Using Bibliometric Analysis for Writing a Research Paper

By I. K. NTI

Presentation outline

Definition and concept of bibliometrics

Bibliometric data sources

Searching and extracting bibliometric data

Analyzing and reporting bibliometric data

Network visualization using VOSViewer

Writing a bibliometric paper

Different quantitative analysis in the field of information science

Bibliometrics: Focuses on the quantitative analysis of bibliographic data, such as citations and publication patterns, to evaluate the impact and influence of scholarly publications and authors.

Scientometrics: Explores the quantitative study of science and scientific research, including the analysis of scientific publications, citations, and collaboration networks.

Informetrics: A broader term that encompasses bibliometrics and scientometrics, as well as other measures of information production, dissemination, and usage.

Cybermetrics: Examines web-based data, including website rankings, links, and online usage patterns, to study and evaluate the impact of academic institutions and researchers in the online environment.

Webometrics: A subfield of cybermetrics, it specifically focuses on the quantitative analysis of the World Wide Web, including website metrics, link analysis, and web presence of academic entities.

Altmetrics: A newer field that explores alternative metrics beyond traditional citations, such as social media mentions, downloads, and online discussions, to measure the impact and attention received by scholarly outputs.
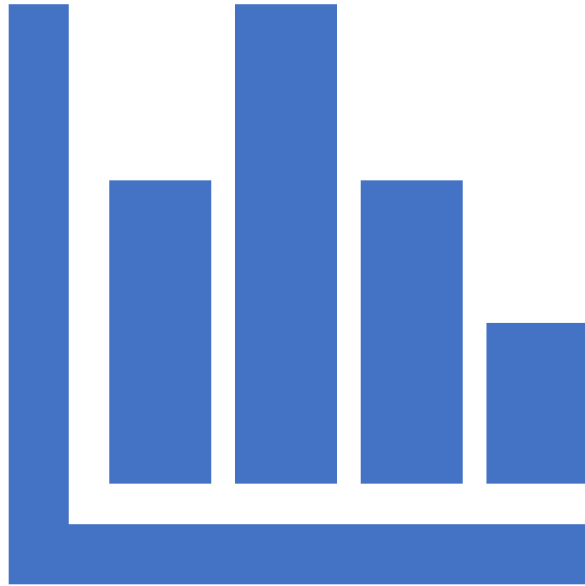
Definition and concept of bibliometrics

- Bibliometrics
  - **Biblio:** derived from "biblion" Greek word for "book"
  - **Metrics:** derived from "metrikos" Greek word for "measurement"

What is Bibliometric (1)

# Alan Pritchard 1969

➢ Coined the term "bibliometrics"

"the application of mathematics and statistical methods to books and other media of communication"

Journal of Documentation (1969) 25(4):348-349

STATISTICAL BIBLIOGRAPHY OR BIBLIOMETRICS?

The term *statistical bibliography* seems to have been first used by E. Wyndham Hulme in 1922 when he delivered two lectures as the Sandars Reader in Bibliography at the University of Cambridge. Subsequently the lectures were published as a book.[1] Although the debt has never been explicitly recognized by means of citations, Hulme anticipated modern work on the history of science. He used the term to mean the illumination of the processes of science and technology by means of counting documents. Hulme both summarized the results of Cole and Eales[2] and produced original work on the growth of UK patents (relating these to social processes in the UK) and on the changes displayed in the *International Catalogue of Scientific Literature* (relating changes in subject and country production of literature to international developments).

The term then seems to have been ignored for twenty-two years until Gosnell used it in a paper on obsolescence of literature[3] without acknowledging its previous use.
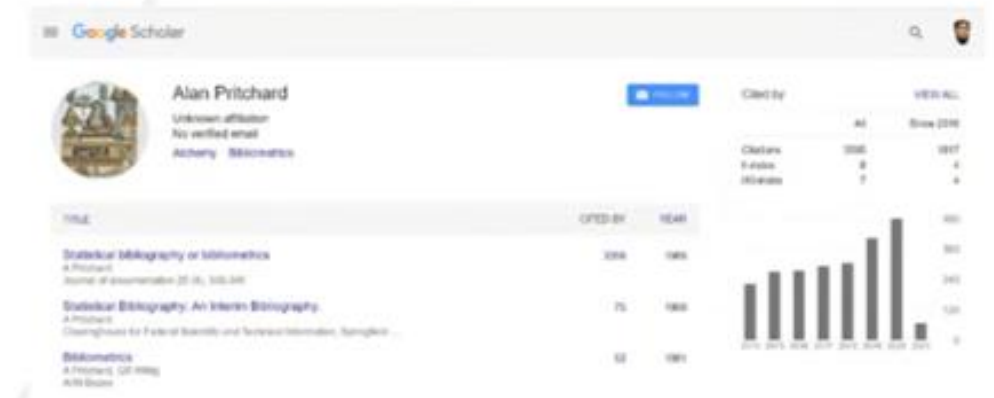
What is Bibliometric (2)

- Bibliometrics is a quantitative and statistical method used to analyze and measure various aspects of scholarly publications, such as
  - books,
  - articles,
  - journals,
  - conference proceedings, and
  - other academic literature.

- It involves the application of mathematical and statistical techniques to bibliographic data to gain insights into patterns, trends, and relationships within the literature.

Users and interested parties of bibliometrics?

- Bibliometrics is an essential tool for various stakeholders in the research ecosystem, helping them make informed decisions, understand research trends, and evaluate the impact of scholarly work.
  - Researchers
  - Academic Institutions
  - Funding Agencies
  - Publishers
  - Libraries
  - Policymakers and Governments
  - Academic Rankings Organizations
  - Academic Societies and Associations

Application of bibliometrics

To:

- Analyze **research patterns and the expansion of knowledge**.
- Examine authorship patterns across different subjects in documents.
- Predict **future publishing patterns**.
- Forecast the **productivity** of publishers, authors, organizations, and countries.
- Evaluate research performance at micro and macro levels.
- Assess the impact of research output.
- Identify **high-impact journals** in various research fields.
- Provide support for promotion, tenure, and grant applications.
- Discover potential opportunities for research collaborations.
- Enhance the global recognition of research endeavors.
- Identify areas of **research strengths and areas that need improvement**.

Bibliometric indicators

- Bibliometric indicators (metrics) are numerical measurements that provide quantitative information about research performance.

- Bibliometric indicators can be used to evaluated the performance and impact of different entities:
  - Authors
  - Departments/ Faculties
  - Institutions
  - Journals
  - Countries
  - Documents
  - Subject area and categories

Journal

Paper Title

Authors

Abstract

Authors affiliation/Institutions

References

Bibliometrics is based on bibliographic data

Using Bibliometric analysis to write a literature review paper

- A bibliometric literature review is a systematic and quantitative analysis of published literature in a specific field or research area.

- It involves applying bibliometric techniques to analyze various bibliographic elements of literature, such as ***authorship, publication, citation, and keyword patterns***.

- The objective of a bibliometric literature review is to gain insights into the intellectual structure, trends, and impact of the research field under investigation.

Typically Steps in Bibliometric Literature Review
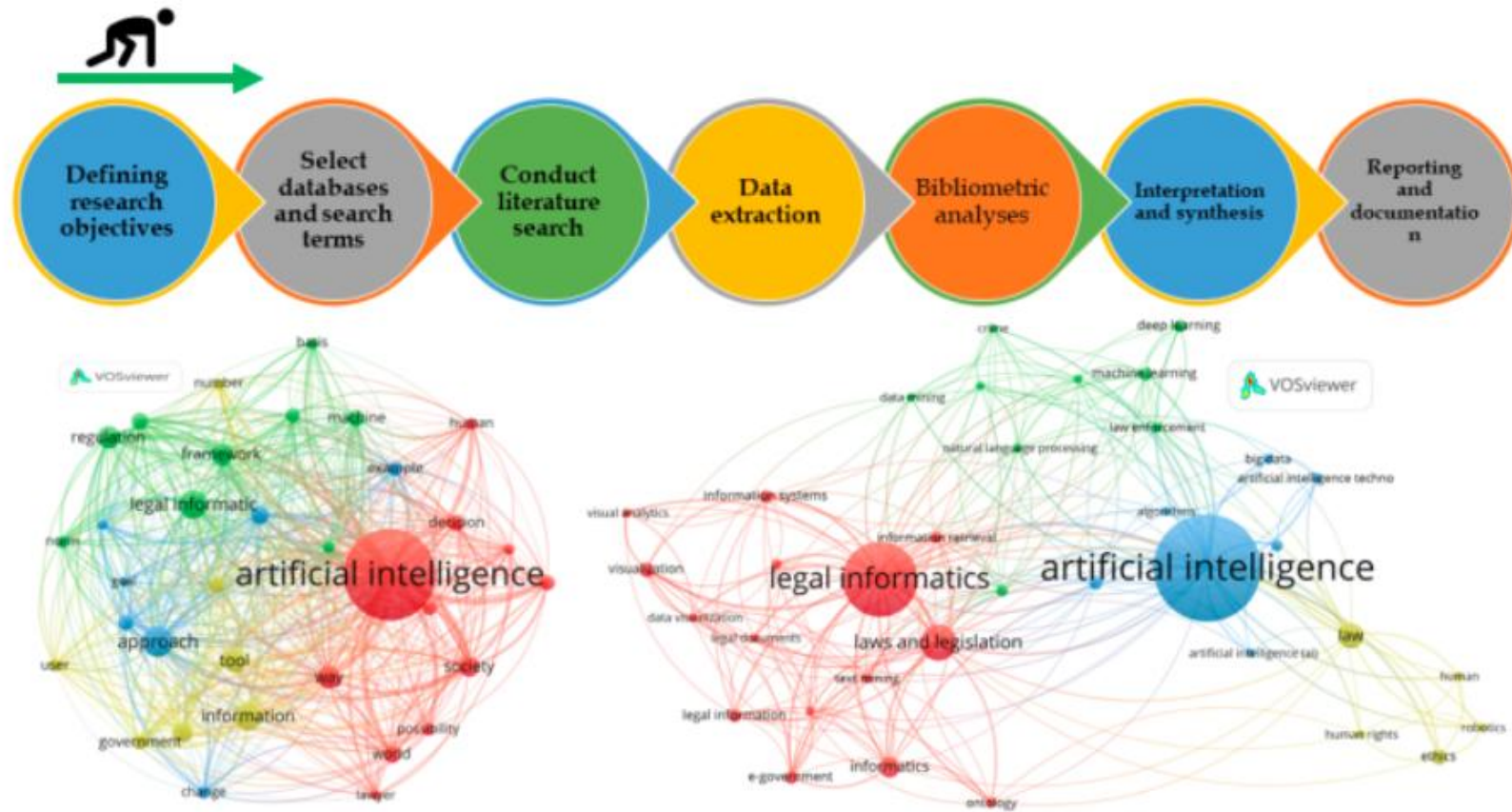
Define research Objectives/Questions (1)

- Determine the specific research questions or objectives you want to explore through bibliometric analysis.

Therefore, the goal of the current study is to assess research on healthcare technology using a scientific mapping review procedure (bibliometric analysis). The following are the questions (ReQ) addressed in this paper.

ReQ1: What are the critical growing trends in the application of technology such as AI, ML, IoT, IoMT, big data, and cloud computing for sustainable healthcare?

ReQ2: Over the last 24 years, which *authors, institutions, papers, and journals* in the literature on TSH have had the most significant impact on citations?

ReQ3: What are the critical themes investigated concerning TSH, and how do they relate?

ReQ4: What issues in TSH have received the most attention and have been studied the most frequently?

ReQ5: What does the nature of cooperation look like in the literature on TSH?

(Nti et al. 2023)

Consequently, it is essential to conduct a thorough and methodical analysis of the use of artificial intelligence in law. This study seeks to achieve the following objectives:

1. Identify the prolific researchers, organizations, and nations/regions involved in artificial intelligence in law throughout the past 62 years, as well as their cooperative partnerships.
2. List the primary areas of focus for artificial intelligence research in the legal field throughout the last six (6) decades.
3. Discover the connections between various study topics and their rate of evolution.
4. List the advantages and difficulties of artificial intelligence in the legal field and suggest exciting new study areas.

(Nti et al. 2023)

Define research objectives/Questions (2)

- Determine the specific research questions or objectives you want to explore through bibliometric analysis.

The purpose of this paper is to review research in the field of healthcare management for sustainability using science mapping review methodology. The review addresses the following research questions (RQ):

- RQ1: What are key growth trends in research on healthcare management for sustainability?
- RQ2: What authors and documents in the literature on sustainable healthcare management have had the greatest impact on citations over the past 25 years?
- RQ3: What is the intellectual structure of the knowledge base on sustainable healthcare management?
- RQ4: What topics in the sustainable healthcare management literature have been studied with the greatest frequency and are currently attracting the greatest attention?

(Punnakitikashem & Hallinger, 2019)

This comparative review of EDLM knowledge production across Asia, Africa, and Latin America was guided by several research questions.

RQ1: What is the volume and distribution by time and geographic source of EDLM research from emerging regions?

RQ2: What authors, institutions, and journal articles from emerging regions have had the greatest influence on EDLM research over the past six decades?

RQ3: What is the intellectual structure of the EDLM knowledge base from emerging regions?

(Hallinger, 2019)

Bibliometric Data Sources

- **Select databases**
  - Identify relevant databases and platforms where you will conduct your literature search.
    - For example:
    - Scopus
    - ACM Digital Library
    - Web of Science
    - Dimensions
    - IEEE Xplore
    - PubMed

Conduct a literature search (1)

✓ Perform a systematic search using your selected databases and search terms.

✓ Apply filters and inclusion/exclusion criteria to narrow down the search results to the most relevant papers.

✓ Document the search strategy and the number of articles retrieved at each stage.

Conduct a literature search (2)

- Case study paper

- **A bibliometric analysis of technology in sustainable healthcare: Emerging trends and future directions**

- **Keywords**
  - Technology
  - Sustainable healthcare
  - Health informatics
  - Machine learning in healthcare
  - Healthcare predictive analytics
  - Disease detection system
  - Health information system
  - Artificial intelligence in healthcare

Conduct a literature search (3)

Define your inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
1. Papers published from 1998 to 2022.	1. Paper not published in the English language.
2. Papers with publication status "final".	2. Papers that are not in the field of healthcare technology.
3. Papers published in journals and conferences.	3. Document type not article or conference proceedings.
4. Papers with full author information.	4. Papers "in press".
	5. Papers with no author(s) names.
	6. Papers with subject area not in healthcare or review.

Search procedures and Operators

- Booleans
- Use of quotation mark
- Use of parenthesis
- Use of wild cards
- Truncation
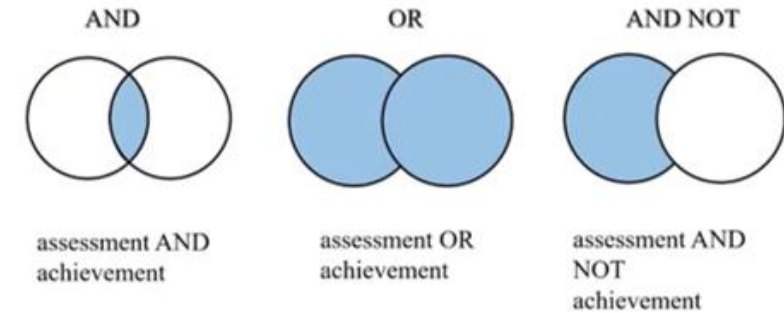
Build your search string

- Build you initial search string

- Keep refining it until you get what you want

- Carefully document every steps in these processes

<mark>**Sample initial search string**</mark>

(TITLE-ABS-KEY (technology AND sustainable AND healthcare) OR TITLE-ABS-KEY (artificial AND intelligence AND healthcare) OR TITLEABS-
KEY (machine AND learning AND healthcare) OR TITLE-ABS-KEY (health AND information AND system) OR TITLE-ABS-KEY (health AND informatics) OR TITLE-ABS-KEY
(sustainable AND healthcare) OR TITLE-ABS-KEY (healthcare AND predictive AND analytics) OR TITLE-ABS-KEY (disease AND prediction) OR TITLE-ABS-KEY (big AND data AND
health AND analytics)) AND PUBYEAR > 1997 AND PUBYEAR < 2023 AND (LIMIT-TO
Study review process. (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "cp") OR LIMIT-TO (DOCTYPE, "re") OR LIMIT-TO (DOCTYPE,
"ch") OR LIMIT-TO (DOCTYPE, "cr") OR LIMIT-TO (DOCTYPE, "bk"))

<mark>**Sample final search string**</mark>

((TITLE-ABS-KEY (technology AND sustainable AND healthcare) OR TITLE-ABS-KEY (artificial AND intelligence AND healthcare) OR TITLE-ABS-KEY
(machine AND learning AND healthcare) OR TITLE-ABS-KEY (health AND informatics) OR TITLE-ABS-KEY (sustainable AND healthcare) OR TITLE-ABS-KEY (healthcare AND predictive
AND analytics) OR TITLE-ABS-KEY (health AND informatics) OR TITLE-ABS-KEY (big AND data AND health AND analytics))) AND PUBYEAR >1997 AND PUBYEAR < 2023 AND
(EXCLUDE (PUBSTAGE, "aip")) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "cp")) AND (EXCLUDE (SUBJAREA, "ENGI") OR EXCLUDE (SUBJAREA, "SOCI") OR
EXCLUDE (SUBJAREA, "MATH") OR EXCLUDE (SUBJAREA, "BUSI") OR EXCLUDE (SUBJAREA, "ENVI") OR EXCLUDE (SUBJAREA, "MATE") OR EXCLUDE (SUBJAREA, "ENER") OR
EXCLUDE (SUBJAREA, "MULT") OR EXCLUDE (SUBJAREA, "AGRI") OR EXCLUDE (SUBJAREA, "CHEM") OR EXCLUDE (SUBJAREA, "ARTS") OR EXCLUDE (SUBJAREA, "ECON") OR
EXCLUDE (SUBJAREA, "EART") OR EXCLUDE (SUBJAREA, "Undefined") OR EXCLUDE (SUBJAREA, "VETE") OR EXCLUDE (SUBJAREA, "CENG") OR EXCLUDE (SUBJAREA, "PHYS")
OR EXCLUDE (SUBJAREA, "PSYC") OR EXCLUDE (SUBJAREA, "HEAL") OR EXCLUDE (SUBJAREA, "NURS")) AND (LIMIT-TO (LANGUAGE, "English")).

Data extraction

- Extract key information from the selected articles, such as:
  - author names
  - publication titles
  - publication years
  - abstracts
  - keywords and
  - citation counts

- Organize the data in a structured format for further analysis

Basic descriptive & Bibliometric Analysis

- Apply various bibliometric analysis techniques to explore patterns and trends within literature (Tools to help [VOSviewer](#), [CiteSpace](#), Scopus, and Web of Science).

Basic descriptive analysis

A total of 18,729 documents were analyzed including 14,444 (77.1%) classified as articles (journal publications) and 4284 (22.8%) conference paper

Summary of results.

Total number of papers	18,729
Total number authors	18,122
Total number of sources	155
Total number of affiliations	160
Total number of keywords	26555
Total number of subject areas	8
Total number of citations	386355
Total number of countries	158
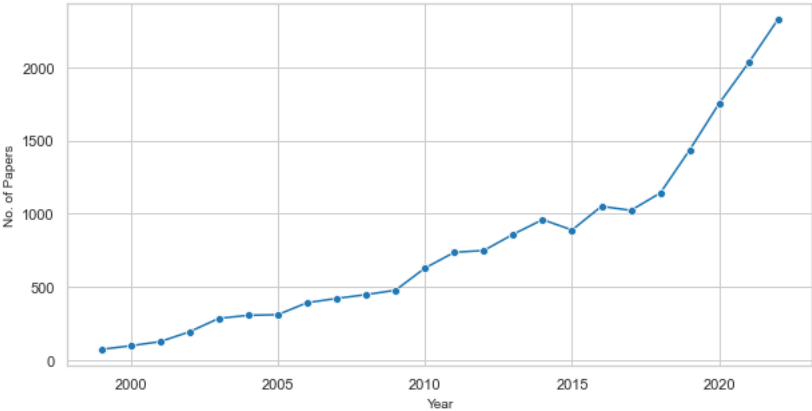
Year-wise publication of papers.
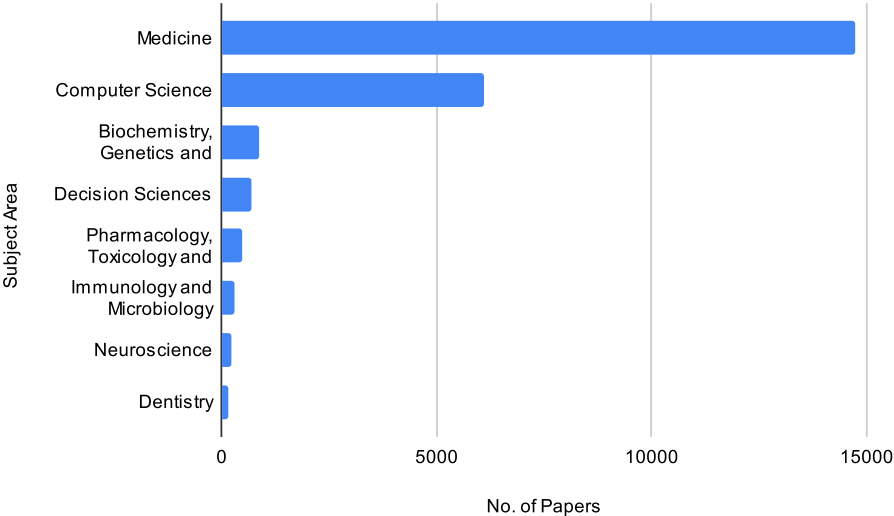
Subject area distribution of documents

Basic descriptive analysis

Most productive publishers

Document counts by countries
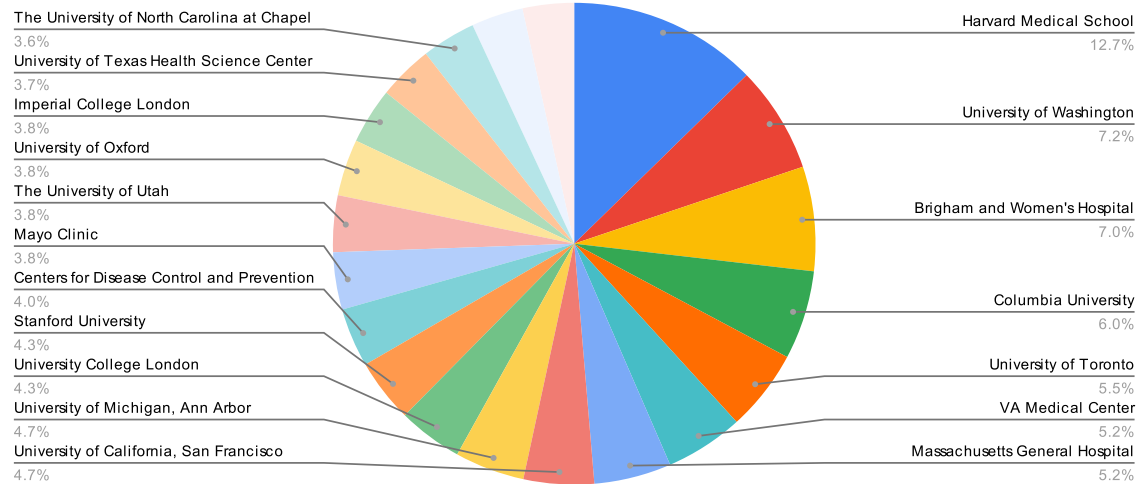
The top eighteen (18) affiliation (institutes) in technology for sustainable healthcare

Bibliometric Analysis

1. Citation analysis: Identify highly cited papers, influential authors, and journals within the field.
2. Co-authorship analysis: Analyze co-authorship networks to identify prolific authors and research collaborations. Explore the evolution of co-authorship patterns over time.
3. Keyword analysis: Identify frequently occurring keywords and their co-occurrence patterns. Visualize keyword timelines to understand the emergence and evolution of research topics.
4. Journal analysis: Analyze the distribution of publications across different journals and their impact factors. Identify top journals in the field and their contribution to the literature.
5. Country bibliographic coupling analysis: Provides a quantitative measure of the shared research interests and collaboration patterns among countries based on the common citations found within their respective scientific publications.

Citation Analysis

Identify highly cited papers, influential authors, and journals within the field.

Table 3: Productive papers in AIL literature.

S/N	Paper (Ref.)	Citations Counts						FWCI*7
		<2019	2019	2020	2021	2022	Total	
1.	(Hacker, 2018)	-	9	15	21	6	51	11
2.	(C.-L. Liu et al., 2004)	9	10	6	7	2	34	0.45
3.	(Robaldo & Sun, 2017)	4	11	5	7	-	27	2.66
4.	(Shih et al., 2008)	20	3	1	-	-	24	2.14
5.	(Hamledari & Fischer, 2021)		1	14	5	20	15.05

* = Scopus Figs as of May 1, 2022

Authors	No. of Papers (P)	Total Citations (TC.)	Avg. Citation (AC.)
Bates, D.W.	72	6093	84.625
Sittig, D.F.	55	2543	46.2364
Ash, J.S.	18	2260	125.5556
Middleton, B.	23	2076	90.2609
Ammenwerth, E.	27	2010	74.4444
Overhage, J.M.	18	1989	110.5
Blumenthal, D.	18	1948	108.2222
Hesse, B.W.	19	1817	95.6316
Wang, F.	36	1634	45.3889

S/N	Title	Ref	Pub. Year	Citations	Source title
1.	The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity	[67]	2010	2716	Arthritis Care and Research
2.	Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data	[68]	2005	1407	Nucleic Acids Research
3.	Utilization of the PICO framework to improve searching PubMed for clinical questions	[69]	2007	1374	BMC Medical Informatics and Decision Making
4.	Acute Physiology and Chronic Health Evaluation (APACHE) IV: Hospital mortality assessment for today's critically ill patients	[70]	2006	1164	Critical Care Medicine
5.	ConSORT-eHealth: Improving and standardizing evaluation reports of web-based and mobile health interventions	[71]	2011	1064	Journal of Medical Internet Research

Basic Descriptive Analysis

Top 10 publication venues ranked based on the number of publications.

#	Journal name	TP	SC	SNIP ᵃ	SJR	Publisher	H-Index
1	International Journal of Medical Informatics	841	9.5	2.019	1.197 Q1	Elsevier	122
2	Journal Of the American Medical Informatics Association	761	11.7	2.324	2.44 Q1	Oxford University Press	169
3	BMJ Open	485	4.4	1.321	1.059 Q1	BMJ Publishing Group	139
4	Journal of Biomedical Informatics	454	8.2	2.115	1.083 Q1	Academic Press Inc.	121
5	ACM International Conference Proceeding Series	405	*	*	0.2	ACM	137

TP = Total Publication; * = not applicable; Figure for 2022 provided by Scopus (SC = Cite Score; SNIP = Source Normalized Impact per Paper; SJR = Journal Rank), H-Index (Figure for 2022 provided by SCImagoJR)

Table 2: The 4 most influential and productive journals and conference proceedings in AIL research (2004–April 2022).

Source	TP	TC	Avg. citations	SJR*	SC	SNIP*	H-index	Country
ACM international conference proceeding series	8	29	3.625	0.182 (Q*)	1.2	0.296	123	United States
Artificial intelligence and law	4	73	18.25	0.856 (Q1)	7.5	3.81	34	Netherlands
International review of law, computers, and technology	3	14	4.6667	0.367 (Q2)	2.2	1.036	12	United Kingdom
Jusletter IT	14	13	0.9286	0.102 (Q4)	0.1	0.019	3	Switzerland

*TP = total publication; * = not applicable; SC = Cite Score; SNIP = Source Normalized Impact per Paper; SJR = SCImago Journal Rank; SNIP* = Fig. for 2020 provided by Scopus; SJR* = Fig. for 2020 provided by SCImagoJR; TC = Total Citations counts on AIL publications*

Hands-on

Using bibliometrics analysis to write a review papaer| I. K. NTI

Interpretation and Synthesis

✓Interpret the results of the bibliometric analysis and synthesize the findings.

✓Identify gaps, trends, and research directions within the literature.

✓Compare and contrast the results with existing theories, frameworks, or models in the field.

Reporting and Documentation

- Prepare a comprehensive report or research paper summarizing the bibliometric analysis.

- Include visualizations, tables, and figures to illustrate key findings.

- Clearly communicate the methodology, results, and implications of the study

Q & A

References

1. Nti, I. K., Boateng, S., Quarcoo, J. A., & Nimbe, P. (2023). Artificial Intelligence Application in Law: A Scientometric Review. Artificial Intelligence and Applications. https://doi.org/10.47852/bonviewAIA3202729

2. Nti, I. K., Adekoya, A. F., Weyori, B. A., & Keyeremeh, F. (2023). A bibliometric analysis of technology in sustainable healthcare: Emerging trends and future directions. Decision Analytics Journal, 100292. https://doi.org/10.1016/j.dajour.2023.100292

3. https://www.linkedin.com/pulse/unveiling-stages-proficient-bibliometric-literature-review-nti%3FtrackingId=Bpf53C7gTbu1Nx812xW9NA%253D%253D/?trackingId=Bpf53C7gTbu1Nx812xW9NA%3D%3D