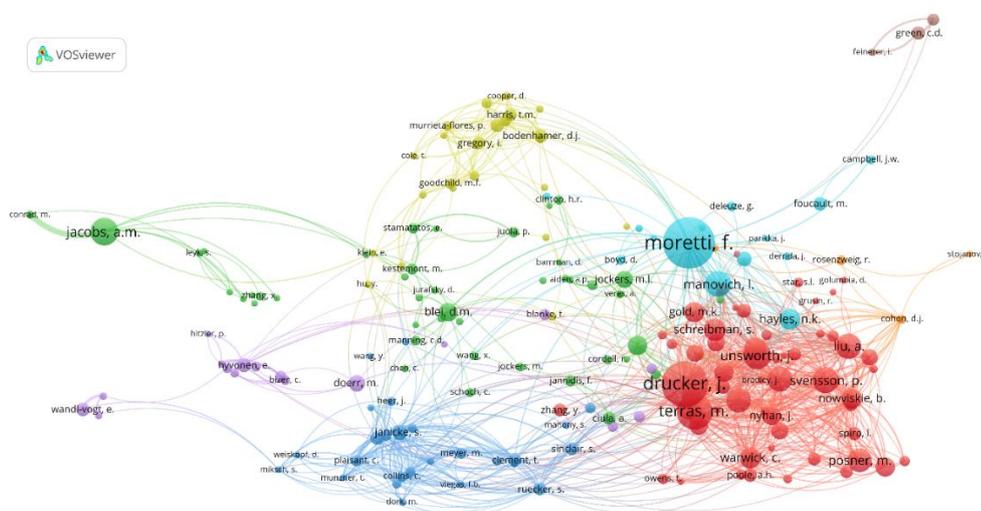


Figure 11. DH publications structure in map view

4. CONCLUSION

A quantitative and systematic literature review is used in this research to assess the DH' reach and effect. It's a great way to reflect on the past while also contributing to the advancement of future knowledge. The study found broad trends in DH research based on the two research topics, such as the most prolific authors, institutions, nations, important literature, and citation structure. These research topics were also addressed using bibliometric indicators published between 1999 and 2021. Bibliometrics R was used to examine the data. The goal of the first step is to identify significant trends and provide an overview of the research base in DH. After developing a history map and utilizing VOSviewer software for DH research in stage 2, a knowledge structure was created using co-word analysis. Literary and Linguistic Computing, Debates in the DH, and DHQ: DH Quarterly are the top three journals in this field. United States, United Kingdom, and Germany are the top three countries on the list. The University of California, the United States' most prolific institution, is followed by the United Kingdom's University College London and the United States' New York University.

In addition, Kitchin is the most important article on DH, followed by Holmberg and Anderson. A strategic diagram created utilizing the authors' keyword co-word analysis was also used in this study to help comprehend the conceptual structure of the DH research base. According to the findings, DH research falls into seven broad categories. In addition to DH and Collaboration, there is also a major topic on Digital libraries included in this list. Also, there is one partial basic and emerging or clining themes on data visualization; there is one partial emerging or clining themes on text mining; there are two niche themes on distant reading and digital preservation. Using the intellectual publishing structure, this following addition aids comprehension of the disciplinary makeup of the DH knowledge base. Three distinct maps or networks were created, such as the historiography map of the top 50 DH research papers. This analysis shows that DH research focused on three major themes: (1) epistemology, (2) digital technology, big data, social networks, and (3) DH, metadata, semantic web. Therefore, understanding how the DH knowledge base has evolved is made easier with the help of this tool. Second, by establishing source and document co-citation networks, this research examined the intellectual underpinnings of the DH research base. Third, using R software, a network of the most prolific writers, the most influential texts, and a thematic map were constructed. These contributed to a better understanding of DH' future (emerging topics).

In addition, future DH research may focus on cultural heritage, visual analytics, and virtual reality, according to the findings of this study. These results add to the growing body of work in DH. Fourth, by creating a network of co-authors, the research showed how publishing academics are now collaborating. DH has a slew of publishing teams, as seen by the graph. These may be beneficial to writers in terms of looking back at past partnerships and creating new ones.

This article provides a thorough overview of the DH knowledge base, allowing scholars to quickly access a variety of trends, topics, and streams. However, this research examined just the bibliometric data associated with the articles, not the whole text. While it is anticipated that the keywords used will accurately represent the primary substance of the text, some restrictions may apply. As a result, integrating this evaluation with more traditional methods would significantly raise the cost of such research. This contribution is intended to assist academics in DH and focus their efforts on future highly relevant research.

ACKNOWLEDGEMENT

This research is supported by Khon Kaen University and Universiti Teknologi Mara.

REFERENCES

- [1] T. Presner, Digital Humanities 2.0: a report on knowledge. *CiteSeer*, 2010.
- [2] P. Mounier, "Manifeste des Digital Humanities. ThatCamp," *Journal des anthropologues. Association française des anthropologues*, no. 122–123, pp. 447–452, 2010, doi: 10.4000/jda.3652
- [3] N. K. Hayles, "How we think: Transforming power and digital technologies," in *Understanding digital humanities*, Springer, Palgrave Macmillan, London, 2012, pp. 42–66, doi: 10.1057/9780230371934_3
- [4] N. Upadhyay and S. Upadhyay, "#RighttoBreathe why not? Social Media Analysis of the Local in the Capital City of India," *Procedia Computer Science*, vol. 108, pp. 2542–2546, 2017, doi: 10.1016/j.procs.2017.05.017
- [5] N. Upadhyay, "UriMartyrs and# IndiaWithForces what is not heard? Social Media Analysis of Tweets," in *Proceedings of IEEE International Conference on "Computing for Sustainable Global Development"*, 1st–3rd March 2017.
- [6] S. Upadhyay and N. Upadhyay, "Future directions and a roadmap in digital computational humanities for a data driven organization," *Procedia computer science*, vol. 122, pp. 1055–1060, 2017, doi: 10.1016/j.procs.2017.11.473.
- [7] M. Terras, "Present, not voting: Digital humanities in the Panopticon: Closing plenary speech, Digital Humanities 2010," *Literary and Linguistic Computing*, vol. 26, no. 3, pp. 257–269, 2011, doi: 10.1093/lc/fqr016.
- [8] R. Risam, "Other worlds, other DHs: Notes towards a DH accent," *Digital Scholarship in the Humanities*, vol. 32, no. 2, pp. 377–384, 2017, doi: 10.1093/lc/fqv063.
- [9] Y. Zhang, S. Liu, and E. Mathews, "Convergence of digital humanities and digital libraries," *Library Management*, 2015, doi:

- 10.1108/LM-09-2014-0116.
- [10] C. A. Sula, "Digital humanities and libraries: A conceptual model," *Journal of Library Administration*, vol. 53, no. 1, pp. 10–26, 2013, doi: 10.1080/01930826.2013.756680.
- [11] L. Spiro, "Knowing and Doing: Understanding the Digital Humanities Curriculum.," in *DH*, 2011, pp. 232–233.
- [12] X. Wang and M. Inaba, "Analyzing structures and evolution of digital humanities based on correspondence analysis and co-word analysis," *Art research*, vol. 9, pp. 123–134, 2009.
- [13] J. Nyhan and O. Duke-Williams, "Joint and multi-authored publication patterns in the digital humanities," *Literary and Linguistic Computing*, vol. 29, no. 3, pp. 387–399, 2014, doi: 10.1093/lc/fqu018.
- [14] A. Quan-Haase, K. Martin, and L. McCay-Peet, "Networks of digital humanities scholars: The informational and social uses and gratifications of Twitter," *Big data & society*, vol. 2, no. 1, p. 2053951715589417, 2015, doi: 10.1177/2053951715589417.
- [15] M. Aria and C. Cuccurullo, "bibliometrix: An R-tool for comprehensive science mapping analysis," *Journal of Informetrics*, vol. 11, no. 4, pp. 959–975, 2017, doi: 10.1016/j.joi.2017.08.007.
- [16] N. J. Van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523–538, 2010, doi: 10.1007/s11192-009-0146-3.
- [17] Q. Wang, "Distribution features and intellectual structures of digital humanities: A bibliometric analysis," *Journal of Documentation*, 2018, doi: 10.1108/JD-05-2017-0076.
- [18] M.-C. Tang, Y. J. Cheng, and K. H. Chen, "A longitudinal study of intellectual cohesion in digital humanities using bibliometric analyses," *Scientometrics*, vol. 113, no. 2, pp. 985–1008, 2017, doi: 10.1007/s11192-017-2496-6.
- [19] M. Thaller, "Controversies around the digital humanities: an agenda," *Historical Social Research/Historische Sozialforschung*, pp. 7–23, 2012.
- [20] D. G. Tracy, "Assessing digital humanities tools: Use of Scalar at a research university," portal: *Libraries and the Academy*, vol. 16, no. 1, pp. 163–189, 2016, doi: 10.1353/pla.2016.0004
- [21] N. J. Van Eck and L. Waltman, "CitNetExplorer: A new software tool for analyzing and visualizing citation networks," *Journal of Informetrics*, vol. 8, no. 4, pp. 802–823, 2014, doi: 10.1016/j.joi.2014.07.006.
- [22] A. H. Poole, "The conceptual ecology of digital humanities," *Journal of Documentation*, vol. 73 no. 1, pp. 91–122, 2017, doi: 10.1108/JD-05-2016-0065.
- [23] A. H. Poole, "'A greatly unexplored area': Digital curation and innovation in digital humanities," *Journal of the Association for Information Science and Technology*, vol. 68, no. 7, pp. 1772–1781, 2017, doi: 10.1002/asi.23743.
- [24] A. H. Poole and D. A. Garwood, "Interdisciplinary scholarly collaboration in data-intensive, public-funded, international digital humanities project work," *Library & Information Science Research*, vol. 40, no. 3–4, pp. 184–193, 2018, doi: 10.1016/j.lisr.2018.08.003.
- [25] A. H. Poole and D. A. Garwood, "Digging into data management in public-funded, international research in digital humanities," *Journal of the Association for Information Science and Technology*, vol. 71, no. 1, pp. 84–97, 2020, doi: 10.1002/asi.24213.
- [26] C. Warwick *et al.*, "Documentation and the users of digital resources in the humanities," *Journal of Documentation*, vol. 65, no. 1, pp. 33–57, 2009, doi: 10.1108/00220410910926112.
- [27] C. Warwick, "Interfaces, ephemera, and identity: A study of the historical presentation of digital humanities resources," *Digital Scholarship in the Humanities*, vol. 35, no. 4, pp. 944–971, 2020, doi: 10.1093/lc/fqz081.
- [28] M. Terras, R. Van den Branden, and E. Vanhoutte, "Teaching TEI: The need for TEI by example," *Literary and Linguistic Computing*, vol. 24, no. 3, pp. 297–306, 2009, doi: 10.1093/lc/fqp018.
- [29] R. Kitchin, "Big Data, new epistemologies and paradigm shifts," *Big data & society*, vol. 1, no. 1, pp. 1–12, 2014, doi: 10.1177/2053951714528481.
- [30] E. Borra and B. Rieder, "Programmed method: Developing a toolset for capturing and analyzing tweets," *Aslib Journal of Information Management*, 2014, doi: 10.1108/AJIM-09-2013-0094.
- [31] Y. Hu, J. Boyd-Graber, B. Satinoff, and A. Smith, "Interactive topic modeling," *Machine learning*, vol. 95, no. 3, pp. 423–469, 2014, doi: 10.1007/s10994-013-5413-0.
- [32] C. W. Anderson, "Towards a sociology of computational and algorithmic journalism," *New media & society*, vol. 15, no. 7, pp. 1005–1021, 2013, doi:10.1177/1461444812465137
- [33] A. Burdick and H. Willis, "Digital learning, digital scholarship and design thinking," *Design Studies*, vol. 32, no. 6, pp. 546–556, 2011, doi: 10.1016/j.destud.2011.07.005
- [34] D. DeLyser and D. Sui, "Crossing the qualitative-quantitative chasm III: Enduring methods, open geography, participatory research, and the fourth paradigm," *Progress in Human Geography*, vol. 38, no. 2, pp. 294–307, 2014, doi:10.1177/0309132513479291
- [35] A. Liu, "The meaning of the digital humanities," *PMLA/Publications of the Modern Language Association of America*, vol. 128, no. 2, pp. 409–423, 2013, doi:10.1632/pmla.2013.128.2.409
- [36] M. Eder, J. Rybicki, and M. Kestemont, "Stylometry with R: a package for computational text analysis," *The R Journal*, vol. 8, no. 1, 2016.
- [37] D. Vanderbilt, "Doing a lot with a little: Making digital humanities work at a small college," *CEA Critic*, vol. 76, no. 3, pp. 327–335, 2014.
- [38] A. Meroño-Peñuela *et al.*, "Semantic technologies for historical research: A survey," *Semantic Web*, vol. 6, no. 6, pp. 539–564, 2015, doi:10.3233/SW-140158
- [39] D. Armitage, "What's the Big Idea? Intellectual History and the Longue Durée," *History of European Ideas*, vol. 38, no. 4, pp. 493–507, 2012, doi:10.1080/01916599.2012.714635
- [40] M. Posner, "No half measures: Overcoming common challenges to doing digital humanities in the library," *Journal of Library Administration*, vol. 53, no. 1, pp. 43–52, 2013, doi:10.1080/01930826.2013.756694
- [41] T. Hitchcock, "Confronting the digital: or how academic history writing lost the plot," *Cultural and Social History*, vol. 10, no. 1, pp. 9–23, 2013, doi:10.2752/147800413X13515292098070
- [42] J. Drucker, "Is there a 'digital' art history?," *Visual Resources*, vol. 29, no. 1–2, pp. 5–13, 2013, doi:10.1080/01973762.2013.761106
- [43] T. Jacobs and R. Tschötschel, "Topic models meet discourse analysis: a quantitative tool for a qualitative approach," *International Journal of Social Research Methodology*, vol. 22, no. 5, pp. 469–485, 2019, doi:10.1080/13645579.2019.1576317
- [44] D. Cooper and I. N. Gregory, "Mapping the English lake district: a literary GIS," *Transactions of the Institute of British Geographers*, vol. 36, no. 1, pp. 89–108, 2011, doi: 10.1111/j.1475-5661.2010.00405.x.
- [45] S. Jänicke, G. Franzini, M. F. Cheema, and G. Scheuermann, "Visual text analysis in digital humanities," in *Computer Graphics Forum*, 2017, vol. 36, no. 6, pp. 226–250, doi:10.1111/cgf.12873
- [46] G. Moretti, R. Sprugnoli, S. Menini, and S. Tonelli, "ALCIDE: Extracting and visualising content from large document

- collections to support humanities studies,” *Knowledge-Based Systems*, vol. 111, pp. 100–112, 2016, doi: 10.1016/j.knosys.2016.08.003
- [47] J. T. Burman, C. D. Green, and S. Shanker, “On the meanings of self-regulation: Digital Humanities in service of conceptual clarity,” *Child Development*, vol. 86, no. 5, pp. 1507–1521, 2015, doi:10.1111/cdev.12395
- [48] K. Holmberg and M. Thelwall, “Disciplinary differences in Twitter scholarly communication,” *Scientometrics*, vol. 101, no. 2, pp. 1027–1042, 2014, doi:10.1007/s11192-014-1229-3.

BIOGRAPHIES OF AUTHORS



Wirapong Chansanam    received his Ph.D. degree in information studies from Khon Kaen University, Khon Kaen, Thailand. He was with Chaiyaphum Rajabhat University as a lecturer for nine years. In 2019, he joined Khon Kaen University, Khon Kaen, Thailand, where he is currently an Associate professor at the Faculty of Humanities and Social Sciences. His research interests include (but not limited to) information science, especially in digital humanities. He can be contacted at email: wirach@kku.ac.th.



Abdul Rahman Ahmad    received Bachelor of Law (Bachelor's Degree (Honors)), International Islamic University, Malaysia, Master of Information Management (Master), Mara University of Technology, Shah Alam Branch, Malaysia, and Doctor of Philosophy In Information Management (Research) (Ph.d.), Universiti Teknologi Mara, Cawangan Shah Alam. His research interests are information management and digital humanities. He can be contacted at email: arahman@uitm.edu.my and arahman@salam.uitm.edu.my.



Chunqiu Li    received her Ph.D. degree in information science from the Graduate School of Library, Information and Media Studies, University of Tsukuba, Tsukuba, Ibaraki, Japan. She is an Assistant professor at the School of Government, Beijing Normal University, Beijing, China. Her current research interests are in metadata, bibliographic data, information organization, government data, and data governance. She can be contacted at email: lichunqiu@bnu.edu.cn.