eISSN : 2287-4577
pISSN : 2287-9099
https://www.jistap.org
Vol. 13 No. 4 December 30, 2025

Journal of Information Science Theory and Practice

Journal of Information Science Theory and Practice

JISTaP

Vol. 13 I No. 4 I December 30, 2025

KISTI www.kisti.re.kr

Indexed/Covered by SCOPUS, LISA, DOAJ, CrossRef, and KCI

JISTaP
Journal of Information Science Theory and Practice

General Information

Aims and Scope

The Journal of Information Science Theory and Practice (*JISTaP*) is an international journal that aims at publishing original studies, review papers and brief communications on information science theory and practice. The journal provides an international forum for practical as well as theoretical research in the interdisciplinary areas of information science, such as information processing and management, knowledge organization, scholarly communication and bibliometrics. JISTaP will be published quarterly, issued on the 30th of March, June, September, and December. JISTaP is indexed in the Scopus, Korea Citation Index (KCI), Directory of Open Access Journals (DOAJ), Korea Science Citation Index (KSCI) and KoreaScience by the Korea Institute of Science and Technology Information (KISTI) as well as CrossRef. The full text of this journal is available on the website at http://www.jistap.org

Indexed/Covered by

Scopus LISA DOAJ DIRECTORY OF OPEN ACCESS JOURNALS Crossref Cited-by 한국학술지인용색인 Korea Citation Index

Publisher

Korea Institute of Science and Technology Information
66, Hoegi-ro, Dongdaemun-gu, Seoul, Republic of Korea
(T) +82-2-3299-6105
(F) +82-2-3299-6067
E-mail: jistap@kisti.re.kr
URL: http://www.jistap.org

Managing Editor: Hyekyong Hwang, Eungi Kim

Copy Editor: Ken Eckert

Design & Printing Company: MEDrang Inc.

2nd Floor, Hyoryeong Building, 32 Mugyo-ro, Jung-gu, Seoul 04521, Korea
(T) +82-2-325-2093
(F) +82-2-325-2095
E-mail: info@medrang.co.kr
Homepage: www.medrang.co.kr

Open Access and Creative Commons License Statement

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee or authors' permission. Open Access articles are automatically archived in the Korea Institute of Science and Technology Information (KISTI)'s Open Access repository (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

This paper meets the requirements of KS X ISO 9706, ISO 9706-1994 and ANSI/NISO Z39.48-1992 (Permanence of Paper).

2025 Copyright © Korea Institute of Science and Technology Information

JISTaP
Journal of Information Science Theory and Practice

Editorial Board

Co-Editors-in-Chief

Gary Marchionini
University of North Carolina, USA

Dong-Geun Oh
Keimyung University, Korea

Co-Associate Editors

Kiduk Yang
Youngnam Old Document Archive Center, Korea

Hyejin Lee
Korea Institute of Science and Technology Information, Korea

Co-Managing Editors

Hyekyong Hwang
Korea Institute of Science and Technology Information, Korea

Eungi Kim
Keimyung University, Korea

Editorial Board

Dan Albertson
University at Buffalo, SUNY, USA

Daniel Martínez Ávila
São Paulo State University, Brazil

Beeraka Ramesh Babu
University of Madras, India

Pia Borlund
University of Copenhagen, Denmark

France Bouthillier
McGill University, Canada

Kathleen Burnett
Florida State University, USA

Ina Fourie
University of Pretoria, South Africa

Boryung Ju
Louisiana State University, USA

Thomas Mandl
Universität Hildesheim, Germany

Lokman I. Meho
American University of Beirut, Lebanon

Jin Cheon Na
Nanyang Technological University, Singapore

Dan O'Connor
Rutgers University, USA

M. P. Satija
Guru Nanak Dev University, India

Ou Shiyan
Nanjing University, China

Paul Solomon
University of South Carolina, USA

Kulthida Tuamsuk
Khon Kaen University, Thailand

Lihong Zhou
Wuhan University, China

Xiao-Zhi Gao
University of Eastern Finland, Finland

Wole Michael Olatokun
University of Ibadan, Nigeria

Consulting Editors

Reysa Alenzuela
Oriental Institute, ASCR, Czech Republic

Wayne Buente
University of Hawaii, USA

Sujin Butdisuwan
Mahasarakham University, Thailand

Seon Heui Choi
Korea Institute of Science and Technology Information, Korea

M. Krishnamurthy
DRTC, Indian Statistical Institute, India

S.K. Asok Kumar
The Tamil Nadu Dr Ambedkar Law University, India

P. Rajendran
SRM University, India

2025 Copyright © Korea Institute of Science and Technology Information

JISTaP
Journal of Information Science Theory and Practice

Table of Contents

Vol. 13 No. 4 December 30, 2025
Journal of Information Science Theory and Practice • https://www.jistap.org

Articles

2025 Copyright © Korea Institute of Science and Technology Information

JISTaP https://www.jistap.org
Journal of Information Science Theory and Practice

RESEARCH PAPER
J Inf Sci Theory Pract 13(4): 1-21, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.4.1

How to Describe the Life's Work of a Scientist? The Œuvre of Christian Schlögl as an Example

Wolfgang G. Stock*
Department of Information Science, Heinrich Heine University Düsseldorf, Düsseldorf, Germany
E-mail: wolfgang.stock@hhu.de

Gerhard Reichmann
Department of Operations and Information Systems, University of Graz, Graz, Austria
E-mail: gerhard.reichmann@uni-graz.at

ABSTRACT

This article attempts to present the entire research output of a scientist in terms of his publication activity and its impact as comprehensively as possible. To this end, a bibliometric overview of the scope and structure of his research output is first provided. Furthermore, the visibility of this output in the various information services is analyzed. Christian Schlögl authored a total of 177 publications, including 77 journal articles. This set of publications is only visible through the use of personal publication lists in an institutional repository (Uni Graz Online). In contrast, only a fraction of these publications are included in the common information services; in the Web of Science, for example, not even a third exist. The publication output is then examined in more detail, including preferred topics, co-authors, and journals. Next, a citation analysis is conducted, revealing, for example, the temporal distribution of citations and the most frequently cited publications. Finally, the seven most important research areas are briefly presented. For each of these areas, a co-author was asked to comment on Christian Schlögl's working methods and collaboration. Overall, this article could serve as a good illustrative example for analyzing the research performance of an individual researcher.

Keywords: oeuvre, researcher, scientometric micro level, bibliometrics, research activities, Christian Schlögl

Received: May 3, 2025
Accepted: September 2, 2025
Revised: August 21, 2025
Published: December 30, 2025

***Corresponding Author:** Wolfgang G. Stock
https://orcid.org/0000-0003-2697-3225
E-mail: wolfgang.stock@hhu.de

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee or authors' permission. Open Access articles are automatically archived in the Korea Institute of Science and Technology Information (KISTI)'s Open Access repository (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

© 2025 [Wolfgang G. Stock, Gerhard Reichmann] This article is distributed under the Creative Commons Attribution License (CC BY), allowing unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. It is published by the Korea Institute of Science and Technology Information (KISTI).

1. INTRODUCTION

How does one describe and analyze the oeuvre of a particular researcher in terms of publication activity and its impact? Should we use multidisciplinary bibliographical information services, a reasonably complete publication list, interviews with peers and colleagues, or all of them? We want to address this topic using a concrete example. Our case study focuses on the research of Christian Schlögl (1961-2024), an Austrian information scientist who was one of the most influential information scientists in the German-speaking world at the beginning of the twenty-first century. He authored nearly 180 publications and was a highly active member of the scientific community. He maintained international contacts with colleagues from Croatia, Germany, Great Britain, Hungary, the United States, and South Korea, often resulting in publications. In addition, he participated in countless scientific conferences as a co-organizer, reviewer, presenter or moderator, regularly reviewed articles for scientific journals, and was also a member of the editorial boards of several relevant journals, such as *Journal of Information Science Theory and Practice (JISTaP)*, published in Korea. As Stefan Dreisiebner told us,

> Christian was an important pillar for the Austrian Information Science community. He leaves a gap that is hard to fill. One example: In a recent project, we aimed to collect all Austrian researchers with information science backgrounds. His unique knowledge of the community and his contacts were [badly] missed in this endeavor (personal communication, March 26, 2025).

In addition to research, he was also very committed to teaching and administration. In this context, the development of the international master's degree program "Global Studies on Management and Information Science", in cooperation with a team from Hildesheim, Germany, and Pai Chai University in Daejeon and Chungbuk National University in Cheongiu, both in Korea, should be highlighted. However, the present article is only about Schlögl's research (for other aspects, e.g., on academic teaching, see Reichmann, 2024).

What is the current state of research in describing and analyzing a researcher's professional life work? In the sense of levels of scientometrics, we are working on the micro level, that is, the level of individual researchers (Sandström & Sandström, 2009), in contrast to the meso level (research institutions) and the macro level (countries or regions of research). Following White (2001), we will build an author-centred bibliometric description with the characterizations automatically made and edited online (White's CAMEO). White's bibliometric measures include the subjects of a researcher (e.g., through subject headings or title terms), the publishing journals and publication years, the researcher's co-authors, and– only applicable in citation databases—the researcher's citation identity (all cited authors), the citation image makers (all citing authors), and the citation image (all co-cited authors). The scope of bibliometric analysis options depends heavily on the information service and the databases offered. White worked with the online service DIALOG and the databases of the former Institute for Scientific Information (Science Citation Index and related databases).

Several bibliometricians followed White's proposal and produced scientometric results concerning the micro level on, e.g., Leydesdorff (Vargas-Quesada et al., 2023), Drucker (Uslay et al., 2009), or Stock (Peters et al., 2020). Other studies are limited to a researcher's production, collaboration, and impact (Ayala-Gascón et al., 2012) or additionally include the topics of the papers of a researcher, such as those of Wilhelm M. Frankl (Stock, 1988) or all the members of the Graz School of philosophy and psychology (Stock & Stock, 1990), including a topic analysis also of the citing works. Jacso (2018), in a bibliometric study on Eugene Garfield, preferred to study the citations to Garfield's works, their sources, the citing institutions, and their countries.

However, we should handle informetric, bibliometric, and scientometric studies carefully, as there are methodological problems and incomplete databases (Stock et al., 2023b). Following the literature (e.g., Stock & Weber, 2006), we speak of "informetrics," if all quantitative aspects of information science are concerned, of "bibliometrics" (as a narrower term of "informetrics"), if written published works are quantitatively analyzed, and of "scientometrics," if we describe scientific research in quantitative ways.

Nevertheless, describing and assessing researchers' performance is "vital in the research community" (Rawat et al., 2024, p. 319). It seems practical and promising to create bibliometric-based heuristics (Bornmann et al., 2022), that is, bibliometric results to produce an overview and give hints on the essential topics of a researcher. In addition to bibliometric descriptions, we apply expert interviews (Meuser & Nagel, 2009). The experts interviewed are proven representatives of their field and co-authors of some of Schlögl's publications.

This paper aims to describe and analyze Schlögl's re-

Table 1. Languages and document types of Schlögl's oeuvre

	Journal articles	Articles in proceedings or edited book	Lexicon entries	Book reviews	Books	Editions	Research data	Total
German	44	16	43	4	1	2	0	110
English	33	31	0	0	0	2	1	67
Total	77	47	43	4	1	4	1	177

Data source: Uni Graz Online, updated; data as of February 2025.

Table 2. Visibility of Schlögl's oeuvre in information services

Information service	Publications	Citations	H-index
Web of Science Core Collection	53	547	11
Scopus	53	673	13
Semantic Scholar[a)]	96	862	-
Dimensions	27	222	-
ResearchGate	78	932	14
ORCiD	30	-	-
For comparison: Uni Graz Online[b)]	177	-	-

Data as of February 2025.
[a)]Aggregation of different entries due to synonymy errors.
[b)]As of June 2024; updated by the authors.

Table 3. Visibility of Schlögl's English-language article publications in selected information services

Information service	Publications	Relative visibility (%)
Web of Science Core Collection	29	45
Scopus	34	53
Uni Graz Online	64	100

Data as of February 2025: only including research articles in journals, proceedings and edited books.

search production (his publications), research impact (the citations the publications attracted), and research topics. In the first step, in Sections 2 to 4, we apply quantitative informetric methods to get a heuristic overview of Schlögl's publications, their topics, and their impact; in the second step, in Section 5, we will shortly sketch Schlögl's main topics, which we found in the informetric results by publication, citation, and topic analyses, and, additionally by the expert interviews. Finally, we will assess the informetric methods used at the micro level.

2. AUTHOR-CENTERED BIBLIOMETRICS OF SCHLÖGL'S OEUVRE: BASIC DATA AND VISIBILITY

We start with a bibliometric overview of Schlögl's oeuvre. In detail, the quantitative overview will include basic data on Schlögl's publications and their visibility in selected information services, his co-authors, the publishing journals, his citation image (that is, the cited authors), the time series of his publications and their citations, the main research topics, Schlögl's citation image makers (that is, the citing authors), their countries and their topics, and, finally, Schlögl's top-cited publications.

Christian Schlögl published 177 papers between 1992 and 2025, of which 62% were in German and 38% in English (Table 1). All raw data is as of February 2025. Most papers were research articles published in journals (44% of all papers) or in proceedings or edited books (27%). German predominates in his journal articles, but the opposite is true for his articles in proceedings and edited books. Schlögl published one book (Schlögl, 2001) and co-edited four proceedings or edited books. Additionally, Schlögl wrote 43 short entries in a lexicon on Library and Information Science (LIS) (Umlauf & Gradmann, 2011/2014), four book reviews, and one set of research data.

An important data source on Schlögl's publications is the research portal of the Karl Franzens University Graz, Austria (Uni Graz Online). The University of Graz was his affiliation during his complete academic life. We worked with a printed version of Schlögl's publication list from mid-2024 for this paper. After his death, further articles appeared, so we updated the list based on Uni Graz Online. This updated version of Schlögl's publication list from Uni Graz Online is our gold standard for all scientometric measures, hoping that this list is indeed correct and true-bounded (Dorsch et al., 2018), which means that the list is complete and that all listed documents are indeed research contributions and authored by Schlögl.

Regarding the visibility (Dorsch, 2017) of Schlögl's work in large multidisciplinary databases, identifying

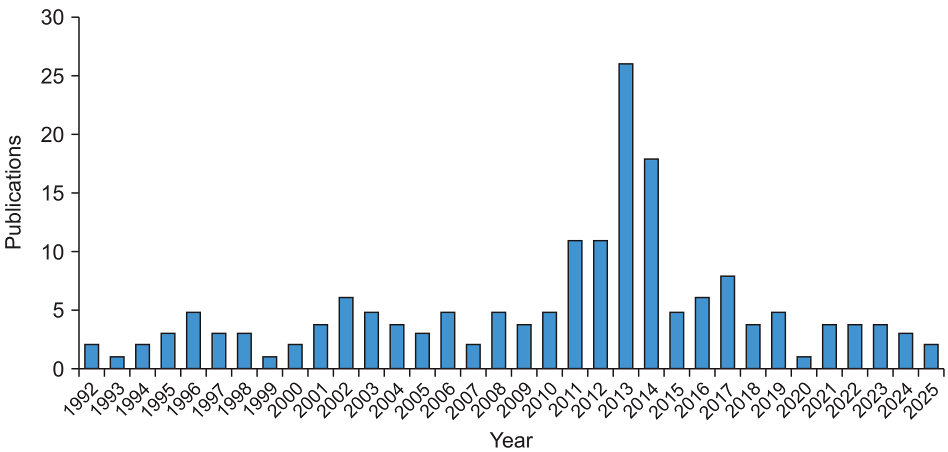

Fig. 1. Time series of Schlögl's publications (1992-2025). Data source: Uni Graz Online, updated; n=177 publications; data as of February 2025.

Table 4. Schlögl's preferred journals

Journal	Number of articles by Schlögl
Information – Wissenschaft und Praxis	20
Scientometrics	7
Zeitschrift für Bibliothekswesen und Bibliographie	4
VÖB Mitteilungen	3
Journal of Information Science Theory and Practice	2
Journal of Informetrics	2
Journal of the American Society for Information Science and Technology	2
Wirtschaftsinformatik	2

Data source: Uni Graz Online, updated; all journals with two or more articles by Schlögl; data as of February 2025.

Table 5. Schlögl's preferred co-authors

Name of co-author	Country of co-author	Number of joint publications
Gerhard Reichmann	Austria	16
Wolfgang G. Stock	Germany	14
Juan Gorraiz	Austria, Spain	10
Kris Jack	UK	6
Peter Kraker	Austria	6
Wolf Rauch	Austria	6
Sandra Boric	Austria	5
Isabelle Dorsch	Germany	5
Christian Gumpenberger	Austria	5
Karin Karlics	Austria, Italy	5
Edgar Schiebel	Austria	4
Stefan Dreisiebner	Austria	3
Robert M. Hayes	United States	3
Franjo Pehar	Croatia	3
DongBack Seo	South Korea	3

Data source: Uni Graz Online, updated; all co-authors with at least three joint publications; data as of February 2025.

Schlögl's publications in information services was difficult. Some services work with identification numbers. In Scopus, Web of Science Core Collection (WoS CC), ResearchGate, and ORCiD, we found one dataset on Schlögl (Scopus) or a manageable number of entries. Scopus works with one identification number (AU-ID=6506241000), while WoS unfortunately uses four different numbers for the same author (A-9343-2008, BBB-9858-2020, JKG-5289-2023, ITM-8469-2023). Semantic Scholar presents many publications by Schlögl, but due to synonym errors, these appear under different entries, which we have intellectually merged. In Dimensions (and similarly in Google Scholar), the problem of homonymy has struck, as many Christian Schlögls are almost inseparably mixed together in the search results lists. Due to incomplete data, Table 2 shows only approximately correct Dimension values. Although there is an entry on Christian Schlögl on Google Scholar, there are inextricably many different Schlögls in a single publication list, so we cannot present a satisfying result.

The most productive information services for Schlögl's oeuvre are Semantic Scholar (96 hits; 54% of the gold standard), ResearchGate (44%), WoS CC (30%), and Scopus (30%), meaning that all mentioned services are more or less incomplete or incorrectly indexed. Therefore, the values for the publications, citations, and H-index in Table

2 can only be viewed as a lower estimate of the "true" results.

We searched again because some multidisciplinary information services prefer to index English-language research articles. We added two search arguments to the name: language (only English) and document type (only articles). We were able to perform such searches in WoS CC and Scopus. Table 3 shows that Schlögl's relative visibility in both WoS CC and Scopus is still limited as there are relative visibility values in Scopus of 53% and WoS CC of only 45% (WoS CC) for all his English-language articles in journals, proceedings, and edited books.

3. RESEARCH PRODUCTION

Schlögl was active in publishing between 1992 and his death in 2024. As a result of the delay in the publication process of some writings, documents are being published even after Schlögl's passing (Fig. 1). There are peaks in the annual publication amount between 2011 and 2014 resulting from publishing short entries in a LIS lexicon. Apart from these short documents, Schlögl published about three to five documents annually on average.

Schlögl published in both German-language and English-language journals (Table 4). His German-language articles concentrate on *Information – Wissenschaft und Praxis (IWP)*, *Zeitschrift für Bibliothekswesen und Bibliographie*, *VÖB Mitteilungen* (Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen & Bibliothekare), and *Wirtschaftsinformatik*. *IWP* is primarily dedicated to information science and practice, while the following two periodicals mentioned above focus more on library topics. *Wirtschaftsinformatik* is dedicated to research in information systems. The English-language articles are spread across many different journals, including *Scientometrics*, *JISTaP*, *Journal of Informetrics*, and *Journal of Association for Information Science and Technology* (formerly *Journal of the American Society for Information Science and Technology*).

Besides his early works and encyclopedia contributions, Schlögl often worked and published in teams. His preferred co-authors include the two authors of this article and Juan Gorraiz from the University of Vienna, Austria (Table 5). To get an overview of Schlögl's co-authors, we

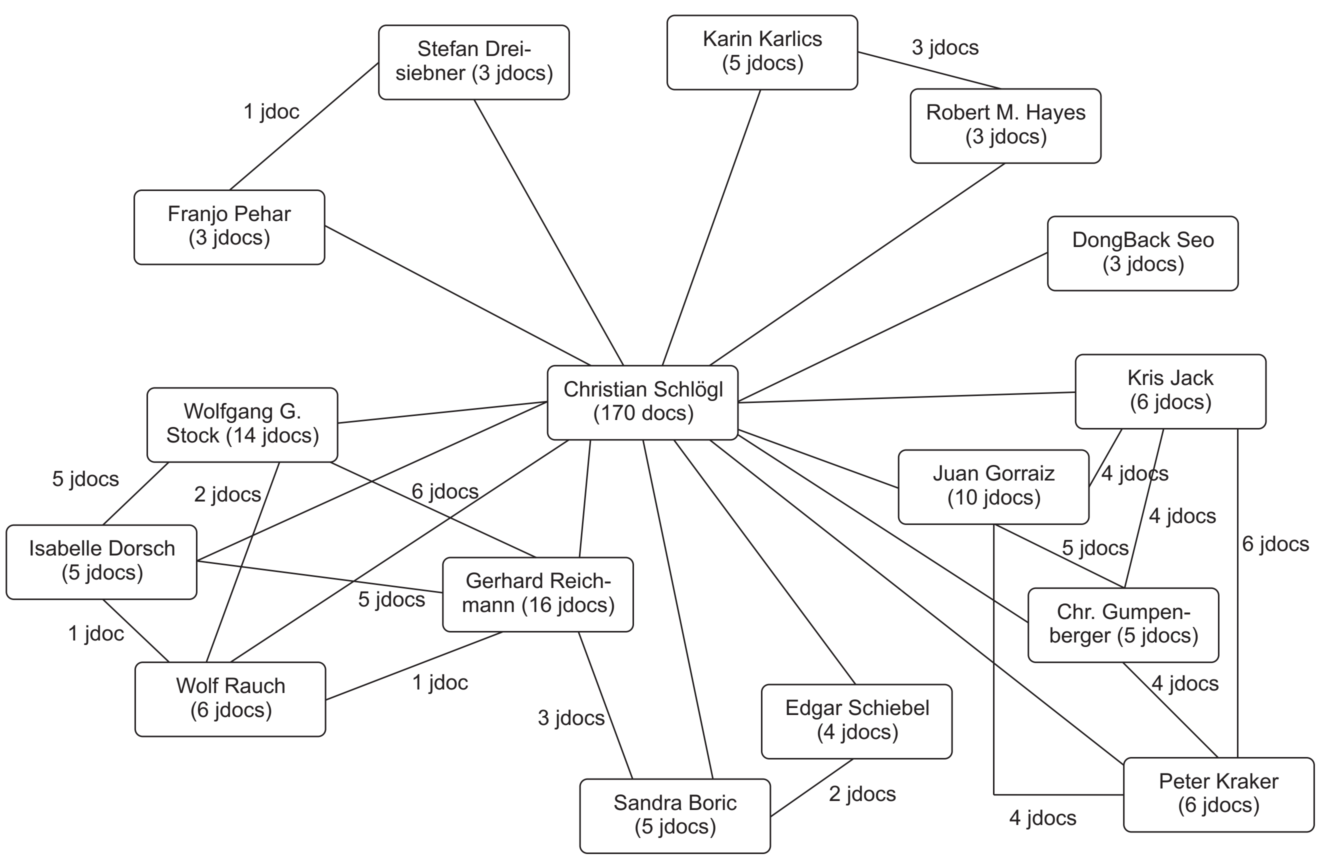

Fig. 2. Cluster of Schlögl's co-authors. jdocs, joint co-authorship documents.

Table 6. Schlögl's citation identity: top-cited authors in Schlögl's articles

Name of cited author	Number of citing articles
Christian Schlögl	36
Wolfgang G. Stock	18
Juan Gorraiz	17
Eugene Garfield	14
Ronald Rousseau	14
Mike Thelwall	11
Judit Bar-Ilan	10
Isabelle Dorsch	10
Christian Gumpenberger	10
Isabella Peters	9
Stefanie Haustein	8
Gerhard Reichmann	8
Ludo Waltman	8
Blaise Cronin	6
Wolfgang Glänzel	6
Chris Jack	6
Peter Kraker	6

Source: Scopus; n=53 articles by Schlögl; all selected authors with six or more mentions in Schlögl's reference lists; data as of February 2025.

created a cluster of his co-authors with whom he frequently collaborated (Fig. 2). Here we see six different author subclusters (A1 to A6):

(A1) Schlögl – Dreisiebner – Pehar,
(A2) Schlögl – Karlics – Hayes,
(A3) Schlögl – Gorraiz – Jack – Kraker – Gumpenberger,
(A4) Schlögl – Reichmann – Boric – Schiebel,
(A5) Schlögl – Reichmann – Stock – Dorsch – Rauch, and
(A6) Schlögl – Seo.

The description of Schlögl's citation identity (White, 2001), that is, authors cited by Schlögl, requires using a citation database. We chose Scopus because there is a search field for authors in cited references (REFAUTH). To find all self-citations of Schlögl, for instance, we worked with the following search argument:

AU-ID("Schlögl, Christian" 6506241000) AND (REFAUTH(Schlögl) OR REFAUTH(Schlogl) OR REFAUTH(Schloegl)).

Applying the REFAUTH field, we had to search for all referenced authors individually, as there is no ranking functionality for cited authors in Scopus. Since we cannot be sure that we have searched for all highly cited authors in Schlögl's publications, we should use the results cautiously. Schlögl's top-cited authors are Schlögl himself (as

Table 7. Top-cited publications of Schlögl in different information services

No.	Short title (year of publication)	Rank (number of citations)		
		Scopus	WoS CC	Google Scholar
1	*A bibliometric analysis of pharmacology and pharmacy journals* (2008)	1 (84)	1 (79)	1 (175)
2	*Comparison of citation and usage indicators* (2010)	2 (81)	2 (73)	4 (130)
3	*Impact and relevance of LIS journals* (2004)	3 (75)	3 (62)	3 (136)
4	*Comparison of downloads, citations and readership data* (2014)	4 (67)	5 (51)	6 (87)
5	*Usage vs citation behavior* (2014)	5 (62)	4 (52)	5 (92)
6	*Global usage vs global citation metrics* (2011)	6 (48)	6 (42)	7 (72)
7	*Information and knowledge management* (2005)	7 (35)	8 (26)	2 (148)
8	*Practitioners and academics* (2008)	8 (33)	7 (31)	8 (53)
9	*Download vs citation vs readership data* (2013)	9 (24)	10 (16)	-
10	*LIS journals: an editor survey* (2005)	10 (22)	9 (19)	9 (46)
11	*Visualization of co-readership* (2015)	-	-	10 (43)

Top ten cited articles of each information service; data as of February 2025.
Wos CC, Web of Science Core Collection; LIS, Library and Information Science.

expected, since Schlögl published very actively in his research areas), Stock, Gorraiz, Garfield, etc. (Table 6). The cited authors from Table 6 are either citation authorities in bibliometrics (e.g., Garfield, Rousseau, Thelwall, Bar-Ilan, or Cronin) or colleagues who knew Schlögl personally, as many of them were his co-authors (Table 5). Another problem with Scopus is that the number of Schlögl's citing articles is stated here, regardless of how many articles by the cited author are mentioned in each document. For instance, in Schlögl and Stock (2004), four different works by Eugene Garfield are cited, but Scopus counts "1."

We decided to analyze the title terms to describe the content of Schlögl's publications. Here, we limited ourselves to the English-language titles to avoid translation errors and to the knowledge that only Schlögl's English-language papers are often cited (see Table 7). Schlögl's top topics are *citation* and *journal*, followed by *information literacy*, *readership*, and *usage* (Table 8).

In Fig. 3, we have created a cluster starting with Schlögl's two most frequent English-language title terms, *citation* and *journal*. From this, it is possible to identify four topic subclusters (T1 to T4):

(T1) citation – journal – downloads – readership,
(T2) citation – usage – behavior,
(T3) journal – readership – LIS, and
(T4) citation – journal – usage – Scopus – WoS – pharmacology.

Table 8. Most frequent terms in the titles of Schlögl's English-language articles

Title term	Frequency
Citation	15
Journal	13
Information literacy	7
Readership	7
Usage	6
Austria	5
Bibliometrics	5
Downloads	5
Graz	5
Business administration	4
Pharmacology	4
Research evaluation	4
Research institution	4
Web of Science	4
Autonomous driving	3
Behavior	3
Information science	3
Library and Information Science	3
Scopus	3
Visualization	3

Data source: Uni Graz Online; n=64; all English-language title terms with three or more mentions; data as of February 2025.

4. RESEARCH IMPACT

Scientometric analyses of research impact normally work with citation studies. The problem is that we are dependent on the large citation databases including WoS and Scopus (Pranckutė, 2021), which are unfortunately all incomplete in coverage of the research literature and partially only weakly indexed (Stock et al., 2023b). As a result, our data are only approximations of more accurate results.

If we analyze Schlögl's impact in terms of citations of his publications based on Scopus, we see a slow but steady increase in the annual numbers of citations until around 2013 (Fig. 4). Between 2014 and 2024, his Scopus-indexed publications attracted about 40 to 50 citations yearly (except in 2022, with fewer than 30 citations). All in all, Scopus counts 673 citations for Schlögl's oeuvre.

Who did Schlögl's work cite? In the sense of White (2001), we asked for Schlögl's citation image makers. For this, we conducted a cited reference search in WoS CC and found more than 200 authors citing aspects of Schlögl's oeuvre (Table 9). Here, we did not use Scopus but rather WoS because there is elaborated cited reference search functionality.

The top citation image makers are Gorraiz (who is one of Schlögl's preferred co-authors and also one of his top-cited colleagues), Schlögl himself (citing some earlier publications in newer articles), Gumpenberger, Costas, Glänzel, Reichmann, Thelwall, Haustein, Stock, and nine other authors with four to five citations each.

If we compare the lists of Schlögl's preferred co-authors (Table 5), his most cited authors, that is, his citation identity (Table 6), and the most citing authors, that is, his citation image makers (Table 9), we see a remarkable similarity between the top positions of all three lists. Social circles are behind Schlögl's references and citations (Milard, 2014). Schlögl often cited and was often cited by personally known colleagues. Those colleagues work, for instance, at the University of Graz (Reichmann and Boric) or the

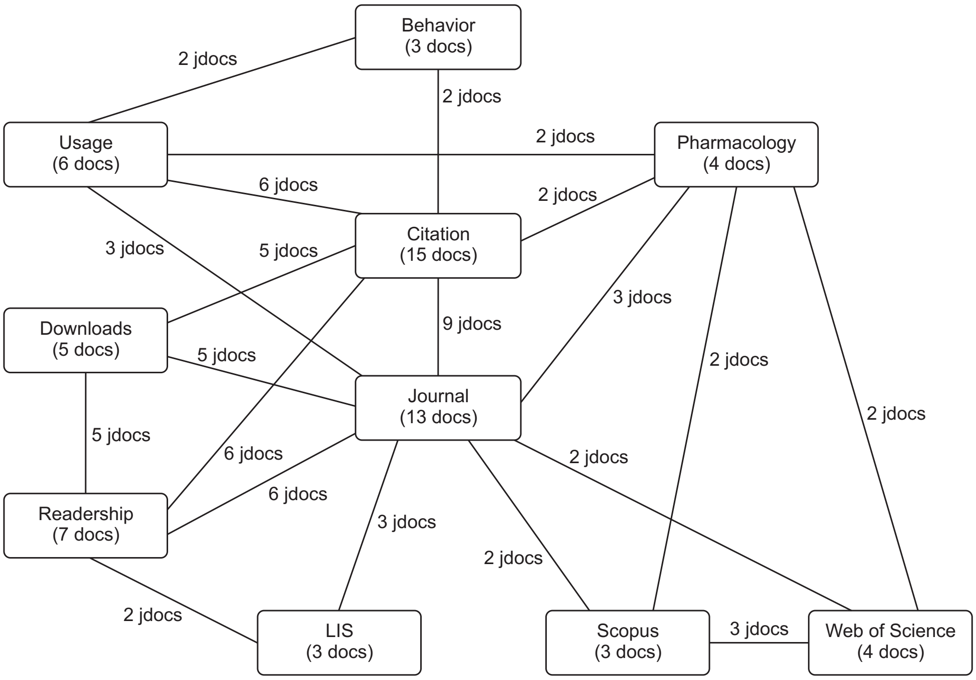

Fig. 3. Cluster of Schlögl's topics around *citation* and *journal*. Topics from Table 8, which co-occur with *citation* or *journal* in three or more English-language document titles; in brackets: number of publications in Schlögl's oeuvre; jdocs: number of co-occurrences of topics in titles; data as of February 2025. LIS, Library and Information Science.

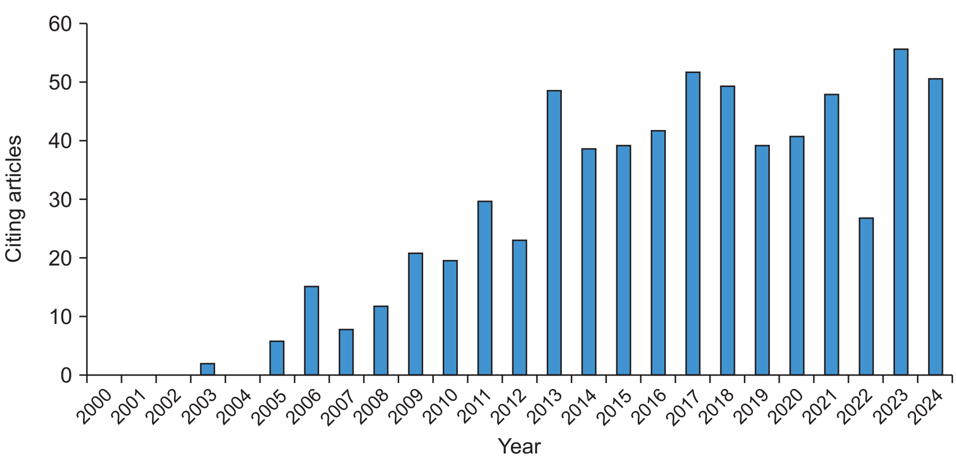

Fig. 4. Time series of citations to Schlögl's publications (2000-2024). Data source: Scopus; n=673 citations; data as of February 2025.

Know Center in Graz (Kraker), in other Austrian cites (e.g., Gorraiz in Vienna), or in institutions all around the world (e.g., Hayes in the USA, Jack in the UK, Sutheo in Hungary, Seo in Korea, Pehar in Croatia, and Stock as well as Dorsch in Germany). Overall, the top-cited and top-citing authors were socially connected to Schlögl (Johnson & Oppenheim, 2007). "Who will cite you back?" Daud et al. (2017) ask. For Schlögl's oeuvre, there is a clear answer: The reciprocal citation behavior of the researchers around Schlögl is strongly influenced by their co-authorships. Cronin and Shaw (2002) found similarities between the authors' citation identities and their citation image makers in their case studies; we have to add the vital role of the co-authors.

In Table 10, we show in which countries Schlögl's research results took effect. The citation image makers' countries of activity include countries from all over the world, starting with Austria and followed by the United States, Germany, China, England, the Netherlands, Canada, and many others.

Using a very rough classification, WoS offers research topics of the citing literature. Schlögl's citation image is dominated by *bibliometrics* and—to a much lesser extent—by *information literacy* and—even less—by *knowledge management* and *entity resolution* (Table 11). Obviously, his primary influence on the international research

Table 9. Schlögl's top citation image makers

Citing author	Number of articles citing Schlögl
Juan Gorraiz	19
Christian Schlögl	18
Christian Gumpenberger	10
Rodrigo Costas	7
Wolfgang Glänzel	7
Gerhard Reichmann	7
Mike Thelwall	7
Stefanie Haustein	6
Wolfgang G. Stock	6
Vincent Larivière	5
Pei-Shan Chi	4
Stefan Dreisiebner	4
Zhichao Fang	4
Peter Kraker	4
Isabella Peters	4
Cassidy R. Sugimoto	4
Daniel Torres-Salinas	4
Martin Wieland	4

Data source: Web of Science CC, cited reference search; n=235 citing authors; all authors with four or more articles citing Schlögl; data as of February 2025.

Table 10. Countries of origin of Schlögl's citation image makers

Country of origin of citation image maker	Number of articles citing Schlögl
Austria	41
USA	33
Germany	28
China	26
England	21
The Netherlands	16
Canada	15
Australia	10
Belgium	7
India	7
Malaysia	7
Hungary	6
South Africa	6

Data source: Web of Science CC, cited reference search; n=235 citing authors; all countries with six or more articles citing Schlögl; data as of February 2025.

Table 11. Broad topic classification of articles citing Schlögl

Web of Science topics	Number of articles citing Schlögl
Bibliometrics	162
Information literacy	27
Knowledge management	7
Entity resolution	5

Data source: Web of Science CC, cited reference search; n=235 citing authors; all WoS topics with five or more articles citing Schlögl; data as of February 2025.

community was (and still is) in *bibliometrics*.

In Table 7, we get a more detailed picture of Schlögl's research impact. We searched for the most cited publications by Schlögl in Scopus, WoS CC, and Google Scholar. Google Scholar revealed the most citations, followed by Scopus and WoS CC; however, the ranks of the top-cited publications are similar in the three information services. An exception is publication no. 7 on information and knowledge management, which ranks second on Google Scholar but seventh and eighth on Scopus and WoS.

Schlögl's most cited publication is "*A bibliometric analysis of pharmacology and pharmacy journals*" (published in 2008), followed by "*Comparison of citation and usage indicators*" (2010) and "*Impact and relevance of LIS journals*" (2004). Because of their age, older publications have a greater chance of being cited than articles from recent years. Therefore, Table 7 is necessarily biased toward older writings. Apart from publication no. 7, which is on information and knowledge management, all frequently cited papers of Schlögl (from Table 7) are in the field of bibliometrics, which confirms the result from Table 11 with bibliometrics as the top-citing research branch.

The citations of Schlögl's publications form an inverse-logistic distribution (Stock, 2006), meaning that there are several frequently cited documents and—below the turning point of the curve—many documents with low citation rates in the form of a long tail. A power-law distribution would have been more expected, as it is the most common in informetrics. There would be one highly cited work at the top of the curve and a steep drop in citation rates. In the context of indexing and retrieval, Peters and Stock (2010) speak about "power tags," meaning the top terms of an inverse-logistic distribution of terms assigned to a document in an information service supporting user-created tags. In our case (Table 7), we can identify "power publications" that are the author's most cited and, thus, most influential works.

5. RESEARCH TOPICS

What were Schlögl's main research topics? We applied the scientometric results to find heuristics on essential themes. In this overview of a complete oeuvre of a researcher, we can only give brief information on the content of Schlögl's research. We will concentrate on the research questions, applied methods, and main results. We also obtained comments on the research style of and collaboration with Christian Schlögl from former co-authors in the respective research area. However, examining the individual research aspects in more detail is a task for future studies.

5.1. Library and Information Science Journals: Authors, Readers, and Editors

We found this theme through articles no. 3, 8, and 10 in Table 7, topic cluster T3, and through parts of author cluster A5 (however, only connected with Stock). Here, the co-authors are Stock and a group of students of Graz and Cologne, who published under the pseudonym "Grazia Colonia." The findings were published between 2002 and 2008.

This project aimed to analyze the relationships between citations and references to scientific journals on the one hand and the readership of the journals on the other (Grazia Colonia, 2002). Subjects of the study were international and German-language LIS periodicals. Schlögl and his research team formulated two research questions, namely (1) "What is the impact of LIS journals?" and (2) "What is the relevance of LIS periodicals to their readers?" (Schlögl & Stock, 2004, p. 1156). Grazia Colonia applied two empirical methods. Besides citation analysis, they worked with questionnaires. Data from the 40 international LIS journals were collected from the Journal Citation Reports, while the data for the German-language periodicals were counted manually—1,494 source articles with 10,520 references. The response rate of the reader survey with 257 usable questionnaires from German-speaking information specialists was relatively high. The main results include a negative correlation (-0.11) between the impact factor and the reading frequency for all analyzed 50 journals and a high correlation (+0.70) between the regional impact factor and the reading frequency of the 10 German-language LIS journals. There were huge differences between the reading behavior of researchers and practitioners (Schlögl & Stock, 2008). The correlation between the regional impact factor of the ten German-language periodicals and the reading frequency was +0.71 for the practitioners and -0.17 for the scientists because the latter preferred to read international journals and not German-language papers. Schlögl checked the results by conducting an empirical survey of editors; however, the editors confirmed the results concerning practitioners and scientists in their roles as readers as well as authors (Schlögl & Petschnig, 2005).

At the end of the project, Juchem et al. (2006, p. 32) determined five basic dimensions of the scientometrics of journals: (1) production (papers and their authors), (2) content (topics), (3) reception (readers), (4) formal research communication (references and citations), and (5) editorial work (editorial policies and publishing houses). These facets provided a template for the structure of Haustein's famous book on multidimensional journal evaluation (Haustein, 2012).

Kerstin Wolff, née Juchem, was student at the University of Applied Sciences in Cologne, Germany, at this time and was one of the participants in the Grazia Colonia project. Now she is the head of two public school libraries in Kerpen, Germany. Wolff states,

> As part of the 'Grazia Colonia' group and beyond, I had the opportunity to work with Christian Schlögl. Initially separated by the two location-based groups in Graz and Cologne and by the various tasks, one thing in particular became apparent in the 'Grazia Colonia' example: Christian Schlögl was passionate about his cause. Achieving a good result was, however, more of a pleasant side effect. In his dealings with 'his' students—among whom we students from Cologne can now count ourselves despite the border—he always had one thing in mind: He offered each and every person involved in the project help to help themselves, was always ready to support wherever help was needed or a perspective needed to be steered in a different direction. His unconventional manner and the way we met on equal terms turned us into a team with a common goal in mind (Wolff, personal communication, May 1, 2025).

5.2. Web of Science and Scopus as Tools for Bibliometric Analyses

We selected this research topic because the primary publication is highly cited (article no. 1 from Table 7 is Schlögl's top-cited paper), and it is part of the author cluster A3 (however, it is only connected with Gorraiz). It is the theme of the topic cluster T4. The co-author of this research project is Gorraiz from the University of Vienna, Austria; the results were published in 2008.

The leading research question of this project is: How suitable are the citation databases WoS and Scopus for scientometric analyses (Gorraiz & Schloegl, 2008)? As the researchers' publication and citation behavior depend on the scientific discipline, Gorraiz and Schlögl considered only one research field: pharmacology and pharmacy. In most cases, the impact factors of the journals were higher in Scopus than in WoS. However, since Scopus covers more journals than WoS, "a journal in Scopus has a higher chance of getting cited in general" (Gorraiz & Schloegl, 2008, p. 717). The authors could demonstrate that both WoS and Scopus have problems with the accuracy of data, e.g., concerning the correct determination of the articles' document types. Furthermore, "the assignment of subject categories to journals is not always transparent and differs between Scopus and WOS" (Gorraiz & Schloegl, 2008, p. 723).

Juan Gorraiz remembers the projects with Schlögl:

I first met Christian Schlögl in Eisenstadt, where he had just begun laying the foundations for the *Fachhochschule* (University of Applied Sciences). At that time, I was teaching a few courses there, and it was during those early encounters that I had the opportunity to spark in him not only an interest but what I believe became a genuine passion for bibliometrics. That shared curiosity quickly turned into a rich and long-standing collaboration. At the time, I was still heading the Department of Information and User Services at the Central Physics Library in Vienna, so my time for research and publishing was limited. Christian more than made up for this. He developed into an ideal co-author—diligent, original, and full of energy—quickly finding his own style and applying bibliometric methods to topics close to his heart.

Our first joint project was in fact a small milestone: We were among the first to use Scopus as a data source in bibliometric analyses. We presented our results at the 11th International Society for Scientometrics and Informetrics (ISSI) Conference in Madrid in June 2007, where our work was met with strong interest and warmly received by the scientometrics community. It felt like a shared entrance onto a bigger stage. What followed were many discussions—often long and lively—whenever we met at conferences or university events. Some of the most memorable ones happened in the context of the University Course for Library and Information Studies, where we both contributed and exchanged views passionately. It was during these meetings that I came to appreciate even more his critical mind, his dry humor, and his generous collegiality.

Among the areas we worked on, usage metrics became our most focused domain—recognizing their importance alongside traditional citation metrics. Christian approached this topic with characteristic pragmatism and intellectual rigor. Later, with my new responsibilities as head of the newly established Bibliometrics Department at the Vienna University Library, our professional paths diverged. But we never lost touch (Gorraiz, personal communication, April 23, 2025).

5.3. Information Usage Behavior of Researchers: Citations, Downloads, and Reference Services

This theme is mirrored in articles no. 4 and 9, topic cluster T1, and author cluster A3. In this project, Schlögl worked together with Gorraiz, Gumpenberger (University of Vienna), and Jack (Mendeley). The publication from 2015 resulted from a cooperation between Schlögl, Jack, Lindstaedt, and Kraker (the latter from Know Center in Graz at the time of publication). The cooperation between Schlögl and Gorraiz on this project lasted between 2010 and 2015 and is strongly connected with the project from paragraph 5.2.; the other authors joined the publication activities on this research between 2013 and 2015.

With the offers of full-text articles and user-generated reference services, new retrieval and document delivery options have emerged as well as new means of scientometric research. Now, there is more data measuring research impact, including:

- citations (as before),
- downloads of article full texts (new),
- metadata on reference services (as a new proxy of readership).

Data sources were Scopus and the Journal Citation Reports of WoS for the citations, Elsevier's Science Direct for downloads, and Mendeley for readerships (Schlögl et al., 2014a). Schlögl and Gorraiz (2010, p. 569) introduced new scientometric indicators with the usage impact factor and the usage half-life. These new indicators were constructed in analogy to the journal impact factor and the journal half-life. Articles have the highest download rates immediately after being available online, while citations come years later. Therefore, downloads have another "meaning" than citations, "which denote only a small aspect of science communication" (Schlögl & Gorraiz, 2010, p. 578), and "usage metrics can be regarded as an interesting complement to citation metrics" (Schlögl & Gorraiz, 2011, p. 161). The article's readership belongs to different sectors, including academics, governmental or industrial readers, and the general public: "As a logical consequence, articles are often downloaded many times but remain uncited due to the fact that they are used for other purposes (pure information, learning, teaching, etc.) apart from the publish or perish 'game'" (Gorraiz et al., 2014, p. 1093).

"Readership" becomes defined with an altmetric measure as the number of mentions of an article in Mendeley user libraries (Schlögl et al., 2014b, p. 1119). A study on two paradigmatic information systems journals shows medium to high rank correlations between downloads and citations, relatively high correlations between downloads and readerships, and much lower correlations between readership and citations (Schlögl et al., 2014b). Schlögl et al. (2014b, p. 1126) found an interesting hint on the relationship between citations and downloads: "Later on citations contribute to a re-increase in the downloads to some extent."

In analogy to co-citations, a further project introduced co-readership, again based on Mendeley user libraries (Kraker et al., 2015). In addition, other data sets and sometimes other data sources (such as SCIMago) were used and the focus was on relational, not evaluative scientometrics. Co-readership means that two articles are found in the library of an individual Mendeley user. The outcomes suggest that "the results of a co-readership analysis may be much more up-to-date than co-citation analysis" (Kraker et al., 2015, p. 179). Although co-readership probably offers a weaker indication of subject similarity than co-citation, it can still be expected to serve as a useful indication of subject similarity and as a sound basis for the visualization of knowledge areas.

Peter Kraker was researcher at the Know Center in Graz; now he is founder and chairman of Open Knowledge Maps in Vienna, Austria. Kraker notes,

> Christian Schlögl sought to move beyond the well-trodden paths of bibliometrics research, which included his interest in exploring new measures. He had a keen eye for identifying promising research directions and actively forged new collaborations to pursue them. One such example is this project, which ultimately brought together four institutions from Austria and the UK and was among the first to compare citations, downloads, and readership metrics. The insights from this work helped shape the discourse around alternative metrics at a time when such approaches were still emerging.
>
> Christian had a profound influence on me throughout my studies, academic career, and later professional life. His work inspired my own research into scholarly communication on the web and its use in visualizing knowledge domains. He supported my interdisciplinary approach and embraced rather than eschewed the complexities it brought. He was an advisor, mentor, and a consistently positive force in my life—and in the lives of many others. He will be sorely missed (Kraker, personal communication, May 2, 2025).

5.4. Evaluation of Research Institutions

This research topic results from the author clusters A5 and (partially) A4. Co-authors were Dorsch (University of Düsseldorf, later Leibniz Information Center for Economics in Kiel, Germany, now GEMASS [Sorbonne University and CNRS] in Paris, France), Reichmann, and Stock, as well as Gorraiz (in the 2003 publication) and Boric from the University of Graz. The main project's publications started in 2022 and lasted until Schlögl's passing; however, Schlögl also researched this topic in earlier stages of his career.

Schlögl et al. (2003) analyzed two paradigmatic research institutions to show the problems with quantitative research evaluation. They found two groups of issues, namely data and indicators, as "different data and different indicators can result in quite different outcomes" (Schlögl et al., 2003, p. 298). For instance, if one cannot find concrete data on an institution's staff, the number of papers per researcher cannot be calculated. The researcher will get different results for publications and citations when searching for information on various services.

Some of Schlögl's studies include the analysis of his home institution in comparison with other research institutions, e.g., the research subjects of information science institutions in Graz and Düsseldorf via a topic analysis (Dorsch et al., 2017) or a long-time study on the research performance of the Graz institution (Reichmann & Schlögl, 2022). Bibliographic databases cover only 13% (WoS) or 23% (Scopus) of all publications. There are different points of view when the analyzed time intervals change. The complete history (33 years) shows the entire research performance. In contrast, shorter intervals (e.g., from 11 years to 3 years or annually) depend more on publication and citation data fluctuations.

It makes a difference in the determination of the performance of an institution if scientometricians work with simple numbers of employees or with full-time equivalents to calculate the institution's labor productivity and impact (Stock et al., 2023a). However, it is not always easy to get such data on employment. How does one search for an institution, for instance, on WoS (Stock et al., 2025)? One can search for the institution's name or for the names of its research staff in a given time interval. This project provided evidence: A search for affiliated authors outperforms a search for institutions in WoS.

Isabelle Dorsch was part of the project team on the scientometrics at the meso level, and adds:

I have pleasant memories of our discussions and planning meetings with colleagues in Graz regarding our collaboration, for example, the scientometric analysis on the influence of the chosen methodology for research evaluation at the institute level or general research topics. I particularly remember his (Christian Schlögl's) friendly, team-focused, and pragmatic working style when I think of our collaborations. Through his work in this area, he actively contributed to analyzing the methodological variations (and biases) for scientometrics studies, apart from his other foci of study (Dorsch, personal communication, March 31, 2025).

Another project co-authored with Boric and Reichmann covered publications of more than 200 business administration professors in Austria. One can get performance data from WoS or Scopus (Schlögl et al., 2024) or from personal publication lists (Dorsch, 2017; Dorsch et al., 2018) of all researchers to be analyzed (Reichmann et al., 2024). Especially for non-English publications, personal publication lists are very useful for bibliometric approaches. "When comparing the individual researchers' ranks based on their personal publication lists (…) with those based on WoS (…) for the basic variant (…), dramatic ranking differences become apparent" (Reichmann et al., 2024, p. 6). In another project module, Boric et al. (2024) analyzed not the individual researchers but the entire group of Austrian university institutions of business administration. In the light of WoS and Scopus, the ranks of the particular institutions are stable in the rankings based on publications per researcher. In the analysis of citation data, "The ranks of the Austrian business schools remain relatively unchanged when moving from WoS to Scopus. There are larger changes when switching from full to adjusted counting, and even larger (…) when switching from citations per researcher to citations per publication" (Boric et al., 2024, p. 18). The project team worked with several performance indicators. The choice of the indicators heavily influences the rankings. As Boric et al. (2024) conclude, "Therefore, the main contribution of our approach to existing international university rankings is that it demonstrates that no composite indicators should be used. Instead, separate and more comprehensive rankings for each considered indicator should be performed" (Boric et al., 2024, p. 23).

Given the problems with scientometric indicators at the meso level, it is clear that rankings of institutions based upon composite indicators, e.g., according to their alleged "importance," make little or no sense. Such problematic rankings are, among others, the Academic Ranking of World Universities (ARWU), the Quacquarelli Symonds World University Rankings (QS), the Times Higher Education World University Ranking (THE), the Leiden Ranking, and the U.S. News Best Global Universities (Schlögl et al., 2025).

Sandra Boric describes the research project:

In our area of research, one might be inclined to stick to describing numerical values and statistics. Prof. Schlögl, however, always paid attention to including a holistic view in our texts and reminding the reader of the bigger picture we are looking at. When it comes to evaluating research institutions, much of his focus was on the comparison of publication- and citation-based rankings. He made an effort to include concrete examples for value differences in institutions, and he always wanted to (if possible) use Austrian or German institutions as examples to raise their research visibility and cover this niche topic of research.

His methodologies heavily focused on reproducibility and describing the data collection process. We also used several levels to classify data, e.g., on the researcher level, subject area level, and institutional level (which again shows the desire to remind the reader of the bigger picture we are looking at). His mantra was that the data sources should always be reliable and (if possible) easily accessible, while the data collection process and used indicators can be as diverse and detailed as the research project requires.

In our research institution evaluations, Scopus generally fared better as a data source than WoS. We knew that the larger institutions would rank higher in absolute terms and visibility, so we adjusted counts, developed a model of ranking variants for business schools, and looked at the ranking stability across institutions. We derived implications for research practice for rankings based on the number of publications, the number of pages per researcher, the share of highly active researchers (e.g., with ten or more publications), the institution's top five most publishing researchers' publication counts, the average citation count (per researcher and publication), the institution's top five most cited researchers' citation counts, and the H-indices (Boric, personal communication, March 21, 2025).

5.5. Estimating the Demand of a Country's Information Professionals

Author cluster A2 gave us hints to this project. Schlögl's co-authors were Hayes (at the time of the project, Professor Emeritus of the UCLA Graduate School of Library and Information Science in Los Angeles, CA, USA) and Karlics (Graz and Bolzano, Italy). This research was published between 2009 and 2013.

The project's leading research question is: What is the size of the information industry in a country? Austria serves as an example of an economy (Hayes et al., 2009). A primary objective is to estimate the need for information management activities. Such activities are essential for university libraries, which serve as models for the complete knowledge industries of a country. Using the Library Planning Model from Hayes, the staffing needs of an average Austrian university library in media processing and user support are determined. These needs represent the starting point for estimating the knowledge industries' demand for information management activities. The number of librarians in university libraries can be used to roughly estimate the demand for information specialists in knowledge-intensive industries (Hayes et al., 2011). Data sources are the official statistics for the national input-output system, including data on, e.g., products of publishing houses and data on the number of employees in libraries and knowledge industries (in full-time equivalents). For the libraries, the actual number of all librarians per library user could be determined—Hayes et al. (2011, p. 328) calculated a value of 0.007733. Following their model, the same value is valid for the number of information specialists per knowledge worker in the economy. In Austria, there are 337,755 knowledge workers, 114,117 students needing knowledge, and 3,494 full-time equivalents of information specialists (Hayes et al., 2013). According to the Austrian official statistics, there were 3,477 information specialists in the Austrian knowledge industry in 2005. Therefore, the calculated value applying the Library Planning Model seems to be a good approximation of the official value.

Karin Karlics describes this research project.

The idea behind the project with the Library Planning Model was to strengthen the role of librarians, in particular, that of information specialists. The goal was to increase their visibility and enhance the value of their professional activities. Christian Schlögl, who was involved in training librarians in various roles throughout his career, was open to this innovative project.

Estimating the demand for a country's information professionals based on data from academic libraries was considered unusual and controversial. However, it was precisely this exploratory research approach that gave the project its unique appeal. The result of approximating the official value surprised even the authors.

The bibliometrician Christian Schlögl demonstrated remarkable openness to new, albeit unconventional, approaches through this project. However, further studies would have been necessary to substantiate the results. In particular, analyses at the micro level or international comparative studies could have provided valuable additional insights (Karlics, personal communication, March 23, 2025).

5.6. Information Literacy

Author cluster A1 led us on the trail of this topic. The primary co-author is Dreisiebner from Graz, who later worked in Hildesheim, Germany, and Villach, Austria; however, Schlögl also worked with colleagues from several countries. The publication activities took place between 2015 and 2019.

The main attempts of this research line were to determine the level of students' information literacy and options for improving them. Schlögl adopted a standardized test for measuring information literacy, which Beutelspacher (2014) created in Düsseldorf, Germany. In his course on information science for business administration students, Schlögl taught fundamentals of information literacy, and the information literacy questionnaire was used instead of the usual written exam (Beutelspacher et al., 2015). In general, the students answered most questions successfully. After finishing this questionnaire, the students received a second (now anonymous) questionnaire on their appreciation of the information literacy questionnaire. This evaluation also had positive outcomes. The next project step was a semi-structured interview to collect data on the perception of information literacy among academic staff at the universities of Graz, Tallinn, and Zagreb (Kirinić et al., 2015). A further step led Schlögl to South Korea. In cooperation with Seo and Rust, Schlögl compared the information literacy of students from the

University of Graz and Chungbuk National University (Rust et al., 2017), applying a modified version of Beutelspacher's test. Another project team with Maurer and Dreisiebner used the Beutelspacher test again to compare the information literacy of beginner students from different study programs (Maurer et al., 2017). The difference in information literacy across the various disciplines was not very high, but the overall competencies were weak or mediocre. As was reflected, "Another interesting result of our study is that students' self-assessed level of information literacy is much higher than it actually is" (Maurer et al., 2017, p. 317).

Are there textbooks for teaching information literacy? Dreisiebner and Schlögl (2019) analyzed discipline-specific teaching materials via a structuring content analysis. The most significant difference between the disciplines is determining the nature and extent of the needed information, especially in identifying potential sources of information, e.g., discipline-specific information services. The authors also discussed the implementation of Massive Open Online Courses (MOOCs) into academic teaching, noting, "Besides regular classes, MOOCs might be a possible solution, which have the advantage to be completely web-based and thus allow the direct integration of practical exercises and examples" (Dreisiebner & Schlögl, 2019, p. 410).

Dreisiebner here remembers:

Christian initially brought my attention to the research field of information literacy during our first consultation session for my Ph. D. studies. He immediately recognized the fitting to my educational background and interests. Indeed, this was a perfect fit, and until today, information literacy stays one of my core research fields. I am very grateful for his guidance in these early days of my academic career. Christian held some popular courses on information literacy for business and Ph. D. students, which were perceived by students as very helpful. The practical experience from his teaching offers, as well as his involvement in the training of librarians in Austria and the information literacy groups among librarians in general inspired his research, where he was especially concerned with the facilitation and evaluation of information literacy in classroom settings. In our work, we also explored new approaches for information literacy facilitation, specifically through MOOCs. He was especially familiar with quantitative approaches to measuring information literacy. Some of his key findings concerned the gap between students' self-perception of their information literacy skills and actual evaluation results, subject-specific differences in information literacy skills, as well as approaches for subject-specific adaptations for information literacy achievement tests (Dreisiebner, personal communication, March 26, 2025).

5.7. Information and Knowledge Management

Article no. 7 from Table 7 and the mention of knowledge management in Table 11 gave hints to this topic, which Schlögl processed alone. Besides this paper from 2005, Schlögl published a book (Schlögl, 2001) and some articles on this topic.

The book (Schlögl, 2001) is an "inventory" of information management and includes a literature overview and a bibliographic study. Schlögl justified the twofold approach: "Since a conventional literature analysis always carries the risk of a certain subjectivity and therefore cannot satisfy fundamental scientific criteria such as intersubjective confirmation, a more objective method was also used: a scientometric study" (Schlögl, 2001, p. 1). Additionally, he analyzed the main dimensions of information management, utilizing an author co-citation study (Schlögl, 2003). This bibliographic study identifies three main branches of information management: technology-oriented information management, content-oriented information management, and knowledge management (Schlögl, 2005).

Since the early days of his research, Schlögl followed an empirical approach to information science, which is always supported by literature studies. By this way, in this article, we adopted Schlögl's approach.

6. DISCUSSION AND CONCLUSION

6.1. The Œuvre of Christian Schlögl in a Nutshell

Our bibliometric analysis reveals that Christian Schlögl was an extremely active and respected researcher. Over the course of his academic career, he authored 177 publications, which have been cited almost 700 times to date (based on Scopus). It should be noted that only the 34 English-language journal articles recorded by this information service have a chance of being counted in the number of citations. The impact of all other publications, including numerous German-language journal articles as well as contributions to proceedings and edited books, cannot be quantified.

It seems positive that Christian Schlögl published relatively consistently over the years, thus showing no research fatigue even in his advanced career. Therefore, it can be assumed that numerous more publications would have appeared had his sudden death not abruptly ended all research activities.

The numerous collaborations and co-authorships are a strong indication that Christian Schlögl was an extremely cooperative and popular researcher. Collaborations were also a welcome opportunity for him to pass on his extensive knowledge to younger colleagues, as emphasized in several comments by co-authors.

Christian Schlögl's research content was quite broad, as confirmed by the topic analysis. This is also reflected in the variety of publication venues he chose. The focus of his journal-based publications was not only information science journals, but also library and computer science journals. If one wanted to reduce the main research topics from the seven previously presented to two, this would be:

a) Scientometrics of journals, including the information behavior of their readers, and
b) Scientometrics of research institutions (which are on the scientometric meso level).

As part of his research activities, Christian Schlögl used a variety of empirical methods and data sources, such as publication analysis, citation analysis, co-citation analysis, questionnaires, semi-structured interviews, usage statistics, official statistics, tests (e.g., for information literacy), and content analysis, as well as altmetrics (Mendeley).

Isabelle Dorsch, one of his many co-authors, characterizes Schlögl's research topics and his research style:

> Based on my personal perception, our joint projects, and personal contact over the last decade, I would characterize Christian Schlögl's research style as an empirical approach focusing on quantitative analysis within scientometrics. For his studies focusing on the micro or meso level, he utilized publication metadata that was self-obtained or provided by different information services. He contributed to the international and national body of information science research (Dorsch, personal communication, March 31, 2025).

Juan Gorraiz also states,

> Christian, to me, embodied what a true colleague and co-author should be: sharp, curious, constructive, and always open to new ideas. He challenged me, supported me, and inspired me. Even now, I often find myself thinking of him when I reflect on a project or prepare a presentation. Just like the "invisible colleges" that shape the intellectual landscape of science, I believe in the presence of "invisible colleagues"—those who, though no longer with us, remain part of our professional and personal journey. Christian is, and will always be, one of those for me (Gorraiz, personal communication, April 23, 2025).

6.2. Scientometrics at the Micro Level: Strengths, Weaknesses, Opportunities, and Threats (SWOT)

Our final section is a SWOT analysis of scientometrics (Rousseau & Rousseau, 2021) in the service of the description of the oeuvre of an individual researcher. Fig. 5 exhibits the scientometric indicators which we applied. Concerning basic data, one must confirm that the lists of publications and citations are truebounded. As we have to work with information services, we should calculate the visibility of an ego's oeuvre on different information services. The bibliometric picture of the research production is mirrored in the ego's publications, their document types, and their time series. Given a truebounded publication list, one can easily determine the co-authors, the cluster of co-authors, and the preferred publication sources.

Additionally, we tried to find the ego's citation identity, that is, the authors of the literature cited by the ego. Regarding the research impact, we searched for literature citing our ego and created a time series. The citation image makers are the authors who cited documents of the ego. Analyzing the countries of activity of the citation image makers, we obtained a picture of the diffusion of the ego's research results all over the world. "Power publications" are the most-cited documents written by the ego. The research topics can be viewed from two perspectives: the ego's publications and the citing literature.

Besides the truebounded Uni Graz Online information service, we applied WoS and Scopus, although underbounded, for data acquisition. The different search functionalities of both information services drove the decision for each.

The *strength* of a bibliometric approach on the micro level is the mostly easy handling of the basic data. However, we needed an appropriate database. For our case study, it was the institutional research repository (Uni Graz Online) and not the commonly used information services such as Scopus or WoS, as they all were incomplete. A further strength of this approach is the offer of citation data by the information services; however, we must assume that this information again is incomplete. But when it comes to citation data, there is nothing better. In line with Rousseau and Rousseau (2021, p. 1436), we can confirm that bibliometric studies are objective and reproducible as

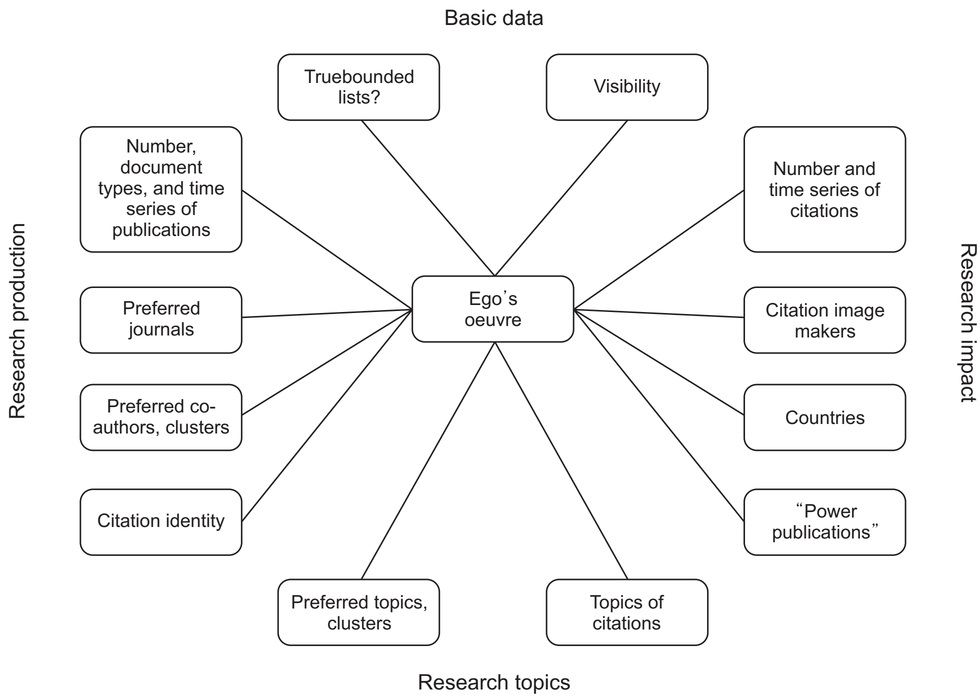

Fig. 5. Scientometric indicators for the description of a researcher's oeuvre.

well as cost- and time-effective. Additionally, they allow for tracing the history of ideas, in our case, the history of the ideas of an individual researcher.

There are many *weaknesses* in bibliometrics (Stock et al., 2023b). The main problems are the incompleteness of information services and the absence of consistent indexing. While Scopus could identify Schlögl through a unique researcher number, all other used databases failed. They failed to correctly handle homonymy (showing and merging other researchers with the same or similar name) and synonymy (indexing name variants of the same author as different researchers, here, e.g., Schlögl, Schloegl, and Schlogl). In our case, Google Scholar was completely unusable. If researchers want to be correctly indexed in Google Scholar, they must necessarily correct their publication lists, that is, adding missing publications, merging synonymous author names, and deleting publications by other researchers with similar names.

It takes some time before a publication can be cited. Therefore, recent works initially have little chance of being highly cited, and so citation databases and, subsequently, bibliometrics are biased towards older publications. We did not find important recent articles by Schlögl via citation analysis but through author clusters or topic clusters. A further weakness is the missing functionality of information services such as WoS or Scopus. We were sorely missing a rank command for all the names of the cited authors when looking for Schlögl's citation identity. However, Semantic Scholar offers ranked lists for a researcher's citing authors, referenced authors, and co-authors. But we could not use these options as we failed to correctly identify Schlögl and his publications due to errors concerning synonymy and homonymy.

A great *opportunity* of bibliometrics, in contrast to other methods (e.g., interviews or literature studies), is the provision of a "big picture" (Rousseau & Rousseau, 2021, p. 1436) of the complete oeuvre of a researcher. This "big picture" will be complete if we use a complete database with all publications by an author (as in our case, Uni Graz Online); it will be sketchy and mutate into a *threat* if one applies information services such as WoS or Scopus. In our case study, these problems concern Schlögl's citation identity, citations of his articles, citation image makers including their countries and topics, and Schlögl's power publications.

We have to mention a practical *limitation* of the applied method. Concerning our illustrative example, our approach works well, especially as there is a reliable and (more or less) clean and complete digital database, which could be validated through collegial insight. However, our approach is not easily generalizable if there is not such a rich data source. Problems in data collection may arise through homonyms and synonyms of author names, researchers publishing in under-indexed languages, or missing institutional repositories or personal publication lists. Additionally, one has to find the best-fitting experts for interviews who are willing to share their knowledge. All the experts we interviewed provided us with their written texts, so no revisions or edits were necessary. We were very happy and grateful here. Our methods work only well when the researcher is well known and has a true-bounded publication list and when he had co-authors in order to provide contextual information.

As professional information services as WoS or Scopus are probably underbounded for all individual researchers, it is very helpful for all scientometric studies to have available truebounded online publication lists for the analyzed researchers, be it an institutional repository (Rothfritz et al., 2025) or personal publication lists (Dorsch et al., 2018).

It would be an interesting research topic to compare researchers' "true" scientometric representations (based upon their time-dependent complete sets of publications and citations for a specific period of time) and their representations in single information services (e.g., WoS and Scopus) or in combinations of these information services. A reviewer of this article called such a representation a "scientometric personhood," being "an algorithmic-bibliometric view of the academic self, shaped as much by indexing systems and visibility infrastructures as by the researcher's own agency" (anonymous reviewer, personal communication, August 15, 2025). Such scientometric personhoods exhibit "how scholarly lives are represented—and sometimes reduced—by quantitative proxies" (anonymous reviewer, personal communication, August 15, 2025).

Scientometric indicators are appropriate for an overview of a researcher's life work and provide us with heuristics for the researcher's research production, research impact, and research topics. However, quantitative data on publications, citations, and topics are only the beginning of the description, analysis, and interpretation of the contents of an ego's publications and the citations those publications attracted. At this point, we have to talk with colleagues and co-authors, and we have to read and understand all the literature that is recognized as important.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGEMENTS

We want to thank our interviewed experts and former co-authors of Christian Schlögl for their valuable statements. Our thanks go in particular to Sandra Boric, Isabelle Dorsch, Stefan Dreisiebner, Juan Gorraiz, Karin Karlics, Peter Kraker, and Kerstin Wolff. Our thanks also go to the reviewers, whose valuable comments were all incorporated into our text.

REFERENCES

Ayala-Gascón, M., Aleixandre-Benavent, R., & Gandía-Balaguer, A. (2012). [Indicators of scientific activity of individual researchers: A bibliometric profile of Eduardo Primo Yúfera, former president of CSIC9]. *Revista Española de Documentación Científica*, 35(2), 209-237. Spanish. https://doi.org/10.3989/redc.2012.2.887

Beutelspacher, L. (2014, October 20-23). Assessing information literacy: Creating generic indicators and target group-specific questionnaires. In S. Kurbanoğlu, S. Špiranec, E. Grassian, D. Mizrachi, & R. Catts (Eds.), *Information Literacy: Lifelong Learning and Digital Citizenship in the 21st Century* (pp. 521-530). Springer Cham. https://doi.org/10.1007/978-3-319-14136-7_55

Beutelspacher, L., Henkel, M., & Schlögl, C. (2015). *Evaluating an information literacy assessment instrument: The case of a bachelor course in business administration*. Paper presented at 14th International Symposium on Information Science (ISI 2015), Zadar, Croatia. https://doi.org/10.5281/zenodo.17980

Boric, S., Reichmann, G., & Schlögl, C. (2024). Possibilities for ranking business schools and considerations concerning the stability of such rankings. *PLoS One*, 19(2), e0295334. https://doi.org/10.1371/journal.pone.0295334

Bornmann, L., Ganser, C., & Tekles, A. (2022). Simulation of the H-index use at university departments within the bibliometrics-based heuristics framework: Can the indicator be used to compare individual researchers? *Journal of Informetrics*, 16(1), 101237. https://doi.org/10.1016/j.joi.2021.101237

Cronin, B., & Shaw, D. (2002). Identity-creators and image-makers: Using citation analysis and thick description to put authors in their place. *Scientometrics*, 54(1), 31-49. https://doi.org/10.1023/A:1015628320056

Daud, A., Ahmed, W., Amjad, T., Nasir, J. A., Aljohani, N. R., Abbasi, R. A., & Ahmad, I. (2017). Who will cite you back? Reciprocal link prediction in citation networks. *Library Hi Tech*, 35(4), 509-520. https://doi.org/10.1108/lht-02-2017-

0044

Dorsch, I. (2017). Relative visibility of authors' publications in different information services. *Scientometrics*, 112(2), 917-925. https://doi.org/10.1007/s11192-017-2416-9

Dorsch, I., Askeridis, J. M., & Stock, W. G. (2018). Truebounded, overbounded, or underbounded? Scientists' personal publication lists versus lists generated through bibliographic information services. *Publications*, 6(1), 7. https://doi.org/10.3390/publications6010007

Dorsch, I., Schlögl, C., Stock, W. G., & Rauch, W. (2017). [Research topics of the information science institutions in Düsseldorf and Graz]. *Information - Wissenschaft & Praxis*, 68(5-6), 320-328. German. https://doi.org/10.1515/iwp-2017-0060

Dreisiebner, S., & Schlögl, C. (2019). Assessing disciplinary differences in information literacy teaching materials. *Aslib Journal of Information Management*, 71(3), 392-414. https://doi.org/10.1108/ajim-07-2018-0183

Gorraiz, J., Gumpenberger, C., & Schlögl, C. (2014). Usage versus citation behaviours in four subject areas. *Scientometrics*, 101(2), 1077-1095. https://doi.org/10.1007/s11192-014-1271-1

Gorraiz, J., & Schloegl, C. (2008). A bibliometric analysis of pharmacology and pharmacy journals: Scopus versus Web of Science. *Journal of Information Science*, 34(5), 715-725. https://doi.org/10.1177/0165551507086991

Grazia Colonia. (2002). *[Information science journals in scientometric analysis]*. German. https://epb.bibl.th-koeln.de/frontdoor/deliver/index/docId/30/file/grazia.pdf

Haustein, S. (2012). *Multidimensional journal evaluation: Analyzing scientific periodicals beyond the impact factor.* De Gruyter Saur. https://doi.org/10.1515/9783110255553

Hayes, R. M., Karlics, K., & Schlögl, C. (2009). *[An analysis of the Austrian information sector].* Paper presented at Research Resources in the Field of Information Science, Konstanz, Germany. German. https://doi.org/10.5281/zenodo.5506491

Hayes, R. M., Karlics, K., & Schlögl, C. (2011). [The demand for information specialists in knowledge-intensive industries of the Austrian economy]. In J. Griesbaum, T. Mandl, & C. Womser-Hacker (Eds.), *Information und Wissen: Global, sozial und frei? Proceedings des 12. Internationalen Symposiums für Informationswissenschaft* (pp. 309-320). Hülsbusch. German. https://zenodo.org/records/4134607

Hayes, R. M., Karlics, K., & Schlogl, C. (2013). University libraries as a model for the determination of the need for information specialists in knowledge industries? An exploratory analysis of the information sector in Austria. *Information Processing & Management*, 49(5), 1008-1018. https://doi.org/10.1016/j.ipm.2013.04.001

Jacso, P. (2018). The scientometric portrait of Eugene Garfield through the free ResearcherID service from the Web of Science Core Collection of 67 million master records and 1.3 billion references. *Scientometrics*, 114(2), 545-555. https://doi.org/10.1007/s11192-017-2624-3

Johnson, B., & Oppenheim, C. (2007). How socially connected are citers to those that they cite? *Journal of Documentation*, 63(5), 609-637. https://doi.org/10.1108/00220410710827727

Juchem, K., Schögl, C., & Stock, W. G. (2006). [Dimensions of the scientometrics of journals]. *Information - Wissenschaft & Praxis*, 57(1), 31-37. German.

Kirinić, V., Schlögl, C., & Virkus, S. (2015, October 19-22). Perception of information literacy among faculty at the University of Graz, Tallinn University, and University of Zagreb. In S. Kurbanoglu, J. Boustany, S. Špiranec, E. Grassian, D. Mizrachi, & L. Roy (Eds.), *Information Literacy: Moving Toward Sustainability* (pp. 467-477). Springer Cham. https://doi.org/10.1007/978-3-319-28197-1_47

Kraker, P., Schlögl, C., Jack, K., & Lindstaedt, S. (2015). Visualization of co-readership patterns from an online reference management system. *Journal of Informetrics*, 9(1), 169-182. https://doi.org/10.1016/j.joi.2014.12.003

Maurer, A., Schlögl, C., & Dreisiebner, S. (2017). Comparing information literacy of student beginners among different branches of study. *Libellarium*, 9(2), 309-319. https://doi.org/10.15291/libellarium.v9i2.280

Meuser, M., & Nagel, U. (2009). The expert interview and changes in knowledge production. In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing Experts* (pp. 17-42). Palgrave Macmillan UK. https://doi.org/10.1057/9780230244276_2

Milard, B. (2014). The social circles behind scientific references: Relationships between citing and cited authors in chemistry publications. *Journal of the Association for Information Science and Technology*, 65(12), 2459-2468. https://doi.org/10.1002/asi.23149

Peters, I., Schmitz, J., Weller, K., Haustein, S., & Trkulja, V. (2020). [Wolfgang G. Stock – A bibliometric review of a career in information science]. In I. Dorsch, K. J. Fietkiewicz, A. Ilhan, C. Meschede, & T. Siebenlist (Eds.), *[Facets of Wolf Stock and their significance for information science]* (pp. 35-57). Hülsbusch. German.

Peters, I., & Stock, W. G. (2010). "Power tags" in information retrieval. *Library Hi Tech*, 28(1), 81-93. https://doi.org/10.1108/07378831011026706

Pranckutė, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications*, 9(1), 12. https://doi.org/10.3390/publications9010012

Rawat, S., Chaturvedi, K. K., Ramasubramanian, V., Farooqi, M. S., Sharma, A., & Pal, S. (2024). Development of a novel metric for productivity assessment of researchers using bibliometric data. *Annals of Library and Information Studies* 71(3), 319-330. https://doi.org/10.56042/alis.v71i3.9010

Reichmann, G. (2024). Christian Schlögl. *Information – Wissenschaft und Praxis*, 75(4), 202-204. German. https://doi.org/10.1515/iwp-2024-2021

Reichmann, G., & Schlögl, C. (2022). On the possibilities of presenting the research performance of an institute over a long period of time: The case of the Institute of Information Science at the University of Graz in Austria. *Scientometrics*, 127(6), 3193-3223. https://doi.org/10.1007/s11192-022-04377-8

Reichmann, G., Schlögl, C., Boric, S., & Nimmerfall, J. (2024). The usefulness of personal publication lists in research evaluation. *The Journal of Academic Librarianship*, 50(4), 102881. https://doi.org/10.1016/j.acalib.2024.102881

Rothfritz, L., Matthias, L., Pampel, H., & Wrzesinski, M. (2025). Current challenges and future directions for institutional repositories: A systematic literature review. An Annual Review of Information Science and Technology (ARIST) paper. *Journal of the Association for Information Science and Technology*, 1-22. https://doi.org/10.1002/asi.70016

Rousseau, S., & Rousseau, R. (2021). Bibliometric techniques and their use in business and economics research. *Journal of Economic Surveys*, 35(5), 1428-1451. https://doi.org/10.1111/joes.12415

Rust, K., Schlögl, C., & Seo, D. (2017, March 13-15). Comparing information literacy of students from University of Graz (Austria) and Chungbuk National University (Republic of Korea). In M. Gäde, V. Trkulja, & V. Petras (Eds.), *Everything Changes, Everything Stays the Same? Understanding Information Spaces. Proceedings of the 15th International Symposium of Information Science (ISI 2017)* (pp. 24-36). Verlag Werner Hülsbusch. https://doi.org/10.18452/1454

Sandström, U., & Sandström, E. (2009, July 14-17). Meeting the micro-level challenges : Bibliometrics at the individual level. In B. Larsen, & J. Leta (Eds.), *Proceedings of ISSI 2009 - 12th International Conference of The International Society for Scientometrics and Informetrics, Vol 2* (pp. 846-856). BIREME/PAHO/WHO and Federal University of Rio de Janeiro.

Schlögl, C. (2001). [*Inventory of information management: A scientometric, qualitative and empirical analysis*]. Gabler. German. https://doi.org/10.1007/978-3-663-08276-7

Schlögl, C. (2003). [Knowledge map of information management]. *Wirtschaftsinformatik*, 45(1), 7-16. German. https://doi.org/10.1007/BF03250879

Schlögl, C. (2005). Information and knowledge management: Dimensions and approaches. *Information Research*, 10(4), no. 235. http://InformationR.net/ir/10-4/paper235.html

Schlögl, C., Boric, S., & Reichmann, G. (2024). Publication and citation patterns of Austrian researchers in operations research and other sub-disciplines of business administration as indexed in Web of Science and Scopus. *Central European Journal of Operations Research*, 32(3), 711-736. https://doi.org/10.1007/s10100-023-00877-x

Schlögl, C., & Gorraiz, J. (2010). Comparison of citation and usage indicators: The case of oncology journals. *Scientometrics*, 82(3), 567-580. https://doi.org/10.1007/s11192-010-0172-1

Schlögl, C., & Gorraiz, J. (2014). A comparison of citations, downloads and readership data for an information systems journal. *Research Trends*, 1(37), 5.

Schlögl, C., Gorraiz, J., Bart, C., & Bargmann, M. (2003). Evaluating two Austrian university departments: Lessons learned. *Scientometrics*, 56(3), 289-299. https://doi.org/10.1023/A:1022318618200

Schlögl, C., Gorraiz, J., Gumpenberger, C., Jack, K., & Kraker, P. (2014a). A comparison of citations, downloads and readership data for an information systems journal. *Research Trends*, 1(37), art. 5. https://www.researchtrends.com/researchtrends/vol1/iss37/5

Schlögl, C., Gorraiz, J., Gumpenberger, C., Jack, K., & Kraker, P. (2014b). Comparison of downloads, citations and readership data for two information systems journals. *Scientometrics*, 101(2), 1113-1128. https://doi.org/10.1007/s11192-014-1365-9

Schlögl, C., & Petschnig, W. (2005). Library and information science journals: An editor survey. *Library Collections, Acquisitions, & Technical Services*, 29(1), 4-32. https://doi.org/10.1080/14649055.2005.10766030

Schlögl, C., & Stock, W. G. (2004). Impact and relevance of LIS journals: A scientometric analysis of international and German-language LIS journals—Citation analysis versus reader survey. *Journal of the American Society for Information Science and Technology*, 55(13), 1155-1168. https://doi.org/10.1002/asi.20070

Schlögl, C., & Stock, W. G. (2008). Practitioners and academics as authors and readers: The case of LIS journals. *Journal of Documentation*, 64(5), 643-666. https://doi.org/10.1108/00220410810899691

Schlögl, C., Stock, W. G., & Reichmann, G. (2025). Scientometric evaluation of research institutions: Identifying the appropriate dimensions and attributes for assessment. *Journal of Information Science Theory and Practice*, 13(2), 49-68. https://doi.org/10.1633/JISTaP.2025.13.2.4

Stock, M., & Stock, W. G. (1990). [Psychology and philosophy of the Graz School. A documentation of the work and impact history of Alexius Meinong, Stephan Witasek, Rudolf Ameseder, Vittorio Benussi, Ernst Schwarz, Wilhelm M. Frankl, and France Veber]. Brill. German.

Stock, W. G. (1988). [Using the example of Wilhelm M. Frankl (1878-1933): Logic, natural philosophy, ontology, and philosophy of science in the Graz School]. In W. L. Gombocz, R. Haller, & N. Henrichs, (Eds.), [International bibliography of Austrian philosophy: IBÖP 1980/81] (pp. 16-89). Brill. German.

Stock, W. G. (2006). On relevance distributions. *Journal of the American Society for Information Science and Technology*, 57(8), 1126-1129. https://doi.org/10.1002/asi.20359

Stock, W. G., Dorsch, I., Reichmann, G., & Schlögl, C. (2023a). Labor productivity, labor impact, and co-authorship of research institutions: Publications and citations per full-time equivalents. *Scientometrics*, 128(1), 363-377. https://doi.org/10.1007/s11192-022-04582-5

Stock, W. G., Reichmann, G., Dorsch, I., & Schlögl, C. (2023b). Counting research publications, citations, and topics: A critical assessment of the empirical basis of scientometrics and research evaluation. *Journal of Information Science Theory and Practice*, 11(2), 37-66. https://doi.org/10.1633/JISTaP.2023.11.2.4

Stock, W. G., Reichmann, G., & Schlögl, C. (2025). Investigating the research output of institutions. *Journal of Informetrics*, 19(2), 101638. https://doi.org/10.1016/j.joi.2025.101638

Stock, W. G., & Weber, S. (2006). Facets of informetrics. *Information - Wissenschaft & Praxis*, 57(8), 385-389.

Umlauf, K., & Gradmann, S. (2011/2014). *[Lexicon of library and information science volume 2: K to Z]*. Hiersemann. German.

Uslay, C., Morgan, R. E., & Sheth, J. N. (2009). Peter Drucker on marketing: An exploration of five tenets. *Journal of the Academy of Marketing Science*, 37(1), 47-60. https://doi.org/10.1007/s11747-008-0099-8

Vargas-Quesada, B., Arroyo-Machado, W., Muñoz-Écija, T., & Chinchilla-Rodríguez, Z. (2023). Science overlay maps: A tribute to Loet Leydesdorff. *Profesional de la información*, 32(7), e320705. https://doi.org/10.3145/epi.2023.dic.05

White, H. D. (2001). Author-centered bibliometrics through CAMEOs: Characterizations automatically made and edited online. *Scientometrics*, 51(3), 607-637. https://doi.org/10.1023/A:1019607522125

JISTaP https://www.jistap.org
Journal of Information Science Theory and Practice

RESEARCH PAPER
J Inf Sci Theory Pract 13(4): 22-45, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.4.2

Exploring the Theory of Information Worlds: A Content Analysis of Its Applications and Developments

Jonathan M. Hollister
Department of Library, Archives & Information Studies, Pusan National University, Busan, Korea
E-mail: hollisterjm@pusan.ac.kr

Jisue Lee
Department of Library & Information Science, Chonnam National University, Gwangju, Korea
E-mail: jislee@jnu.ac.kr

Sungjae Park*
Knowledge and Information Culture Track, Hansung University, Seoul, Korea
E-mail: spark@hansung.ac.kr

ABSTRACT

This study explores the applications and developments of the theory of information worlds (TIW), developed by Burnett and Jaeger (2008), Jaeger and Burnett (2010), in peer-reviewed research published between 2008 and 2023. Building on the previous systematic review by Park et al. (2022), a qualitative content analysis was conducted on 27 articles that employed TIW at substantial theoretical levels, including theory application, theory conversation, and theory generation. The findings reveal that TIW serves as a robust and scalable theoretical framework, applied across diverse contexts and methodological approaches, often integrated with qualitative and mixed methods. The core concepts of normative information behavior and information value were most frequently employed and concepts such as social types were less frequently applied, indicating potential areas for further theoretical refinement. TIW has also informed new theoretical contributions, including a framework of information access in local communities and a theory of local information landscapes, and has been used in conjunction with complementary theories across disciplines. While criticisms regarding its limited treatment of individual agency and affective dimensions have been raised, recent theoretical developments incorporating cognitive and signification domains into the framework have sought to address these limitations. This study contributes to understanding the role and evolution of LIS theory and highlights the potential of TIW for future theoretical advancement and interdisciplinary application.

Keywords: theory of information worlds, theory application, theory development, theory generation, content analysis

Received: May 14, 2025
Accepted: July 17, 2025
Revised: June 26, 2025
Published: December 30, 2025

***Corresponding Author:** Sungjae Park
https://orcid.org/0000-0002-8352-068X
E-mail: spark@hansung.ac.kr

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee or authors' permission. Open Access articles are automatically archived in the Korea Institute of Science and Technology Information (KISTI)'s Open Access repository (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

© 2025 [Jonathan M. Hollister, Jisue Lee, Sungjae Park] This article is distributed under the Creative Commons Attribution License (CC BY), allowing unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. It is published by the Korea Institute of Science and Technology Information (KISTI).

1. INTRODUCTION

The field of library and information science or studies (LIS) has been criticized for its lack of theory development, as well as reliance on theories from other disciplines. McKechnie et al. (2001), Pettigrew and McKechnie (2001), and Kim and Jeong (2006) found that the use of theory in LIS research in the 1990s and early 2000s was trending upward. However, Pettigrew and McKechnie (2001) identified notable use of theories from other disciplines and a lack of use of theories from within LIS. Additionally, Kim and Jeong (2006) noted a decline in articles focusing specifically on theory development. A decline in theory use in LIS conference papers (Vakkari, 2008; VanScoy et al., 2022) and LIS research and professional articles (Julien et al., 2011) has also been noted. Other research has pointed to the increased use of theories from external disciplines, prompting calls to reinforce theory use and development from within LIS (Julien et al., 2011; Kassim, 2024; McKechnie et al., 2001; Pettigrew & McKechnie, 2001; Ukwoma & Ngulube, 2021; VanScoy et al., 2022).

To address these challenges, the present authors decided to explore and describe the use and development of a theory native to LIS: Burnett and Jaeger (2008)'s and Jaeger and Burnett (2010)'s theory of information worlds (TIW hereafter). Burnett and Jaeger (2008; 2011)'s and Jaeger and Burnett (2010)'s TIW characterizes the social contexts of information behavior through five core concepts, as described in the literature review. This study represents the third study in the series. In the first study, Park (2022) explored the use of TIW across 14 doctoral dissertations and found that TIW is often used with qualitative methods, with the interactions between the core concepts helping to describe and explain individuals' social information worlds. The second study, Park et al. (2022) expanded on the first through a systematic review of peer-reviewed research articles and found a similar trend of TIW and qualitative methods. They also found that while the adoption of TIW is trending upwards in LIS and across disciplines among scholars worldwide, relative awareness and adoption of TIW appears to be low.

The purpose of this study is to continue exploring and describing how and in what ways TIW has been applied and developed in peer-reviewed research articles. The previous study by Park et al. (2022) primarily describes quantitative trends, such as author affiliations and disciplines, publication venues, author keywords, research method types, and sample populations. This study iterates on the previous by re-using and updating the original sample and revising the analytical framework, and differs by looking more closely at how and why TIW has been applied and developed, as well as how it has informed LIS theory development and scholarship through qualitative content analysis. The qualitative content analysis of this study allows for a more nuanced and detailed exploration of the ways in which TIW has been applied to guide research design or interpret data, how it has been developed or expanded, and how it is discussed in terms of rationales for use as well as strengths and weaknesses. The research questions for this study are:

RQ1: In research using the TIW, how is it applied?
1-1 How does TIW inform research design?
1-2 Is TIW applied in part or as a whole?
1-3 What other concepts or theories are used alongside TIW?

RQ2: In research that generates or develops TIW, in what ways is TIW revised, expanded upon, or used to inform other models or theories?
2-1 What revisions or expansions to TIW are proposed?
2-2 What alternative formulations of TIW are proposed?
2-3 What new concepts, models, or frameworks are informed by TIW?

RQ3: In research that applies, develops, or substantially discusses TIW, how are its implications and applications considered?
3-1 What strengths of TIW are reported?
3-2 What weaknesses of TIW are reported?

This exploratory and descriptive study reports on the status of TIW's empirical application and theoretical development, and offers recommendations for continued application and development of TIW in future in light of the findings. This study aims to contribute to ongoing conversations about theory use and development in LIS by offering an in-depth examination of how TIW has been applied, interpreted, and adapted in LIS scholarship. This study may build awareness of a relatively newer, lesser known theory native to LIS and address concerns about the potential dearth or decline of theory in LIS.

2. BRIEF LITERATURE REVIEW

2.1. Theory Use in LIS

The concept of "theory" can be defined differently depending on the academic context and purpose. Buckland

(1991) viewed theory as a framework for explaining or expanding the essence of phenomena, whereas Popper regarded theory as a set of laws that can be empirically tested. A model, which is often used interchangeably with theory, is defined as a framework for analyzing problems and describing relationships (Wilson, 1999). However, it is distinct from a theory. Frankfort-Nachmias and Nachmias (1996) specify the difference in that a theory explains phenomena and a model describes phenomena. Therefore, the use of theory as a tool for analyzing and explaining phenomena, rather than merely describing them, can serve as a measure of research maturity and reliability (Brookes, 1980; Houser, 1988; Van Maanen, 1998).

The field of LIS has advanced through the development, application, and expansion of its own theories. LIS theories enable deeper understanding of human information behaviors, provide critical insights for the design of effective information systems, improve library collections and services, and inform the education and training of librarians and other information professionals. Since the 1990s and early 2000s, the use and application of theories in LIS research have increased (Kim & Jeong, 2006). However, Kim and Jeong (2006) note that despite this increase, articles specifically focused on developing theory were trending down. A decline in theory use has been noted by other researchers as well. Julien et al. (2011) noted decreasing theory use in information behavior research articles published between 1999 and 2008. Vakkari (2008) and VanScoy et al. (2022) identified overall decreases in theory use in LIS conference papers. Additionally, researchers have identified the trend of importing theories from neighboring disciplines, such as sociology, psychology, and business, for use in LIS (Pettigrew & McKechnie, 2001; Ukwoma & Ngulube, 2021; VanScoy et al., 2022). Interdisciplinary approaches may enhance the diversity and depth of LIS research, create new research opportunities, and speak to the meta-disciplinarity of information and information behaviors. However, these trends have spawned some existential dread among LIS scholars about the identity and theoretical foundations of the discipline.

Vakkari et al. (2022) looked at articles published in connection with authors' disciplines and found that LIS articles published by LIS scholars accounted for about one-third of the publications. They noted significant contributions from computer science, business, economics, the humanities, the natural sciences, and medicine. Vakkari et al. (2022) argue that their findings support previous arguments from Nolin and Åström (2010) that LIS is a fragmented adhocracy (Whitley, 1984), meaning that LIS is a field in which there is inconsistency or disagreement on research areas and methods and less dependence of scholars upon one another in building knowledge. Bates (2009) and Raber (2003) describe a variety of different definitions, theories, and paradigms within the discipline but note that LIS continues to progress.

Studies have been conducted to examine the extent to which theories are utilized in LIS research. Lund (2019) found that Kuhlthau (1991)'s information search process (ISP) and Bates (1989)'s berrypicking are by far the most frequently cited theories in LIS. However, the sample of theories used by Lund (2019) was determined by those included in Fisher et al. (2005)'s textbook, *Theories of Information Behavior*, which does not include TIW because it had not yet been published. Citation data in the previous study (Park et al., 2022) found that TIW was cited 68 times over 13 years and 9 months (between June 2008, when it was first introduced, to March 2022, the end of data collection), which is roughly 4.95 citations per year, suggesting low awareness or adoption of TIW. Regionally, theoretical applications in LIS studies have been investigated in countries outside the West, such as Korea (Jeong & Kim, 2005), Taiwan (Wu et al., 2017), Tanzania (Kassim, 2024), and Nigeria and South Africa (Ukwoma & Ngulube, 2021). Jeong and Kim (2005) noted that about 10% of studies published in Korean LIS journals apply theories. In Taiwanese LIS journals published between 2010 and 2015, Wu et al. (2017) found that 49.9% of articles applied theory. Ukwoma and Ngulube (2021) found that most studies from LIS postgraduate students and researchers in Nigeria and South Africa used theories from outside disciplines, primarily in management, psychology, and sociology.

The depth or level of theory application has also been explored. Notably, Kumasi et al. (2013) categorized theory talk levels into seven stages: theory dropping, theory positioning, theory diversification, theory conversation, theory application, theory testing, and theory generation. VanScoy et al. (2022) adapted Kumasi et al. (2013)'s theory talk framework to explore the use of theory in ISIC conference papers published between 1996 and 2020. They found that most (around 78%) of the conference papers mention theory, with about 69% percent of papers using theory at a substantial level. VanScoy et al. (2022) also found that most of the theories used originated from LIS, though some from other disciplines, such as sociology, psychology, communication, and business or management, were noted. Similar to Lund (2019), VanScoy et al. (2022) found that Kuhlthau's ISP was the most used LIS theory. Lund (2021) identified the prominence of ISP in dissertations as well.

Kassim (2024) also applied Kumasi et al. (2013)'s theory talk model in an exploration of theory use in articles published in Tanzanian LIS journals from 1990 to 2023. Kassim also identified an upward trend of theory use after 2006, but only 154 out of 616 (25%) of Tanzanian LIS articles overall included theory. Kassim (2024) also found that 109 (70.3%) applied theory, 29 (18.7%) dropped theories, and 17 (10.9%) generated theory. Contrary to other studies, Kassim (2024) found that the technology acceptance model (TAM; Davis, 1987) and the unified theory of acceptance and use of technology (UTAUT; Venkatesh et al., 2003) were the most often used theories at 17.4% and 14.1%, respectively.

VanScoy et al. (2022) note that while Kumasi et al. (2013)'s theory talk framework was useful for providing terminology and framing the discussion of their findings, there were a variety of limitations. In particular, they describe difficulties in applying the framework to analyze or categorize studies that may not clearly articulate why or how a theory is used, or studies that may not clearly fall into one category. They argue that more development of the theory talk model could make it useful for analyzing the use of theory. In another study applying Kumasi et al. (2013)'s theory talk framework, Park et al. (2022) also note that the framework poses challenges in identifying the usage of specific theories because it considers all theories used in a study. Park et al. (2022) adapted and revised the original framework into the single theory talk framework by applying mutual exclusivity between certain categories and repositioning and redefining the seven original levels into five levels (theory dropping, theory positioning, theory conversation, theory application, and theory generation) to analyze the application of the TIW. Park et al. (2022) found that in the context of articles featuring TIW (n=68), 51.47% were briefly name-dropping TIW; 17.65% were using TIW to position their own study or theory; 4.41% conversed on TIW; 26.47% applied TIW for study design or data analysis or interpretation; and 16.18% used TIW in some way to generate theory. This study will explore the use and application of TIW in greater detail.

2.2. Theory of Information Worlds

The TIW is a framework for analyzing the social aspects of human information behavior. The theory is based on Chatman (1999)'s small worlds theory and Habermas (1992)'s concepts of the lifeworld. Normative information behaviors within Chatman's small worlds are shaped by social norms, social types, and worldview. Burnett and Jaeger (2008; Jaeger & Burnett, 2010) modified the worldview concept into information value to account for variations in information value systems across worlds because different worlds may evaluate information differently. Additionally, the concept of boundaries was added to account for potential social interaction and exchange of information across worlds, in consideration of Habermas's concept of the lifeworld. These revisions allow TIW to provide a more holistic perspective on the broader social and cultural contexts of information behavior in a way that small worlds cannot. To study information behavior in social contexts, the TIW (Jaeger & Burnett, 2010, p. 8) presents five key core concepts, defined as follows:

- Social norms: "a world's shared sense of the appropriateness of social appearances and observable behaviors"
- Social types: "the roles that define actors and how they are perceived within a world"
- Information value: "a world's shared sense of a scale of the importance of information"
- Information behavior: "the full range of normative behaviors and activities related to information available to members of a world"
- Boundaries: "the places at which information worlds come into contact with each other and across which communication and information exchange can—but may or may not—take place."

It should be noted that there are no propositions for TIW presented alongside the core concepts. This suggests that TIW is not intended to be predictive in nature aside from the implied potential interactions amongst factors or variables associated with the core concepts. Personal communication between Burnett and Hollister in 2013 suggested that while some propositions for TIW were developed as part of a grant proposal, it was unexpected then that they would be formally published (Hollister, 2016). To date, it appears that official propositions for TIW have yet to be published and may remain unlikely to be.

The TIW has been utilized in many studies, as discussed in Park et al. (2022) and Park (2022), but at the same time, its theoretical limitations have been mentioned. There has been criticism that consideration of individual perspectives is necessary, given that TIW primarily analyzes information behavior from a social perspective (Yu, 2012). In response, Burnett (2015) proposed the distinction and integration of three domains.

Composed of the cognitive and affective domain of the individual, the domain of the social (which includes the original five core concepts), and the domain of signification and communication practices, the information domains framework includes the individual realm that was criticized as lacking in the existing theory. However, the information domains framework goes beyond simply including the individual perspective and attempts to approach from the perspective of the information lifecycle (Burnett, 2015). Rather than focusing only on information seeking and utilization, it includes five stages: information production, information encoding through media, information activity processes, interpreting agent, and acknowledgment, response, or impact, encompassing the changes that occur from information production to the final user. Just as the five concepts in TIW interact with each other, the framework's three domains are closely connected and interact with each other, making it a theory that attempts to comprehensively understand information behavior.

For a more detailed discussion of TIW's development and comparisons to other social information behavior theories, the research team recommends Burnett and Jaeger (2008; 2011), Jaeger and Burnett (2010), and Burnett (2015). As the goal of this research is to describe the use and development of TIW in peer-reviewed research literature, comparisons to other theories are outside the scope and space limitations of this manuscript.

3. METHOD

This follow-up study extends the analysis of a systematic literature review conducted by the authors (Park et al., 2022) to explore research and publication trends of peer-reviewed studies citing or applying Burnett and Jaeger (2008)'s and Jaeger and Burnett (2010)'s TIW. The focus on peer-reviewed research literature proceeded under the assumption that said literature completed higher-level, more competitive and thorough peer review processes; though, of course, this may vary across publication venues. Previous work by Park (2022) focused on TIW use in doctoral dissertations, and so Park et al. (2022) and this study opted to explore the use of TIW in peer-reviewed research articles.

This study builds on the previous by conducting a qualitative content analysis of studies featured in peer-reviewed research articles that have applied or used the TIW at higher levels, such as for theoretical conversation (coded as TC); data collection, analysis, or interpretation (coded as TA); or to inform the development of other concepts or theories (coded as TG). Through a systematic literature review approach informed by Kumasi et al. (2013), Lund (2021), and Williams et al. (2015), Park et al. (2022) identified a sub-sample of 26 peer-reviewed research articles published between 2008, the year the TIW was introduced by Burnett and Jaeger, and 2022 that used the theory at higher levels. Research articles considered for inclusion must have been peer-reviewed, published in English or Korean, and available in full-text via online websites, institutional repositories, author websites, and academic databases accessible to the researchers through their respective institutional libraries.[1] Given the research team's language proficiencies, the sample included articles published in English or Korean. Sampled articles must have cited at least one of the theory's core works: Burnett and Jaeger (2008; 2011) or Jaeger and Burnett (2010).

The core sample of 26 previously identified articles was expanded using the same systematic literature review approach to include relevant articles published through the end of 2023, which yielded 23 potential new titles given the original selection criteria, as mentioned above. However, after applying the revised single theory talk model, as discussed below, the new TC, TA, and TG additions were narrowed to three titles, and two of the previously collected articles from the original subsample were removed as they no longer met the qualifications to be coded as TG papers. Upon further review during revisions to the single theory talk model, two papers (Oh & Butler, 2019; Riley et al., 2022) previously coded as TP/TG were recategorized as TP only and thus removed from the sample of this study, as while they did identify TIW, it was for positioning (TP) and not for generating new theory (TG). As such, the sample for this study includes 27 articles (n=27).

The previous study (Park et al., 2022) categorized and coded articles using theory at higher levels with a single theory talk model that was adapted from the theory talk framework by Kumasi et al. (2013). During initial codebook development conversations for this study, the single theory talk model was found to need revision and expansion to guide qualitative content analysis and provide additional depth to account for the different ways theory can be used at higher levels. The single theory talk model, as

[1]These included ResearchGate: https://www.researchgate.net; GoogleScholar https://scholar.google.com; EBSCOHost; Emerald Specialist Collection; Library, Information Science and Technology Abstracts [LISTA]; ProQuest Central; ScienceDirect; SCOPUS; Web of Science; and Wiley databases.

noted in Table 1, was revised primarily to expand the Theory Application (TA) and Theory Generation (TG) codes from the previous study to provide great specificity to account for the different ways theory can be applied, developed, or used to generate new theoretical constructs. The higher-level code of Theory Conversation (TC) and the lower levels of use (0, TD, and TP) were left unchanged.

In this study, the revised TA code was split into two subcategories: Theory Application 1 - Research Design & Data Analysis (TA1), which refers to the use of TIW for

Table 1. Revised single theory talk model

Category	Code	Adapted operational definition	Notes
Not Applicable	0	Not Applicable (0) is applied when the manuscript does not directly identify or address TIW or its core concepts, even if the core TIW publications are cited. This is used as an exclusionary criterion for screening purposes	Unchanged
Theory Dropping	TD	Theory Dropping (TD) is the mention of TIW or its core concepts within the manuscript (with accompanying citations to core TIW publications) but without a clear connection to the main focus of the article, research design, or analysis	Unchanged
Theory Positioning	TP	Theory Positioning (TP) is the short discussion, description, or critique of TIW or its core concepts alongside or in relation to other theories used to provide context for the study/manuscript at hand but not as a main focus of the manuscript, nor to guide the research design or analysis. This may include using TIW or core concepts to discuss the findings/results	Unchanged
Theory Conversation	TC	Theory Conversation (TC) refers to a focused, in-depth discussion or critique of TIW, its core concepts, or its theoretical implications or practical applications within LIS or in other fields. This category is limited to articles that specifically discuss TIW or its core concepts as part of the core focus of the paper but do not report empirical research using the theory. This may occur in the literature review, discussion, or conclusion sections	Unchanged
Theory Application	TA	Theory Application (TA) refers to the application of TIW or its core concepts to the research design or methodology (research questions, design, or method); the analysis or interpretation of empirical data; or the evaluation, assessment, testing, or validation of TIW or its core concepts	Split into TA-1 & 2
Theory Application 1 - Research Design & Data Analysis	TA1	Theory Application 1 (TA-1) refers to the application of TIW or its core concepts to the research design or methodology (research questions, design, or method), or for the analysis or interpretation of empirical data	
Theory Application 2 - Theory Convergence	TA2	Theory Application 2 (TA-2) refers to the use of TIW or its core concepts alongside other theories or theoretical concepts for the analysis or interpretation of data	Automatically includes TA-1
Theory Generation	TG	Theory Generation (TG) refers to the theoretical development, revision, or expansion of TIW or its core concepts through empirical research or inductive reasoning. This may entail the expansion or revision of TIW, the proposal of propositions, or the building or creation of new theories	Split into TG-1, 2, 3, & 4.
Theory Generation 1 - Refinement/ Revision	TG1	Theory Generation 1 (TG-1) - Refinement refers to a proposed refinement or revision to one or more of the existing 5 core concepts of TIW	May overlap with TC
Theory Generation 2 - Expansion	TG2	Theory Generation 2 (TG-2) - Expansion refers to the proposed addition of one or more new concepts alongside the original 5 core concepts of TIW	New concepts may be derived from the existing theory. May overlap with TC, TA-2
Theory Generation 3 - Alternative	TG3	Theory Generation 3 (TG-3) - Alternative refers to the proposal of an alternative information worlds theory or another similarly scoped theory	May overlap with TC
Theory Generation 4 - Informed	TG4	Theory Generation 4 (TG-4) - Informed refers to new or emergent concepts, models, or frameworks developed through the application or use of TIW to explain or interpret a phenomenon, but not intended as replacements or alternatives to TIW	May overlap with TP, TC, TA-1, TA-2, and other TG categories

TIW, theory of information worlds.

research design or data analysis; and Theory Application 2 - Theory Convergence (TA2), which refers the use of TIW alongside other theories for the analysis or interpretation of data. Additionally, the revised TG was split into four subcategories: Theory Generation 1 - Refinement/Revision (TG1), when revisions or refinements to TIW are introduced; Theory Generation 2 - Expansion (TG2), when new concepts beyond the original five core concepts are introduced; Theory Generation 3 - Alternative (TG3), when alternative information worlds theories are proposed; and Theory Generation 4 - Informed (TG4), when TIW is used to inform the development of new or emergent concepts, models, or theories that are not alternatives to TIW. The TG1 and TG2 codes are meant to capture the theoretical development of TIW itself through revision or expansion. TG3 and TG4 are meant to capture how TIW may inspire, inform, or build other concepts, models, or theories. The TG3 code is meant to identify theories proposed as alternatives or replacements for Burnett and Jaeger's TIW, and the TG4 code captures concepts, models, frameworks, or theories that are informed by TIW but are not intended as alternatives or replacements. As noted in Table 1, the codes are not mutually exclusive. Furthermore, while the research team considers the continued application of the theory across disparate research contexts, use with different methods, and use alongside other theories as ways to indirectly contribute to theory development as forms of theory testing, application of the theory does not always imply or guarantee theory development or generation. As seen in Table 2 (Bao & Ke, 2023; Brattland, 2017; Burnett, 2015; Burnett et al., 2009; Burnett et al., 2013; Froggatt, 2015; Gibson, 2016; Gibson & Kaplan, 2017; Hollister, 2020; Hollister et al., 2023; Hovious, 2018; Jaeger et al., 2014; Känsäkoski & Huotari, 2016; Lee, 2019; Lee & Butler, 2019; Ndumbaro & Ochieng, 2021; Ndumu, 2019; 2020; Pang et al., 2020; Park, 2022; Park et al., 2022; Rohman & Pitaloka, 2025; Welch, 2019; Worrall, 2019; Worrall et al., 2021; Xie et al., 2021; Yu, 2012) below, coding an article with TA does not always mean it was also coded with TG.

The revised single theory talk model provided the framework for directed content analysis to categorize and analyze the sampled articles (Hsieh & Shannon, 2005). Additional qualitative content analysis was conducted using open coding to identify content relevant to the research questions and any related emergent themes.

Intercoder reliability testing and consensus building were used to ensure the consistency and trustworthiness of the codebook and data analysis. A collaborative approach similar to Cascio et al. (2019) was used to establish intercoder consensus with each of the authors of this study serving as coders. During codebook development, each article was assigned two coders in a sequence that allowed the coders to be evenly distributed across the sample, resulting in each author coding two-thirds of the sample. Coding during initial codebook development was done in batches of 3-5 articles, followed by a meeting to discuss the coding as a whole research team and adjust the codebook as needed. Given the small size and nature of the sample (n=27), the entire research team applied the final codebook to each of the articles, yielding an intercoder reliability rating of 92.59%, which is well above an acceptable level of agreement (80%) as argued by Lombard et al. (2002). Open coding and the resulting emergent themes identified during additional content analysis, as guided by research questions, were vetted through cumulative discussion and consensus-building. Additional intercoder reliability testing was conducted to explore whether TIW is applied fully (91.67% intercoder reliability rating, n=20) in TA-coded articles, as well as whether the core concepts are present in the results/findings or discussion sections of the TA-coded articles (89.167% agreement, n=20).

This study is limited in scope by focusing on peer-reviewed research articles published between 2008 and 2023, and is limited by the availability of full-text access and library database subscriptions. The limitations in scope, sampling, and the use of qualitative methods mean the findings are not generalizable or predictive. However, the trends and themes identified may also exist in types of research materials employing TIW, such as conference papers, dissertations, or research articles not included in the sample of this study. Additionally, the systematic literature review approach, intercoder reliability testing, and consensus building enhance the dependability and confirmability of the findings.

To ensure transparency and credibility, it should be noted that all members of the research team have previous work related to TIW. Hollister (2020) and Lee (2019) have applied TIW in empirical research and worked with Burnett in a TIW-focused theory study group (Hollister et al., 2023), all of which are included in the sample of this paper. As noted in the introduction and literature review, previous studies led by Park (2022) (Park et al., 2022) have explored the use of TIW, and this study extends those explorations. The core analysis of this paper is structured around the revised single theory talk model, not TIW. As noted in Section 2.1, Kumasi et al. (2013)'s theory talk framework and adaptations thereof have been applied in

Table 2. Higher TIW theory use articles coding results

Sampled articles' authors & titles (n=27)	Coding
Bao, X., & Ke, P. (2023). Chaos, expansion, and contraction: The information worlds of depression patients during the COVID-19 pandemic lockdown.	TA2
Brattland, K. (2017). Information preferences of Reddit communities surrounding the Brock Turner case.	TA2
Burnett, G. (2015). Information worlds and interpretive practices: Toward an integration of domains.	TC, TG2
Burnett, G., Whetstone, M., & Jaeger, P. T. (2013). Personal health record interfaces: A hermeneutic analysis.	TA1
Burnett, K., Subramaniam, M. M., & Gibson, A. (2009). Latinas cross the IT border: Understanding gender as a boundary object between information worlds.	TA2
Froggatt, D. L. (2015). The informationally underserved: Not always diverse, but always a social justice advocacy model.	TA2, TG4
Gibson, A. N. (2016). Building a progressive-situational model of post-diagnosis information seeking for parents of individuals with Down syndrome.	TA2, TG4
Gibson, A. N., & Kaplan, S. (2017). Place, community and information behavior: Spatially oriented information seeking zones and information source preferences.	TA2, TG4
Hollister, J. M. (2020). The information worlds of online role-players.	TA1
Hollister, J. M., Burnett, G., Lee, J., & Skinner, J. (2023). Collaborative codebook development across information worlds: Intersections between pedagogy, theory, and method in three dissertations.	TC
Hovious, A. (2018). Toward a socio-contextual understanding of transliteracy.	TG4
Jaeger, P. T., Gorham, U., Bertot, J. C., Taylor, N. G., Larson, E., Lincoln, R., Lazar, J., & Wentz, B. (2014). Connecting government, libraries and communities: Information behavior theory and information intermediaries in the design of LibEGov.org.	TA2, TG2
Känsäkoski, H., & Huotari, M. L. (2016). Applying the theory of information worlds within a health care practise in Finland.	TA1
Lee, J. (2019). An exploratory study on the political information behaviors of Korean opinion leaders on Twitter: Through the lens of theory of information worlds.	TA1
Lee, M., & Butler, B. S. (2019). How are information deserts created? A theory of local information landscapes.	TP, TG4
Ndumbaro, F., & Ochieng, L. M. (2021). Access to information on family planning (FP) methods among married women of reproductive age in Ilala District, Dar es Salaam Tanzania.	TA1
Ndumu, A. (2019). Linkages between information overload and acculturative stress: The case of Black diasporic immigrants in the US.	TA2
Ndumu, A. (2020). Toward a new understanding of immigrant information behavior: A survey study on information access and information overload among US Black diasporic immigrants.	TA1
Pang, N., Karanasios, S., & Anwar, M. (2020). Exploring the information worlds of older persons during disasters.	TA2
Park, S. (2022). A reflection on social perspective of information behavior: Around the theory of information worlds.	TC
Park, S. Lee, J., & Hollister, J. M. (2022). A systematic review on the application of the theory of information worlds.	TC
Rohman, A., & Pitaloka, D. (2025). Disconnected and disabled during the pandemic: Toward more inclusive pandemic response plans in the Global South.	TA1
Welch, C. (2019). Circling the reality of public institutions: An e-Delphi study of information, culture, and community stakeholders voicing the way forward in Detroit, Michigan.	TA2
Worrall, A. (2019). "Connections above and beyond": Information, translation, and community boundaries in LibraryThing and Goodreads.	TA2
Worrall, A, Cappello, A, & Osolen, R. (2021). The importance of socio-emotional considerations in online communities, social informatics, and information science.	TA1, TG4
Xie, J., He, Z., Burnett, G., & Cheng, Y. (2021). How do mothers exchange parenting-related information in online communities? A meta-synthesis.	TA1
Yu, L. (2012). Towards a reconceptualization of the 'information worlds of individuals.'	TP, TG3

TIW, theory of information worlds.

other explorations of theory use in the research literature. The dependability of the qualitative analysis is supported by a robust approach to inter-coder reliability testing and consensus building, which used the entire sample rather than a subsample and yielded high levels of agreement (O'Leary, 2005; Pickard & Childs, 2013). To further enhance trustworthiness and confirmability and to avoid or mitigate potential interpretive bias, the findings and interpretations thereof are supported with evidence from the data (Pickard & Childs, 2013).

4. FINDINGS

4.1. Applications of TIW (RQ1)

This study analyzed 27 papers that applied the TIW. The codes used in TA papers, which applied the theory, and TG papers, which provided new theoretical concepts, are TP, TC, TA1, TA2, TG2, TG3, and TG4. Although TP papers, which use TIW to position a study without applying it to research design or data analysis, were not the subjects of analysis, the corresponding codes were included in this study's analysis when used alongside other TC, TA, or TG codes. A total of 35 codes were assigned across the 27 papers, with an average of approximately 1.3 codes per paper. The most frequently assigned codes were TA1 and TA2, which appeared in 10 papers each. Following those were TG4, which appeared six times, TC four times, TG2 and TC two times each, and TG3 one time. Abbreviated citations of the articles and their associated codes are listed in Table 2 (Bao & Ke, 2023; Brattland, 2017; Burnett, 2015; Burnett et al., 2009; Burnett et al., 2013; Froggatt, 2015; Gibson, 2016; Gibson & Kaplan, 2017; Hollister, 2020; Hollister et al., 2023; Hovious, 2018; Jaeger et al., 2014; Känsäkoski & Huotari, 2016; Lee, 2019; Lee & Butler, 2019; Ndumbaro & Ochieng, 2021; Ndumu, 2019; 2020; Pang et al., 2020; Park, 2022; Park et al., 2022; Rohman & Pitaloka, 2025; Welch, 2019; Worrall, 2019; Worrall et al., 2021; Xie et al., 2021; Yu, 2012).

Since a paper can have multiple codes, the study analyzed papers that applied the theory through combinations of codes. When considering combinations of codes, 10 code combinations were used, and the number of papers per combination is shown in Fig. 1. The code that appeared in most papers was TA1 (in eight papers). This was followed by TA2 in seven papers, and TC and TG4 in three papers each.

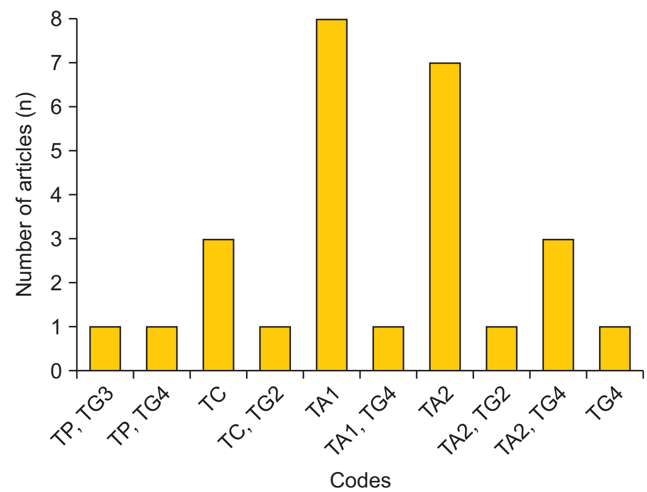

Fig. 1. The number of articles per unique code. Please see Section 3 or Table 1 for more information. TP, Theory Positioning; TC, Theory Conversation; TA1, Theory Application 1 - Research Design & Data Analysis; TA2, Theory Application 2 - Theory Convergence; TG2, Theory Generation 2 - Expansion; TG3, Theory Generation 3 - Alternative; TG4, Theory Generation 4 - Informed.

4.1.1. RQ1-1: TIW-Informed Research Designs

In the research design process, TIW was utilized for grounded theory data collection guidance (Gibson, 2016), the development of survey instruments (Froggatt, 2015), and the design of interviews and their protocols (Froggatt, 2015; Gibson & Kaplan, 2017). Although the analysis was constrained by the limited space available in academic journal articles, which restricts access to detailed survey or interview questions, the analysis techniques addressed key concepts of TIW. Additionally, the authors of those works explicitly mentioned the use of TIW in creating their research instruments, confirming its application in the research design phase.

4.1.2. RQ1-2: The Application of TIW Core Concepts

To investigate how the TIW was applied in the sampled papers, the analysis was divided into two sections: research design and research results. The research design provides the overall framework for conducting the study, including the literature review and methodology. The literature reviews of the sampled papers were examined to determine whether TIW and its five key concepts were explained, and the methodology sections of the sample papers were assessed as to whether TIW was used during the development of research instruments. The analysis revealed that most studies, except one, referred to TIW and its key concepts at the research design stage. Among the 20 analyzed papers, 12 studies (60%) established their research questions based on TIW. Notably, Känsäkoski and

Huotari (2016) formulated the following research question: "Which concepts of the theory by Burnett and Jaeger can be identified as elements of the organizational culture that enhance or inhibit the value creation of the ICP?" This serves as a direct example of TIW being used in research question development. Furthermore, Burnett et al. (2009) employed TIW to explore how successful Latina IT professionals construct and reconstruct their conceptions of gender before, during, and after engaging with the information world of the IT industry.

To illustrate how the five core concepts of the TIW were applied across the selected studies, we present representative excerpts from the analyzed literature. These excerpts demonstrate how each concept—social norms, social types, information behavior, information value, and boundaries—was interpreted and operationalized within different research contexts.

Social norms were often reflected in community expectations and implicit behavioral codes. In Xie et al. (2021), mothers participating in online parenting communities "generally achieved a sense of what is and is not 'normal' through reading about others' experiences" (pp. 5-6). This finding illustrates how shared narratives and interactions contribute to constructing social norms that shape acceptable information practices within a particular information world.

Social types were used to classify individuals according to their perceived roles and patterns of engagement in information environments. In Bao and Ke (2023), participants were observed to "play multiple information roles and 'social types'; they are information demanders, receivers, and providers" (p. 982). This conceptualization of role-based identities influenced how information was circulated, negotiated, or restricted within the group.

Information behavior, as expected, appeared most frequently across the studies and encompassed practices of seeking, avoiding, or sharing information. Ndumu (2019) emphasizes that "stress, anxiety, and feelings are vital considerations when analyzing immigrants' engagement with information" (p. 4), underscoring the importance of affective dimensions in shaping individual information practices.

Information value was discussed in relation to the criteria individuals use to assess the relevance, credibility, or ideological alignment of information. For example, in Welch (2019), among the top five issues listed by both community advocates and library, archives, and museum (LAM) practitioners was "Foundations exhibit White supremacist values that frame communities as needing to be saved or be complacent" (p. 254). This observation highlights how deficit-oriented narratives were perceived to be devalued by foundations, while critical, anti-racist, and community-empowering knowledge was regarded as more meaningful to both community advocates and LAM practitioners. Welch (2019) provides examples of how information values can be shared or differ across and between stakeholders, community groups, and organizations within interrelated or overlapping information worlds.

Boundaries, defined as conceptual or structural limits to participation or access, consistently emerged as a salient factor, particularly among marginalized groups. In their study on disaster information access, Pang et al. (2020) found that older adults "did not report using social media, the Internet, or mobile applications," a finding the authors attribute to the "age and socioeconomic profile of the subjects" (p. 628). This demonstrates how demographic and technological divides can function as barriers to information engagement in urgent contexts.

Together, these excerpts offer a glimpse into the diversity of ways TIW concepts are interpreted and embedded in empirical work, reinforcing the theory's adaptability across domains and populations.

The results of examining the usage of the five concepts from the TIW in the results and discussion sections are shown in Table 3. The concept most frequently discussed in the results is "information behavior," appearing in 16 TA papers (80%). This seems to be because the TIW was designed to analyze people's information behavior, leading to the frequent use and analysis of the concept of information behavior. Next, the concept of "information value" appeared in the results of 15 papers, accounting for 75% of all TA papers. The least frequently mentioned concept was "social types," which appeared in the results section of

Table 3. Use of TIW core concepts in findings and discussion sections (n=20)

	Social norms	Social types	Information value	Information behavior	Boundaries
Findings	11 (55)	7 (35)	15 (75)	16 (80)	11 (55)
Discussion	15 (75)	12 (60)	16 (80)	17 (85)	15 (75)

Values are presented as number (%).
TIW, theory of information worlds.

seven papers, or 35% of the total papers.

In the discussion sections of the sampled articles, the most frequently mentioned concept was also "information behavior," followed by "information value," which appeared in 80% of the papers. Similar to the trends in the results section, the concept of "social types" appeared in the discussion section of 12 papers, but it was still the least frequently mentioned among the five concepts.

4.1.3. RQ1-3: Complementary Concepts or Theories

TIW has been used in conjunction with other theoretical frameworks or concepts in scholarly literature. Table 3 summarizes the frequencies of complementary theories and concepts that appeared alongside TIW in the 27 articles analyzed. This analysis addresses Research Questions 1-3 (RQ1-3), which seek to identify what other theories or concepts are used alongside TIW and to what extent TIW operates independently as a theoretical framework.

Table 4 (Bowker & Star, 1999; Brattland, 2017; Burnett et al., 2009; Charmaz, 2014; Corbin & Strauss, 2008; Currie & Suhomlinova, 2006; Dunne, 2002; Elder et al., 2003; Engeström, 1987; Froggatt, 2014; 2015; Gibson, 2016; Gibson & Kaplan, 2017; Harding, 2015; Hovious, 2018; Huotari & Chatman, 2001; Jaeger et al., 2014; Jones, 1997; Känsäkoski, 2014; Känsäkoski & Huotari, 2016; Kuhlthau, 1991; Ndumu, 2019; 2020; Pang et al., 2020; Rudmin, 2009; Ruthven, 2022; Savolainen, 2008; Sonnenwald, 1999; Star, 1989; Star & Griesemer, 1989; Strauss, 1987; Welch, 2019; Wenger, 1998; Worrall, 2019) presents 19 distinct theories and concepts that were used in combination with TIW across the sample. The category with the highest frequency in Table 4 (Bowker & Star, 1999; Brattland, 2017; Burnett et al., 2009; Charmaz, 2014; Corbin & Strauss, 2008; Currie & Suhomlinova, 2006; Dunne, 2002; Elder et al., 2003; Engeström, 1987; Froggatt, 2014; 2015; Gibson, 2016; Gibson & Kaplan, 2017; Harding, 2015; Hovious,

Table 4. Frequency of complementary concepts or theories

# of papers mentioning (TA2, n=10)	Concepts or theories and primary cited sources	As cited in
7	None	Not applicable
2	Acculturative stress (Rudmin, 2009)	Ndumu (2019; 2020)
	Activity theory (Engeström, 1987)	Hovious (2018) Pang et al. (2020)
	Social worlds perspective/boundary objects/translation (Bowker & Star, 1999; Star, 1989; Star & Griesemer, 1989; Strauss, 1987)	Burnett et al. (2009) Worrall (2019)
	Grounded theory (Charmaz, 2014; Corbin & Strauss, 2008)	Gibson (2016) Gibson & Kaplan (2017)
	Information overload (Bawden & Robinson, 2009; 2013)	Ndumu (2019; 2020)
1	Bridges (Jaeger et al., 2014)	Jaeger et al. (2014)
	Everyday life information seeking (Huotari & Chatman, 2001)	Känsäkoski & Huotari (2016)
	Communities of practice (Wenger, 1998)	Brattland (2017)
	Dynamic structural model of racism (Jones, 1997)	Welch (2019)
	Everyday information practice (Savolainen, 2008)	Froggatt (2015)
	Information horizon theory (Sonnenwald, 1999)	Gibson & Kaplan (2017)
	Information search process (Kuhlthau, 1991)	Froggatt (2015)
	Life course theory (Elder et al., 2003)	Gibson (2016)
	Organisational culture (Currie & Suhomlinova, 2006)	Känsäkoski & Huotari (2016)
	Value creation (Huotari & Chatman, 2001; Känsäkoski, 2014)	Känsäkoski & Huotari (2016)
	Persons-in-progressive-situations (Dunne, 2002)	Gibson (2016)
	Rape culture (Harding, 2015)	Brattland (2017)
	Theory of informationally underserved (Froggatt, 2014)	Froggatt (2015)
	Transition (Ruthven, 2022)	Bao & Ke (2023)

 https://doi.org/10.1633/JISTaP.2025.13.4.2

2018; Huotari & Chatman, 2001; Jaeger et al., 2014; Jones, 1997; Känsäkoski, 2014; Känsäkoski & Huotari, 2016; Kuhlthau, 1991; Ndumu, 2019; 2020; Pang et al., 2020; Rudmin, 2009; Ruthven, 2022; Savolainen, 2008; Sonnenwald, 1999; Star, 1989; Star & Griesemer, 1989; Strauss, 1987; Welch, 2019; Wenger, 1998; Worrall, 2019) is "None" (seven occurrences), indicating that many studies used TIW as a standalone theory. This suggests that TIW can function effectively as an independent theoretical framework. Acculturative stress (Rudmin, 2009), activity theory (Engeström, 1987), social world perspective, boundary objects, and translation (Bowker & Star, 1999; Star, 1989; Star & Griesemer, 1989; Strauss, 1987), grounded theory (Charmaz, 2014; Corbin & Strauss, 2008), and information overload (Bawden & Robinson, 2009; 2013) each appear twice, making them the most commonly used complementary theories alongside TIW in the sample. Notably, acculturative stress and information overload were repeatedly used in two papers by the same author on similar research projects, and activity theory and boundary objects were applied in two separate papers by different authors.

In addition, 14 concepts or theories appeared only once alongside TIW. In the field of library and information science, theories such as Kuhlthau (1991)'s ISP, Chatman's Everyday Life Information Seeking (Huotari & Chatman, 2001), and Information Horizon Theory (Sonnenwald, 1999) were used in conjunction with TIW. There were also cases where new concepts, rather than established theories or models, were introduced. For instance, Jaeger et al. (2014) incorporated the concept of bridges to conduct their analysis. Froggatt (2015)'s study applied TIW alongside Savolainen (2008)'s everyday information practices and Kuhlthau (1991)'s ISP to develop a theory of the informationally underserved (IU), which describes student success and academic achievement in relation to the accessibility or availability of school library programs.

The diverse range of complementary theories reveals the epistemological versatility of TIW. Drawing on TIW, researchers have successfully integrated frameworks that span cognitive (e.g., ISP), affective (e.g., acculturative stress), social-structural (e.g., boundary objects), and methodological paradigms (e.g., grounded theory). The analysis of concepts or theories used with TIW demonstrates that TIW is a robust tool for analyzing information phenomena on its own. Moreover, it indicates that when combined with other tools, TIW serves as a useful framework for addressing research questions in diverse contexts.

4.2. Theory Revision, Expansion & Generation with the Theory of Information Worlds (RQ2)

As noted above, nine articles were involved in theory generation using TIW. However, none of the articles propose revisions to the existing core concepts of TIW (TG1, n=0).

4.2.1. RQ2-1: Expansions to TIW (TG2)

Expansions to TIW are proposed in two articles (TG2, n=2): Burnett (2015) and Jaeger et al. (2014). Burnett (2015) proposes an expansion to the TIW by introducing the information domains framework. Two new domains, the domains of the individual and of signification, are added, with the original TIW being repositioned unchanged as the domain of the social. The domain of the individual is meant to capture individual factors, such as cognition, affect/emotion, agency, etc., to account for, in part, limitations raised by Yu (2012), as well as to encompass myriad other theories and research on individual information behaviors, such as Dervin (1992)'s sense-making and Kuhlthau (1991)'s ISP, among others. The domain of signification accounts for the various languages, representations, tools, and technologies through which information is stored and communicated. The domain of signification is influenced by semiotics and hermeneutics and can also account for physical representations of information as addressed in the LIS literature, such as Buckland (1991)'s information-as-thing, among others.

Jaeger et al. (2014) introduce a potential sixth core concept: bridges. Bridges are defined as "the tangible (physical or digital) mechanisms and channels that facilitate the exchange of information and interaction across boundaries" (Jaeger et al., 2014; para. 13). Through an analysis of an e-government website (LibEGov.org) using TIW, they found that the concept of bridges could characterize the roles in supporting librarians to use and provide e-government services and resources in their libraries, communicate with government agencies, and establish and follow best practices. Bridges may also help describe how information moves across or between information worlds, across boundaries.

4.2.2. RQ2-2: Alternative Formulations of TIW (TG3)

Alternative formulations of TIW are presented in one article (TG3, n=1): Yu (2012). As noted in the literature review, Yu (2012) introduces a TIW of individuals because they argue that Burnett and Jaeger's TIW does not explicitly define an individual's information world nor differentiate one's individual information world from one's social worlds (Yu, 2012, p. 4). Yu (2012)'s information worlds of the individuals can be defined as the "space-time-intellect delimited life sphere in which information or knowledge afforded by Popper's worlds 1, 2, and 3 is converted into personal information assets through intentional, conscious and involuntary information practices that are performed by the individual as an information creator, provider, transmitter, seeker, receiver, and user" (Yu, 2012, p. 15). Yu's TIW of individuals is informed by Popper's three worlds: physical, objective knowledge, and mental states (Popper, 1979). Yu argues that the boundaries of the information worlds of individuals are based on an individual's information practices rather than social norms, in contrast to Chatman (1991)'s information or small worlds.

4.2.3. RQ2-3: New Concepts or Theories Informed by TIW (TG4)

New concepts or theories informed by TIW are identified in six articles (TG4, n=6): Froggatt (2015); Gibson (2016); Gibson and Kaplan (2017); Hovious (2018); Lee and Butler (2019); and Worrall et al. (2021).

In a mixed-methods study, Froggatt (2015) explores the impact of access to active school library programs (ASLP) on students' academic performance. Froggatt (2015) uses TIW as the basis of a proposed model for the theory of the IU. The IU model proposes four information worlds of students: IW-A, where students have access to ASLP and achieve academic success; IW-B, where students have access to ASLP but have no or low academic success; IW-C, where students do have limited or no access to ASLP and still have academic success; and IW-D, where students have limited or no access to ASLP and have no or low academic success. Froggatt (2015) notes that their study does not adequately address IW-B. Access to public libraries and familial support may account for IW-C, and IW-D correlates with low standardized testing scores. Froggatt (2015) argues that the implications of the IU theory and findings stress the importance of information access and support Rioux (2010)'s calls for integrating social justice meta-theory into LIS scholarship, education, and professional practice.

In interviews with the parents of children with Down syndrome, Gibson (2016) used TIW as one of the sensitizing theories to guide data collection and derived a model of post-diagnosis information-seeking behavior. The model represents a progressive process impacted by cognitive, affective, behavioral, situational, and contextual factors. Further analysis and interpretation were informed in part by TIW by Gibson and Kaplan (2017) on the same interview sample with the addition of social network mapping to create a framework for information access in local communities. The framework seeks to understand how place and social contexts impact information seeking (Gibson & Kaplan, 2017, p. 138). The framework consists of four facets: "1. Spatial information access zones in terms of distance from the information seeker; 2. Information seeking effort (in this case measured by minutes of travel); 3. Preferences for fulfillment of information needs in certain zones; 4. Measuring fulfillment of those preferences or expectations."

Through a systematic literature review, Hovious (2018, p. 183) combines TIW with activity theory (Engeström, 1987) to redefine transliteracy as the "ability to adapt a literacy practice successfully across activity systems." Hovious (2018, p. 182) argues that TIW is a "useful framework for understanding literacy within information technology-based contexts."

Lee and Butler (2019) use TIW to position and propose a theory of local information landscapes. They note that TIW helps to explain how social and contextual factors influence information access and behaviors. However, the theory of local information behavior is not derived from TIW, but rather informed or situated by it. The theory of local information landscapes is presented as a supplement to TIW and related theories rather than as an alternative and focuses more on information provision and its impacts on information behavior and access.

Through a mixed methods study using TIW, Worrall et al. (2021) propose and advocate for a paradigm shift in LIS research towards an integration of the social and affect paradigm, and propose a preliminary and informal model of a socio-emotional information science research paradigm. They argue that social theories and paradigms within information science, presumably TIW included, do not adequately develop or integrate affective factors into social information behavior research.

4.3. Theory Conversations with TIW (RQ3)

TC-coded articles discuss the theoretical implications and practical applications of TIW as well as offer critiques and other considerations, but do not use TIW to inform research design or collect and analyze empirical data. TC-coded articles, as well as those coded with TA and TG, identified and discussed the strengths and weaknesses of TIW. Consequently, the strengths and weaknesses of TIW identified across the sample are discussed in Sections 5.3 and 5.4, respectively, with supporting evidence.

Four articles within the sample are coded for theory conversation (TC): Burnett (2015); Hollister et al. (2023); Park (2022); and Park et al. (2022). Burnett (2015) presents a theory development paper that introduces an expansion of core TIW with the addition of the domains of individual and signification; the original core concepts are repositioned as the domain of the social. The new information domains framework can account for other theoretical paradigms and existing research in LIS, such as cognitive and physical paradigms (using Raber's 2003 interpretations).

Hollister et al. (2023) describe the development of a TIW general codebook through a collaboration of three (then) doctoral candidates and Burnett. They found that TIW provides a scalable and flexible theory that could be successfully applied with a variety of research methods and in disparate research contexts. Park (2022) and Park et al. (2022) are precursors to the study presented here, exploring the use and applications of TIW in dissertations and peer-reviewed research articles, respectively. Both studies find that while TIW is often employed with qualitative methods, it has and continues to be applied with mixed and quantitative methods and is used to study social information behavior in a variety of populations around the world.

5. DISCUSSION

The findings are first summarized and discussed in order of the primary research questions. Given the nature of the TC code, articles featuring substantial discussion or critique of TIW will be described and discussed. Strengths and weaknesses identified during the full analysis are also discussed.

5.1. Theory Application

TIW has been demonstrated to be a versatile and robust theoretical framework applicable across various research contexts. The findings show that TIW effectively operates as an independent analytical tool, with many studies employing it exclusively to explain information phenomena. This highlights the theory's self-contained theoretical structure, enabling researchers to leverage its value in specific research settings. TIW's core concepts, particularly "information behavior" and "information value," were frequently utilized, underscoring the theory's optimization for analyzing information behavior. In particular, information value can be considered a unique concept of TIW, as it represents a shift from the world view concept in the process of developing Chatman's original theory into TIW (Jaeger & Burnett, 2010). In the TC article, Hollister et al. (2023), there are specific examples of how to operationalize and apply the information value concept across three different research projects, which may be useful for those thinking about using the information value concept or TIW more fully. However, concepts such as social types and boundaries have been discussed relatively less frequently, and research on social types appears to be limited due to the difficulty of measuring the perception of assigned roles within society. Therefore, to enhance the extensibility of TIW, it may be necessary to clearly define and operationalize the concept of social types. As noted in Hollister et al. (2023), a draft of the general TIW codebook is available by request, which may help scholars to operationalize and apply the TIW concepts to guide their research and interpret their data.

TIW has also made contributions to research design by guiding the development of research questions, data collection methods, and research instruments. For example, TIW was useful in the creation of interview protocols (Gibson & Kaplan, 2017) and survey tools (Worrall, 2019), demonstrating its practical utility in empirical research. When combined with qualitative research methodologies, such as grounded theory (Gibson, 2016), TIW facilitated systematic analysis of complex social phenomena. Its core concepts were explicitly referenced during study design, ensuring that data collection and interpretation were aligned with the theory's foundational principles. Methodologically, TIW has been leveraged in diverse research designs, including mobile ethnography (Pang et al., 2020), sentiment analysis (Worrall et al., 2021), and cognitive mapping (Bao & Ke, 2023). These applications illustrate TIW's adaptability to both qualitative and quantitative methods, reinforcing its role as a versatile analytical framework. However, despite its strengths, challenges remain in quantifying individual cognition and emotions within TIW. Future research should focus on refining operational definitions of social types, integrating quantitative models, and developing theoretical frameworks that incorporate effective computing to enhance TIW's explanatory power.

TIW has demonstrated strong synergy with various theoretical frameworks, enhancing its applicability across diverse research contexts. It has been integrated with Boundary Object Theory (Worrall, 2019), Information Overload (Ndumu, 2019; 2020), and Grounded Theory (Gibson, 2016; Lee, 2019) to provide deeper insights into information behavior. Additionally, it has been applied alongside activity theory to analyze transliteracy (Hovious, 2018) and the information behaviors of rural seniors during disasters (Pang et al., 2020), and with Social Informatics to examine socio-emotional interactions in online communities (Worrall et al., 2021). These integrations collectively highlight TIW's flexibility in capturing both structural and affective dimensions of information behavior, demonstrating its ability to adapt to various research paradigms while maintaining theoretical coherence.

5.2. Theory Revision, Expansion & Generation

As seen in Section 4.3, although no modifications to the five core concepts of TIW were identified, TIW itself has seen some development and has been used to inform the development of other concepts, models, and theories. Jaeger et al. (2014)'s introduction of the bridges concept and Burnett (2015)'s own addition of domains are the only substantive theoretical developments of TIW, the latter especially so. Burnett (2015)'s information domains are coded as TC and TG2 as the domains framework absorbs the core TIW into the domain of the social and expands the framework with the introduction of the domains of the individual and signification. As such, the information domains can be seen as an expansion and not a replacement. As seen in the results, TIW is still functional as an independent theory and has continued to be used well after the introduction of the information domains framework. Yu (2012)'s TIW of individuals is positioned as an alternative to TIW, but is based on Popper (1979)'s three worlds rather than Chatman (1999; 2000)'s small worlds; and Yu also admits that research with Burnett and Jaeger's TIW or other social information theories may still better address social aspects of and societal impacts on information behavior, as noted in Section 5.4.

TIW also informed the development of other theories, concepts, and models as seen in articles coded with TG4. Froggatt (2015) identifies different information worlds through interactions of student academic success and access to school library programs. Gibson (2016) uses TIW as one sensitizing theory to develop a model of post-diagnosis information seeking. Gibson and Kaplan (2017) develop a framework of information access in local communities. Similarly, Lee and Butler (2019) consult TIW to develop a theory of local information landscapes. Hovious (2018) combines TIW with activity theory (Engeström, 1987) to redefine transliteracy. Worrall et al. (2021) argue for a social-affective paradigm shift in LIS research, in part informed by TIW. These findings suggest that TIW is useful for informing the development of other theories.

5.3. Strengths: Research Contexts, Methodological Flexibility & Theoretical Amicability

As seen above and in the previous study (Park et al., 2022), TIW has been applied in a variety of research contexts, methods, and alongside other concepts, models, and theories. The scalability of TIW appears useful for exploring and describing multiple-level sociotechnical contexts, enabling its use across various research populations and phenomena ranging from social media users, medical practitioners, immigrants, online gamers, health information systems, to topics such as information access, digital inclusion, and information behavior (Park et al., 2022). Bao and Ke (2023, p. 974), pointing to Burnett and Jaeger's 2011 chapter, explain that TIW is useful for understanding the social dimensions of information use as well as it "endows it with meaning through the multilevel context in which it occurs, so as to deepen the understanding the role of social and structural factors in shaping information behavior in society." Känsäkoski and Huotari (2016, p. 336)'s study was one of the first to apply TIW to a micro-level information world setting of health care professionals and patients, and conclude that TIW "could be applied in wider contexts to increase understanding of the intersection among individuals in small worlds, in the meso worlds in which they work, and in the macro world in which they live." Gibson and Kaplan (2017, p. 138) note that TIW, among other theories, "acknowledge[s] the social multiplicity inherent in physical information places." Welch (2019, p. 246) found TIW useful for understanding the "conceptual, social, technological, and political context of [information behavior]" in their mixed-methods study on LAM professionals in Detroit. Lee and Butler (2019) found that TIW is helpful for understanding issues related to information access and information behaviors. Ndumu (2019; 2020) also found that TIW helps both to describe the information behaviors and to contextualize the information overload and acculturative stress experienced by Black immigrants in the US. Park (2022) noted similar strengths identified through dissertation projects using TIW.

Another area of strength is TIW's methodological flexibility, allowing it to be applied with various research methodologies, whether quantitative, qualitative, or some combination thereof. This is in line with the original intention for TIW to be amethodological (Jaeger & Burnett, 2010). Although TIW was more often used with qualitative methods, as noted in Park (2022) and Park et al. (2022), it has been applied with quantitative methods as well. Froggatt (2015) developed a survey questionnaire regarding students' information behaviors and information access from their contextual small, meso, and life worlds. Jaeger et al. (2014, para. 10) used surveys, site visits, usability and accessibility testing methods to examine user's information needs, access, and behaviors "from the micro (the localized small worlds of information), to the meso (the intermediate levels of information in society), to the macro (the largest level of information flow in society)" while using e-government websites and services provided by public libraries. However, it should be noted that quantitative data collection and analysis methods were mostly used as parts of mixed and or multiple methods to triangulate findings from multiple data sources and provide holistic views of phenomena.

The most frequently used research methods among the sampled articles are qualitative, including data collection and analysis approaches, such as interviews, focus groups, content analyses of documents/texts, field diaries, and observation. Gibson (2016) and Gibson and Kaplan (2017)'s work used TIW to develop interview protocols in their grounded theory studies; Gibson (2016, p. 3) reported that "[TIW] is useful for guiding grounded theory data collection as it provides a framework for describing actor, information behaviors, information values, and boundaries within and between information worlds without imposing specific coding categories." The five core concepts of social types, social norms, information behaviors, information value, and boundaries allow researchers to capture a wide spectrum of normative information behaviors and practices in social worlds, information values influenced by social types and norms, as well as conflicts and dynamics in, out, between, and across the multiple boundaries. TIW can be used as a guiding lens without specifying certain social types, social norms, information behaviors, information values, and boundaries; these are based on the unique data from the persons, communities, and contexts under study.

However, some studies feature a shared codebook to aid data collection and analysis. Hollister (2020) and Lee (2019) used a collaboratively developed codebook to analyze information worlds of Massively Multiplayer Online Role-Playing Game users and political opinion leaders on Twitter from their respective doctoral dissertation projects. Although the codebooks used in Hollister (2020) and Lee (2019) were initially based on inductively chosen codes for five core concepts, new codes were added through collaborative work and discussions with one of the theorists of TIW, Dr. Gary Burnett, and another (then) doctoral student, Julia Skinner, as well as deduced through coding research data across three different dissertation projects, as described in Hollister et al. (2023). Park (2022, p. 438) pointed out that "codebook development process using both inductive and deductive approaches contributed to expanding TIW," and Hollister et al. (2023) maintain that collaborative doctoral education and dissertation work effectively assisted with further development and application of TIW into various research settings.

As seen in Sections 4.1.3 and 4.2.3 and noted in Sections 5.1 and 5.2, TIW exhibits theoretical amicability. Theoretical amicability is exhibited by TIW's versatility to be used alongside other theories or theoretical concepts as well as to inform or guide the development of other concepts, models, or theories. Pang et al. (2020) combine Burnett (2015)'s expanded TIW with activity theory to better understand the information worlds of seniors during natural disasters, identifying significant roles of individual and social factors as well as the influence of space, proximity, and technology. In their conclusion, Pang et al. (2020, p. 629) argue that "[d]rawing on concepts from activity theory with the TIW allowed us to seek out substantial connections between information worlds and purposeful activities."

Blending methodological and theoretical versatility, TIW was also combined with grounded theory approaches. As stated by Gibson and Kaplan (2017, p. 133), "Because this theory [TIW] is not predictive or prescriptive at the level of specific populations, it does not violate Corbin and Strauss (2014)'s or Charmaz (2014)'s suggestion that initial grounded theory should not test the validity of currently existing theoretical frameworks." As noted in Section 2.2, TIW has no official propositions and is not intended to be predictive. Even though Park (2022) noted that the lack of official propositions could be considered a weakness, the omission appears to further support TIW's strengths in methodological flexibility and theoretical amicability. Contrarily, the development of propositions may encourage more TIW use with quantitative methods.

5.4. Weaknesses: Social vs. Individual Information Worlds

As noted in Sections 4.2.2 and 5.2, Yu (2012) offers the most direct challenge to TIW with their introduction of the information worlds of individuals, arguing that Burnett and Jaeger's formulation of information worlds does not adequately account for individual agency and other personal factors. However, Yu (2012) notes that issues related to parsing social and individual information worlds, the formation of information worlds, and understanding how information worlds impact or interact with society remain. Yu (2012, p. 16) acknowledges that continued research and development of Burnett and Jaeger's TIW, or those developed by others, may address such issues. Burnett (2015)'s later integration of the domain of the individual simultaneously addressed prior limitations identified by Yu (2012) as well as potentially anticipated calls from Worrall et al. (2021) to integrate emotional or affective factors into social information behavior theory and research. Further, the domain of signification may account for Jaeger et al. (2014)'s suggested sixth core concept of bridges. Burnett (2015)'s added domain of signification focuses on the signs, languages, tools, and technologies that may or may not enable the exchange of information in a way similar in intent to the concept of bridges, which focuses on the ways in which information may or may not be moved across boundaries. However, broader adoption of the bridges concept is outside the scope of this project and may warrant further research.

The ability of the revised TIW, Burnett (2015)'s information domains framework, to address these concerns also requires further exploration. Within this study's sample, just two research articles applied the information domains framework: Bao and Ke (2023) and Pang et al. (2020). Bao and Ke (2023) employ the concept of transition (Ruthven, 2022) alongside the expanded version of TIW (Burnett, 2015), in part to overcome the perceived social focus of TIW and to better understand how an individual's cognitive, emotional, and physiological factors and activities impact information behavior during different stages of a transition process. In their study, the transitions were focused on the changing information worlds through different phases of lockdown as experienced by depression patients during the COVID-19 pandemic. Bao and Ke (2023, p. 983) note in their conclusion that they "understand the effect of emotion, body, and embodiment in crisis situations on individual meaning construction as a motivating factor for individuals to create a picture of the information worlds and the ongoing impact it is thought to have on individuals' lives." Additionally, Pang et al. (2020) specifically note that the combination of activity theory (Engeström, 1987) and TIW's domain of the individual helped them understand not just the social contexts of behaviors, but also the individual characteristics that could influence individuals' information behaviors "rather than treating the subject as a black box" (Pang et al., 2020, p. 628). With some additional conceptual and theoretical support, the expanded TIW shows some potential in overcoming the limitations stated by Yu (2012), though more work is needed to explore the applications and developments of Burnett (2015)'s expanded TIW framework. TIW's theoretical amicability allows it to work alongside other concepts, models, and theories in cooperative and constructive ways.

Although not in the sample of this or the previous study, Savolainen (2021) briefly discusses TIW in relation to three levels of critique for the purposes of developing theoretical concepts and models of human information behavior. Savolainen (2021) argues for three distinct levels of critique based on depth: the lowest level focusing on gaps or absences in the research literature; the middle level focusing on finding and interpreting limitations of existing theories; and the highest level that addresses identified limitations by proposing alternatives or modifications to strengthen human information behavior theories. Savolainen (2021) positions TIW in the second level, arguing that TIW's development, as described in Burnett and Jaeger (2008; 2011), is based on a critique of the scope of Chatman (1999; 2000)'s small worlds theories due to their inability to address interactions between other small worlds and societal worlds. However, using Savolainen's levels of critique, it seems like TIW should fall into the third level, as Burnett and Jaeger's initial development of TIW builds on the limitations of small worlds through modification of the information value concept and expansion through the addition of the concept of boundaries, and not on critique without proposed solutions. Burnett (2015)'s later development of TIW into the information domains framework also features constructive criticism through the integration of domains to strengthen the theory by addressing weaknesses, such as those identified by Yu (2012). However, Savolainen (2021) did not capture these developments.

5.5. Emergent Themes, Limitations, and Future Work

The primary emergent themes of this study are largely represented by the strengths and weaknesses identified in Sections 5.3 and 5.4. Themes regarding TIW's use across myriad research contexts, methodological flexibility, and theoretical amicability are meaningful, but are not unexpected. In fact, these strengths align with Jaeger and Burnett (2010)'s original intentions for the theory to be useful in these ways. The results of this study provide some evidence to support those intended goals. These features may explain TIW's steady adoption. However, this study is not equipped nor intended to compare TIW with other social information behavior theories or to compare the adoption rate of TIW with related theories. Foundational TIW texts, such as Burnett and Jaeger (2008) and Jaeger and Burnett (2010) provide more information on the development and context of TIW. A larger, more recent, and comprehensive data set is needed to compare TIW's adoption rate with other related theories.

TIW's focus on social contexts at the perceived expense of individual cognitive and affective factors emerged as the primary weakness, with Yu (2012)'s direct challenge as well as the indirect critiques from Jaeger et al. (2014)'s introduction of the bridges concept and Worrall et al. (2021)'s calls for a socio-emotional paradigm shift. Burnett (2015)'s introduction of the information domains framework seems to address these concerns. The domain of the individual can account for an individual's agency as well as cognitive and affective factors. The domain of signification, which focuses on the tools, technologies, symbols, and languages that facilitate the use and exchange of information, seems to account for the concept of bridges, which focuses on how information is exchanged across boundaries. As stated in Section 5.4, although two studies in this sample (Bao & Ke, 2023; Pang et al., 2020) reported useful applications of information domains, further research is needed to explore the use of the information domains framework to assess whether the framework addresses TIW's weaknesses. Similarly, future work could explore the impact or use of the bridges concept as well.

Some uses and interpretations of TIW that emerged during the analysis related to these larger themes are interesting. In particular, and as noted above, Gibson (2016) and Gibson and Kaplan (2017) used TIW in their grounded theory studies because TIW is not prescriptive. This is an interesting finding because early grounded theory work by Glaser and Strauss (1967) and later Glaser (1998), specifically, discourages the use of pre-existing theories for coding and encourages delaying explorations of related research literature until near the end of data analysis (Thistoll et al., 2016; Thornberg, 2012). However, Gibson (2016) and, in later work, Gibson and Kaplan (2017) use a relatively more recent, post-positivist or constructivist form of grounded theory drawing from Charmaz (2014) and Corbin and Strauss (2008). See Thistoll et al. (2016), Thornberg (2012), and Kelle (2007) for more on the development and differentiation of grounded theory. As noted in Section 2.2 and discussed in Section 5.3, TIW's lack of propositions is an apparent strength that permits TIW to be used with various methods and alongside other theories.

Even though TIW has been used to study a variety of research populations, few studies seem to focus on libraries, librarians, library educators, or other information professionals. Froggatt (2015)'s work demonstrates how TIW can be helpful for understanding how IU students benefit from access to school librarian programs. Jaeger et al. (2014) argue that the concepts of TIW alongside the introduced concept of bridges may help librarians and libraries to better understand and perform their intermediary roles in supporting access and use of e-government resources and the challenges faced by their library users in terms of information value, boundaries, and bridges. One of the studies included in this sample, Hollister (2020), is based on a portion of his dissertation research (Hollister, 2016), which explores the interactions between the core concepts of TIW and digital literacies. Hollister (2016) argues that social types, social norms, and information values can impact an individual's information behaviors and digital literacy practices, such as which sources of information are considered important and how to appropriately use information within a community based on its social norms and information values. Considering practical implications, Hollister's findings suggest that librarians and library educators should consider social types, social norms, and information value systems when engaging in information and digital literacy instruction and developing curricula. The social norms and information values of an information world may enable or hinder the effectiveness of its members' information or digital literacy skills or acceptance of information from outside sources.

Building on those implications, a stronger understanding of a community's social types, social norms, information behaviors, information value systems, and boundaries could help libraries and librarians be more effective with community outreach, programming, and instruction. TIW may also be a useful tool for librarians and libraries to assess and meet the needs of diverse communities, creating more inclusive and representative collections, services, and staffing. Future work on TIW should explore the potential applications in professional practice in libraries, archives, schools, universities, and other settings.

The design of this study introduces limitations as well. The small sample size could be augmented by including more languages beyond English and Korean, including more paywalled academic databases not currently available to the research team, and exploring TIW use in other publication formats, such as conference proceedings, books and book chapters, dissertations and theses, and grey literature. Interviews or focus groups conducted with scholars who have used or are using TIW might provide additional insight and depth not captured in the articles.

Given the findings, themes, and limitations of this study, the researchers propose several areas of future research and work:

(1) Continued application, development, and testing of TIW across new research contexts and more methods. Although TIW is amethodological, it is primarily used with qualitative methods. More quantitative and mixed methods studies using TIW may aid in adoption and development.

(2) Broader explorations and comparisons of TIW with other LIS theories. The findings of this study suggest that TIW plays well with a variety of methods and with other theories, and can be used to inform new concepts and theories, yet awareness and use of TIW seem to be comparatively low. The relationship of TIW to other theories and its overall impact on LIS scholarship could be explored and evaluated in the future. Relatedly, future systematic reviews of TIW could expand their scope to include more theories, more articles in other databases, other publication types, additional languages, etc. as well as conducting bibliometric or scientometric studies to examine comparative use and citation rates, research trends, interdisciplinarity, and the evolution of theory use in LIS more broadly.

(3) Exploration and description of the application and development of TIW, its core concepts, and expansions. Burnett (2015)'s information domains framework and Jaeger et al. (2014)'s concept of bridges seem to be the primary theoretical developments of TIW. Future work could explore and assess the use of these revisions or expansions to TIW and whether they improve upon TIW's strengths, and address its weaknesses.

(4) Application of TIW to inform professional LIS practice and education. Given TIW's strengths, it may be a useful tool in professional LIS and educational settings, such as libraries, archives, and other information organizations. TIW may aid libraries, librarians, and other information professionals in building and providing inclusive, equitable, and socially just collections, services, instruction, outreach, policies, and staffing that are responsive to and representative of the information worlds of the diverse communities they serve. More work is needed to explore the practical applications of TIW, to evaluate its use as a tool for working professionals as well as its usefulness in preparing future librarians, archivists, and other information professionals.

6. CONCLUSION

Through directed qualitative content analysis, this study expands on the findings from a systematic review (Park et al., 2022) on the use of TIW in peer-reviewed research articles with qualitative exploration of the ways TIW has been applied, developed, and discussed. The findings of this study provide further context into how and why TIW is being used. TIW's strengths in its abilities to be useful for understanding and interpreting complex, multi-level sociotechnical research contexts, its (a) methodological flexibility, and its robustness when used as a primary theoretical framework, as well as its theoretical amicability to mesh with other concepts, models, and theories have all encouraged and supported its continued use and development. These core findings provide support for and align with the original intentions of TIW as specified by Burnett and Jaeger (2008) (Jaeger & Burnett, 2010). TIW is also shown to be useful for advancing the development of new context-aware concepts and theories. Despite criticisms of TIW's emphasis on social aspects over individual factors (Yu, 2012), Burnett (2015)'s integration of the domain of individuals may address this concern as well as more directly acknowledge or address cognitive and emotional factors, as promoted by Worrall et al. (2021). The domain of signification may also account for Jaeger et al. (2014)'s proposed sixth core concept of bridges.

Exploration and evaluation of the continued use and development of TIW is warranted in future research. Introduced originally in 2008, TIW is a relatively young theory and its use is likely to grow over time. Circling back to review its use and growth in another decade or two may reveal new insights. Future work should also consider the adoption and application of Burnett (2015)'s information domains framework. Regardless of its sole focus on TIW, this study also demonstrates that theory use and development exist within LIS. In turn, the research team echoes calls from other scholars to promote theory application and development in LIS. The research team hopes this study inspires interest and adoption of the TIW in the distinct, yet intrinsically connected information worlds of LIS research and professional practice, as well as across boundaries into other disciplines.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

This research was financially supported by Hansung University.

REFERENCES

Bao, X., & Ke, P. (2023). Chaos, expansion, and contraction: The information worlds of depression patients during the COVID-19 pandemic lockdown. *Journal of the Association for Information Science and Technology*, 74(8), 971-989. https://doi.org/10.1002/asi.24769

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, 13(5), 407-424. https://doi.org/10.1108/eb024320

Bates, M. J. (2009). Information. In M. J. Bates & M. N. Maack (Eds.), *Encyclopedia of library and information sciences* (3rd ed.) (pp. 14). CRC Press.

Bawden, D., & Robinson, L. (2009). The dark side of information: Overload, anxiety and other paradoxes and pathologies. *Journal of Information Science*, 35(2), 180-191. https://doi.org/10.1177/0165551508095781

Bawden, D., & Robinson, L. (2013). *Introduction to information science*. Neal-Schuman.

Bowker, G. C., & Star, S. L. (1999). *Sorting things out: Classification and its consequences*. MIT Press.

Brattland, K. (2017). Information preferences of Reddit communities surrounding the Brock Turner case. *Progressive Librarian*, 2017/18(46), 86-100.

Brookes, B. C. (1980). The foundations of information science. Part I. Philosophical aspects. *Journal of Information Science*, 2(3-4), 125-133. https://doi.org/10.1177/016555158000200302

Buckland, M. K. (1991). *Information and information systems*. Greenwood Press.

Burnett, G. (2015). Information worlds and interpretive practices: Toward an integration of domains. *Journal of Information Science Theory and Practice*, 3(3), 6-16. https://doi.org/10.1633/JISTaP.2015.3.3.1

Burnett, G., & Jaeger, P. T. (2008). Small worlds, lifeworlds, and information: The ramifications of the information behaviour of social groups in public policy and the public sphere. *Information Research*, 13(2), 346.

Burnett, G., & Jaeger, P. T. (2011). The theory of information worlds and information behaviour. In A. Spink & J. Heinström (Eds.), *New directions in information behaviour* (pp. 161-180). Emerald Group Publishing Limited. https://doi.org/10.1108/S1876-0562(2011)002011a010

Burnett, G., Whetstone, M., & Jaeger, P. T. (2013). Personal health record interfaces: A hermeneutic analysis. *First Monday*, 18(8). https://doi.org/10.5210/fm.v18i8.4748

Burnett, K., Subramaniam, M. M., & Gibson, A. (2009). Latinas cross the IT border: Understanding gender as a boundary object between information worlds. *First Monday*, 14(9). https://doi.org/10.5210/fm.v14i9.2581

Cascio, M. A., Lee, E., Vaudrin, N., & Freedman, D. A. (2019). A team-based approach to open coding: Considerations for creating intercoder consensus. *Field Methods*, 31(2), 116-130. https://doi.org/10.1177/1525822x19838237

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). SAGE Publications.

Chatman, E. A. (1991). Life in a small world: Applicability of gratification theory to information-seeking behavior. *Journal of the American Society for Information Science*, 42(6), 438-449. https://doi.org/10.1002/(SICI)1097-4571(199107)42:6<438::AID-ASI6>3.0.CO;2-B

Chatman, E. A. (1999). A theory of life in the round. *Journal of the American Society for Information Science*, 50(3), 207-217. https://doi.org/10.1002/(SICI)1097-4571(1999)50:3<207::AID-ASI3>3.0.CO;2-8

Chatman, E. A. (2000). Framing social life in theory and research. *New Review of Information Behaviour Research*, 1, 3-17.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). SAGE Publications. https://doi.org/10.4135/9781452230153

Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. SAGE Publications.

Currie, G., & Suhomlinova, O. (2006). The impact of institutional forces upon knowledge sharing in the UK NHS: The triumph of professional power and the inconsistency of policy. *Public Administration*, 84(1), 1-30. https://doi.org/10.1111/j.0033-3298.2006.00491.x

Davis, F. D. (1987). *User acceptance of information systems : The technology acceptance model (TAM)*. https://hdl.handle.net/2027.42/35547

Dervin, B. (1992). From the mind's eye of the user: The sense-making qualitative-quantitative methodology. In J. D. Glazier & R. R. Powell (Eds.), *Qualitative research in information management* (pp. 61-84). Libraries Unlimited.

Dunne, J. E. (2002). Information seeking and use by battered women: A "person-in-progressive-situations" approach. *Library & Information Science Research*, 24(4), 343-355. https://doi.org/10.1016/S0740-8188(02)00132-9

Elder, G. H., Johnson, M. K., & Crosnoe, R. (2003). The emergence and development of life course theory. In J. T. Mortimer & M. J. Shanahan (Eds.), *Handbook of the life course* (pp. 3-19). Springer US. https://doi.org/10.1007/978-0-306-48247-2_1

Engeström, Y. (1987). *Learning by expanding: An activity theoretical approach to developmental research*. Orienta-Konsultit.

Fisher, K. E., Erdelez, S., & McKechnie, L. (2005). *Theories of information behavior*. American Society for Information Science and Technology.

Frankfort-Nachmias, C., & Nachmias, D. (1996). *Research methods in the social sciences*. Arnold.

Froggatt, D. L. (2014). *Theory of the informationally underserved: A pragmatic model for social justice* (doctoral dissertation). Simmons College, Boston, MA, USA.

Froggatt, D. L. (2015). The informationally underserved: Not always diverse, but always a social justice advocacy model. *School Libraries Worldwide*, 21(1), 54-72. https://doi.org/10.29173/slw6883

Gibson, A. N. (2016). Building a progressive-situational model of post-diagnosis information seeking for parents of individuals with Down syndrome. *Global Qualitative Nursing Research*, 3, 2333393616680967. https://doi.org/10.1177/2333393616680967

Gibson, A. N., & Kaplan, S. (2017). Place, community and information behavior: Spatially oriented information seeking zones and information source preferences. *Library & Information Science Research*, 39(2), 131-139. https://doi.org/10.1016/j.lisr.2017.03.001

Glaser, B. G. (1998). *Doing grounded theory: Issues and discussions*. Sociology Press.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

Habermas, J. (1992). Further reflections on the public sphere. In C. Calhoun (Ed.), *Habermas and the public sphere* (pp. 421-461). MIT Press.

Harding, K. (2015). *Asking for it : The alarming rise of rape culture--and what we can do about it.* Da Capo Lifelong.

Hollister, J. M. (2016). *In- and out-of-character: The digital literacy practices and emergent information worlds of active role-players in a new massively multiplayer online role-playing game* (doctoral dissertation). Florida State University, Tallahassee, FL, USA.

Hollister, J. M. (2020). The information worlds of online role-players. *Journal of Korean Biblia Society for Library and Information Science*, 31(2), 223-266. https://doi.org/10.14699/kbiblia.2020.31.2.223

Hollister, J. M., Burnett, G., Lee, J., & Skinner, J. (2023). Collaborative codebook development across information worlds: Intersections between pedagogy, theory, and method in three dissertations. *Journal of Education for Library and Information Science*, 64(1), 18-36. https://doi.org/10.3138/jelis-2020-0095

Houser, L. (1988). A conceptual analysis of information science. *Library and Information Science Research*, 10(1), 3-34.

Hovious, A. (2018). Toward a socio-contextual understanding of transliteracy. *Reference Services Review*, 46(2), 178-188. https://doi.org/10.1108/rsr-02-2018-0016

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Huotari, M. L., & Chatman, E. (2001). Using everyday life information seeking to explain organizational behavior. *Library & Information Science Research*, 23(4), 351-366. https://doi.org/10.1016/S0740-8188(01)00093-7

Jaeger, P. T., & Burnett, G. (2010). *Information worlds: Behavior, technology, and social context in the age of the Internet*. Routledge. https://doi.org/10.4324/9780203851630

Jaeger, P. T., Gorham, U., Bertot, J. C., Taylor, N. G., Larson, E., Lincoln, R., Lazar, J., & Wentz, B. (2014). Connecting government, libraries and communities: Information behavior theory and information intermediaries in the design of LibEGov.org. *First Monday*, 19(11). https://doi.org/10.5210/fm.v19i11.4900

Jeong, D. Y., & Kim, S. J. (2005). Knowledge structure of library and information science in South Korea. *Library & Information Science Research*, 27(1), 51-72. https://doi.org/10.1016/j.lisr.2004.09.004

Jones, J. M. (1997). *Prejudice and racism* (2nd ed.). McGraw-Hill.

Julien, H., Pecoskie, J., & Reed, K. (2011). Trends in information behavior research, 1999–2008: A content analysis. *Library & Information Science Research*, 33(1), 19-24. https://doi.org/10.1016/j.lisr.2010.07.014

Känsäkoski, H. (2014). *Value creation in childhood obesity care and prevention.* University of Oulu.

Känsäkoski, H., & Huotari, M. L. (2016). Applying the theory of information worlds within a health care practise in Finland. *Journal of Documentation*, 72(2), 321-341. https://doi.org/10.1108/jd-05-2015-0065

Kassim, M. (2024). Application of theories in library and information science research in Tanzania: A content analysis. *Alexandria*, 34(1-2), 14-27. https://doi.org/10.1177/09557490241250297

Kelle, U. (2007). "Emergence" vs. "Forcing" of empirical data? A crucial problem of "grounded theory" reconsidered. *Historical Social Research, Supplement,* (19), 133-156.

Kim, S. J., & Jeong, D. Y. (2006). An analysis of the development and use of theory in library and information science research articles. *Library & Information Science Research*, 28(4), 548-562. https://doi.org/10.1016/j.lisr.2006.03.018

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42(5), 361-371. https://doi.org/(SICI)1097-4571(199106)42:5<361::AID-ASI6>3.0.CO;2-#

Kumasi, K. D., Charbonneau, D. H., & Walster, D. (2013). Theory talk in the library science scholarly literature: An exploratory analysis. *Library & Information Science Research*, 35(3), 175-180. https://doi.org/10.1016/j.lisr.2013.02.004

Lee, J. (2019). An exploratory study on the political information behaviors of Korean opinion leaders on Twitter: Through the lens of theory of information worlds. *Journal of the Korean Society for Library and Information Science*, 53(1), 83-108. https://doi.org/10.4275/KSLIS.2019.53.1.083

Lee, M., & Butler, B. S. (2019). How are information deserts created? A theory of local information landscapes. *Journal of the Association for Information Science and Technology*, 70(2), 101-116. https://doi.org/10.1002/asi.24114

Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, 28(4), 587-604. https://doi.org/10.1111/j.1468-2958.2002.tb00826.x

Lund, B. (2021). The structure of information behavior dissertations 2009–2018: Theories, methods, populations, disciplines. *Journal of Librarianship and Information Science*, 53(2), 225-232. https://doi.org/10.1177/0961000620935499

Lund, B. D. (2019). The citation impact of information behavior theories in scholarly literature. *Library & Information Science Research*, 41(4), 100981. https://doi.org/10.1016/j.lisr.2019.100981

McKechnie, L. E. F., Pettigrew, K. E., & Joyce, S. L. (2001). The origins and contextual use of theory in human information behaviour research. *New Review of Information Behaviour Research*, 2(November), 47–63.

Ndumbaro, F., & Ochieng, L. M. (2021). Access to information on family planning (FP) methods among married women of reproductive age in Ilala district, Dar es Salaam Tanzania. *University of Dar es Salaam Library Journal*, 16(1), 98-114. https://doi.org/10.4314/udslj.v16i1

Ndumu, A. (2019). Linkages between information overload and acculturative stress: The case of Black diasporic immigrants in the US. *Journal of Librarianship and Information Science*, 0961000619857115. https://doi.org/10.1177/0961000619857115

Ndumu, A. (2020). Toward a new understanding of immigrant information behavior: A survey study on information access and information overload among US Black diasporic immigrants. *Journal of Documentation*, 76(4), 869-891. https://doi.org/10.1108/jd-04-2019-0066

Nolin, J., & Åström, F. (2010). Turning weakness into strength: Strategies for future LIS. *Journal of Documentation*, 66(1), 7-27. https://doi.org/10.1108/00220411011016344

Oh, C. Y., & Butler, B. (2019). Small worlds in a distant land: International newcomer students' local information behaviors in unfamiliar environments. *Journal of the Association for Information Science and Technology*, 70(10), 1060-1073. https://doi.org/10.1002/asi.24185

O'Leary, Z. (2005). *Researching real-world problems: A guide to methods of inquiry*. SAGE Publications.

Pang, N., Karanasios, S., & Anwar, M. (2020). Exploring the information worlds of older persons during disasters. *Journal of the Association for Information Science and Technology*, 71(6), 619-631. https://doi.org/10.1002/asi.24294

Park, S. (2022). A reflection on social perspective of information behavior: Around the theory of information worlds. *Journal of the Korean Society for Library and Information Science*, 56(1), 425-447. https://doi.org/10.4275/KSLIS.2022.56.1.425

Park, S., Lee, J., & Hollister, J. M. (2022). A systematic review on the application of the theory of information worlds. *Journal of Information Science Theory and Practice*, 10(4), 87-109. https://doi.org/10.1633/JISTaP.2022.10.4.7

Pettigrew, K. E., & McKechnie, L. (2001). The use of theory in information science research. *Journal of the American Society for Information Science and Technology*, 52(1), 62-73. https://doi.org/10.1002/1532-2890(2000)52:1<62::AID-ASI1061>3.0.CO;2-J

Pickard, A. J., & Childs, S. (2013). *Research methods in information* (2nd ed.). Neal-Schuman.

Popper, K. R. (1979). *Objective knowledge: An evolutionary approach*. Clarendon Press.

Raber, D. (2003). *The problem of information: An introduction to information science*. Bloomsbury Academic.

Riley, F., Allen, D. K., & Wilson, T. D. (2022). When politicians and the experts collide: Organization and the creation of information spheres. *Journal of the Association for Information Science and Technology*, 73(8), 1127-1139. https://doi.org/10.1002/asi.24618

Rioux, K. (2010). Metatheory in library and information science: A nascent social justice approach. *Journal of Education for Library and Information Science*, 51(1), 9-17.

Rohman, A., & Pitaloka, D. (2025). Disconnected and disabled during the pandemic: Toward more inclusive pandemic response plans in the Global South. *Journal of Librarianship and Information Science*, 57(1), 211-222. https://doi.org/10.1177/09610006231207657

Rudmin, F. (2009). Constructs, measurements and models of acculturation and acculturative stress. *International Journal of Intercultural Relations*, 33(2), 106-123. https://doi.org/10.1016/j.ijintrel.2008.12.001

Ruthven, I. (2022). An information behavior theory of transitions. *Journal of the Association for Information Science and Technology*, 73(4), 579-593. https://doi.org/10.1002/asi.24588

Savolainen, R. (2008). *Everyday information practices: A social phenomenological perspective*. Bloomsbury Academic.

Savolainen, R. (2021). Levels of critique in models and concepts of human information behaviour research. *Aslib Journal of Information Management*, 73(5), 772-791. https://doi.org/10.1108/ajim-01-2021-0028

Sonnenwald, D. H. (1999). Evolving perspectives of human information behavior: Contexts, situations, social networks and information horizons. In T. D. Wilson & D. K. Allen (Eds.), *Exploring the contexts of information behavior: Proceedings of the second international conference in information needs* (pp. 176-190). Taylor Graham.

Star, S. L. (1989). The structure of Ill-structured solutions: Boundary objects and heterogeneous distributed problem solving. In L. Gasser & M. N. Huhns (Eds.), *Distributed artificial intelligence* (pp. 37-54). Morgan Kaufmann. https://doi.org/10.1016/B978-1-55860-092-8.50006-X

Star, S. L., & Griesemer, J. R. (1989). Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19(3), 387-420. https://doi.org/10.1177/030631289019003001

Strauss, A. L. (1987). *Qualitative analysis for social scientists. Cambridge University Press*. https://doi.org/10.1017/CBO9780511557842

Thistoll, T., Hooper, V., & Pauleen, D. J. (2016). Acquiring and developing theoretical sensitivity through undertaking a grounded preliminary literature review. *Quality & Quantity*, 50(2), 619-636. https://doi.org/10.1007/s11135-015-0167-3

Thornberg, R. (2012). Informed grounded theory. *Scandinavian Journal of Educational Research*, 56(3), 243-259. https://doi.org/10.1080/00313831.2011.581686

Ukwoma, S. C., & Ngulube, P. (2021). To borrow or not to borrow is the question? Theory borrowing in library information science postgraduate research in Nigeria and South Africa. *The International Information & Library Review*, 53(1), 48-62. https://doi.org/10.1080/10572317.2020.1790261

Van Maanen, J. (1998). Different strokes: Qualitative research in the *Administrative Science Quarterly* from 1956 to 1996. In J. Van Maanen (Ed.), *Qualitative studies of organizations* (pp. ix-xxxii). SAGE Publications.

Vakkari, P. (2008). Trends and approaches in information behaviour research. *Information Research*, 13(4), paper 361. http://InformationR.net/ir/13-4/paper361.html

Vakkari, P., Chang, Y. W., & Järvelin, K. (2022). Disciplinary contributions to research topics and methodology in library and information science—Leading to fragmentation? *Journal of the Association for Information Science and Technology*, 73(12), 1706-1722. https://doi.org/10.1002/asi.24690

VanScoy, A., Julien, H., Buckley, A., & Goodell, J. (2022). Theory usage in empirical research in ISIC conference papers (1996-2020). *Information Research*, 27(Special issue), isic2227. https://doi.org/10.47989/irisic2227

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478. https://doi.org/10.2307/30036540

Welch, C. (2019). Circling the reality of public institutions: An e-Delphi study of information, culture, and community stakeholders voicing the way forward in Detroit, Michigan. *Libri*, 69(4), 241-261. https://doi.org/doi:10.1515/libri-2019-0108

Wenger, E. (1998). *Communities of practice: Learning, meaning, and identity*. Cambridge University Press. https://doi.org/10.1017/CBO9780511803932

Whitley, R. (1984). *The intellectual and social organization of the sciences*. Clarendon Press.

Williams, M. D., Rana, N. P., & Dwivedi, Y. K. (2015). The unified theory of acceptance and use of technology (UTAUT): A literature review. *Journal of Enterprise Information Management*, 28(3), 443-488. https://doi.org/10.1108/jeim-09-2014-0088

Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation*, 55(3), 249-270. https://doi.org/10.1108/eum0000000007145

Worrall, A. (2019). "Connections above and beyond": Information, translation, and community boundaries in LibraryThing and Goodreads. *Journal of the Association for Information Science and Technology*, 70(7), 742-753. https://doi.org/10.1002/asi.24153

Worrall, A., Cappello, A., & Osolen, R. (2021). The importance of socio-emotional considerations in online communities, social informatics, and information science. *Journal of the Association for Information Science and Technology*, 72(10), 1247-1260. https://doi.org/10.1002/asi.24489

Wu, M. M., Liang, H. S., Hsieh, H. H., & Lee, V. (2017). *Theory use in library and information science journals in Taiwan 2010 – 2015*. iConference 2017: Effect • Expand • Evolve: Global Collaboration across the Information Community,

Wuhan, China.
Xie, J., He, Z., Burnett, G., & Cheng, Y. (2021). How do mothers exchange parenting-related information in online communities? A meta-synthesis. *Computers in Human Behavior*, 115, 106631. https://doi.org/10.1016/j.chb.2020.106631
Yu, L. (2012). Towards a reconceptualization of the 'information worlds of individuals'. *Journal of Librarianship and Information Science*, 44(1), 3-18. https://doi.org/10.1177/0961000611424586

JISTaP https://www.jistap.org
Journal of Information Science Theory and Practice

RESEARCH PAPER
J Inf Sci Theory Pract 13(4): 46-57, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.4.3

Beyond Authorship: Exploring the Use of Contributor Roles Taxonomies in LIS Journals

Eun-Ja Shin*

Department of Media & Communication, Sejong University, Seoul, Korea
Institute of Platform-Media, Sejong University, Seoul, Korea
E-mail: ejshin@sejong.ac.kr

ABSTRACT

Traditional author attribution often inadequately represents researchers' varied contributions, potentially causing dissatisfaction among collaborators. Contributor Role Ontologies and Taxonomies (CROTs) address this by specifying researcher roles and offering a clearer framework for recognizing diverse contributions, including innovative methodologies. As Library and Information Science (LIS) research becomes more complex and interdisciplinary with diverse participant roles, clearly delineating these roles in LIS journals is crucial. To explore the current landscape of the LIS field, this study collected and examined 82 journals indexed in the 2023 Journal Citation Reports. The analysis of these journals revealed that only nine of them have actively adopted CROTs, within which a total of 749 instances of reported contributor roles were analyzed. While the use of CROTs in collaborative papers within these nine journals varied, not all collaborative research followed this practice, with the lowest adoption rate observed being 46%. Moreover, while most CROTs offered a precise and comprehensive presentation of participants' roles, several listed roles more generically, implying equal contributions from all authors. Furthermore, CRediT, a widely used CROTs model, may not fully capture the specific nuances of LIS research, especially in areas like literature reviews. To encourage broader and more equitable adoption, a tailored model that reflects the characteristics of LIS research and seamlessly integrates with online submission systems is essential. These efforts will ultimately foster fairer recognition of contributions and strengthen the collaborative research and publishing culture within the LIS community.

Keywords: authorship, contributor roles, contributorship, Contributor Role Ontologies and Taxonomies, CRediT, collaborative research

Received: May 18, 2025
Accepted: August 14, 2025
Revised: July 23, 2025
Published: December 30, 2025

***Corresponding Author:** Eun-Ja Shin
https://orcid.org/0000-0002-0607-9587
E-mail: ejshin@sejong.ac.kr

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee or authors' permission. Open Access articles are automatically archived in the Korea Institute of Science and Technology Information (KISTI)'s Open Access repository (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

© 2025 [Eun-Ja Shin] This article is distributed under the Creative Commons Attribution License (CC BY), allowing unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. It is published by the Korea Institute of Science and Technology Information (KISTI).

1. INTRODUCTION

1.1. The Background of Contributor Role Ontologies and Taxonomies' Emergence

Collaborative research has increased significantly over the past few decades, though the extent of this trend varies across academic fields. Prior studies suggest that research conducted collaboratively often yields productivity advantages compared to solo research (Jokić, 2020; Larivière et al., 2015; Lee & Bozeman, 2005; Mwantimwa & Kassim, 2023). This growth in collaboration can be largely attributed to the increasing complexity of research questions, which frequently necessitates expertise from multiple disciplines. Furthermore, collaboration among researchers from diverse backgrounds is particularly valuable as it fosters the generation of novel ideas and creative solutions (Allen et al., 2014).

However, collaboration introduces challenges less common in solo research. A significant issue is the inflation in the number of authors, which can lead to dissatisfaction over authorship attribution and instances of improper contributorship (Brand et al., 2015; Holcombe et al., 2021; Jabbehdari & Walsh, 2017; Paul-Hus et al., 2017). For example, in biomedicine, ghost authorship—where a researcher contributes but is not credited—accounts for approximately 20% of cases (Wislar et al., 2011). Similarly, honorary authorship, including guest and gift authorship, is estimated to occur in 20-40% of cases (Hardjosantoso et al., 2020; Jabbehdari & Walsh, 2017; Mowatt et al., 2002). Both practices undermine the integrity of authorship and pose risks to researchers' morale and the academic publishing ecosystem.

To address these challenges, adopting Contributor Role Ontologies and Taxonomies (CROTs) offers a potential solution. CROTs are a framework for clearly classifying and defining the roles of each contributor in research, and CRediT is the most widely used CROT. CROTs can explicitly specify each researcher's role in a publication, thereby ensuring transparency and accountability. Including contributorship statements in academic papers clarifies contributions for readers and stakeholders alike, providing a clear account of contributors' responsibilities (Alpi & Akers, 2021; Ding et al., 2021; Hosseini et al., 2024; Teixeira da Silva et al., 2023). This transparency benefits journal editors, who can easily identify individual contributions, and funders, who can assess whether researchers fulfilled their assigned roles. University administrations could also use CROTs to evaluate faculty members' contributions for hiring, promotion, and tenure decisions. Furthermore, CROTs assist project managers in identifying research personnel for specific tasks (Hosseini et al., 2023).

Despite these advantages, consensus on the necessity of CROTs has not yet been reached across all academic fields. Disciplines such as biomedicine, life sciences, and medicine have actively embraced CROTs, driven by the prevalence of large research teams and the associated challenges of contributorship (Allen et al., 2014; Hosseini et al., 2024; Larivière et al., 2021). Discussions about CROTs began in these fields as early as the late 1990s, gaining traction when the International Committee of Medical Journal Editors (ICMJE) recommended their inclusion in academic publications (Larivière et al., 2021; Vasilevsky et al., 2021). In 2012, the Wellcome Trust spearheaded efforts to standardize CROTs, which culminated in the development of the CRediT system, specifying 14 distinct roles. By 2014, CRediT was finalized, and its adoption by several medical journals in 2015 marked a turning point. Major publishers, including Cell, Public Library of Science (PLOS), Springer, and BMJ, soon joined this initiative, leading to the adoption of CRediT by over 120 journals. The movement gained further momentum in 2019 when Elsevier began strongly advocating for its implementation (Reller, 2019).

1.2. Need for CROTs in LIS Research

Library and information science (LIS) is undergoing a significant transformation. While it may not be dominated by massive research teams like the medical field, LIS research is increasingly collaborative and interdisciplinary, naturally leading to a growing interest in how contributions are recognized. LIS uniquely draws theories and methods from a wide array of fields, including sociology, psychology, business administration, and computer science (Petras, 2023; Vakkari et al., 2023). This interdisciplinary nature is further emphasized by reports indicating that a substantial two-thirds of LIS master's and doctoral theses originate from non-LIS departments (Prebor, 2010). Moreover, the field, particularly information science, now deeply integrates IT domains such as big data, artificial intelligence, and digital preservation (Chang, 2018; 2021; Petras, 2023). As LIS research grows more complex, experts from diverse backgrounds contribute, resulting in more layered and varied roles for participants. Fitzgerald et al. (2020) argue that for collaborative research, it is generally considered standard practice and aligns with research ethics to clearly define each researcher's individual contributions, and it is based on this principle that *College & Research Libraries* intends to adopt such a policy.

While specific reports on the prevalence of ghost or honorary authorship in LIS are not yet available, these evolving shifts strongly suggest the field is vulnerable to such problems. Fortunately, some valuable early studies are already urging LIS to develop its own suitable model for handling contributions (Fitzgerald et al., 2020; Herbert & Kaspar, 2019). Given these challenges, implementing CROTs is essential for ensuring transparency and accountability in LIS research publications.

This study, therefore, explores how current LIS journals record individual contributions. Specifically, the CROTs practices of journals within the Journal Citation Reports' Information Science & Library Science categories were examined. The focus was on those typically indexed in the Social Science Citation Index (SSCI), which are referred to as SSCI LIS journals in this study. This selection reflects their high visibility within the academic community and strong author publication preferences (Lee et al., 2020; Nicholas et al., 2022; Rowley et al., 2022).

1.3. Research Questions

This study examined the CROTs practices of prominent LIS journals, typically SSCI-indexed, to explore how they document individual contributions, leading to the following research questions:

- Which, and what are their publishers?
- How do these journals label their CROTs sections?
- What is the CROTs specification rate in LIS journal articles adhering to a CROTs policy?
- Are there notable differences in role distribution across journals?
- Is the application of CRediT, a leading CROTs model, feasible for LIS publishing?

In the LIS field, where collaborative and multidisciplinary research is increasingly common, accurately crediting individual contributions is essential. This goes beyond merely listing author names; instead, specific roles and contributions should be clearly defined (Fitzgerald et al., 2020; Herbert & Kaspar, 2019). Recognizing the vital importance of documenting CROTs in LIS journal publications, this study offers concrete data collection and analysis to serve as a foundational step. Continued research is expected to contribute to the formation of a gradual consensus on this critical aspect of scholarly communication. Ultimately, this study seeks to promote a publishing culture within the LIS community that accurately specifies individual roles and contributions, thereby ensuring fair credit allocation for all collaborators.

2. LITERATURE REVIEW

As research competition intensifies, academic interest in specifying CROTs to clarify contributions in collaborative projects and ensure fair credit allocation continues to grow. Most prior studies on CROTs emphasize their importance, evolution, and current state, alongside recommendations for adapting them to specific academic fields.

Allen et al. (2014) noted the frequent difficulty in discerning the precise contributions of individual researchers from the author lists, acknowledgments, and contribution sections currently provided by many academic journals. Highlighting the importance of clarity in participant roles, especially in large-scale studies, they suggested that adopting CROTs could address this issue. By explicitly adopting CROTs, researchers can highlight their specialized expertise and relevant skills, thus increasing opportunities such as grant applications and academic appointments. Additionally, CROTs may provide a platform for junior researchers to emerge as key authors and for methodological innovators—who might otherwise go unnoticed in traditional author lists—to gain recognition. Funding agencies could benefit from clearer insights into each team member's specific role, while universities and research institutions could leverage CROTs for innovations in academic appointments and promotions. As a result, Allen et al. (2014) underscored the necessity of implementing CROTs. They also emphasized the potential value of implementing such systems, proposing the addition of a feature to online submission systems allowing authors to specify their detailed roles.

Furthermore, Ware and Mabe (2015) stressed the need for standardizing CROTs. In most fields, papers typically involve multiple authors with varied roles, such as study design, data analysis, or writing. While these roles are sometimes detailed in acknowledgments, especially in medical journals, descriptions are often inconsistent and unstructured. To address this, editors, journals, and publishers are working on developing a standardized taxonomy for contributor roles in academic journals. The evolution of CROTs is discussed in Allen et al. (2019), who found that the dissemination of CROTs largely depends on how well editors, journals, and publishers understand their purpose and take steps to implement them. For instance, the PLOS played a pivotal role by adopting the 14-role CRediT structure. Similarly, Cell Press encourages authors to specify their contributions, while *Learned Pub-*

lishing has advocated for documenting contributors' role allocations since 2015. The implementation of CRediT by Elsevier in 2019 further accelerated the use of CROTs, marking a significant step in its evolution (Reller, 2019).

In addition to publishers, online submission systems have played a critical role in the dissemination of CROTs. Since 2015, select journals using Editorial Manager have asked authors to specify contributions via CRediT. Committees involved in academic publishing have also facilitated the spread of CROTs. According to Larivière et al. (2021), the widespread adoption of CROTs can be attributed to efforts by the ICMJE and the Committee on Publication Ethics, which have encouraged journal editors to implement these standards. As a result, medical journals began implementing CRediT in 2015, marking a turning point in the broader adoption of CROTs.

Analyzing PLOS journal data from 2017 to 2018, Larivière et al. (2021) investigated how research contributions are distributed within teams. Their findings revealed that while fundamental tasks such as conceptualization, methodology, and writing are prevalent across research, disciplinary variations highlight the significance of more specialized roles, including visualization and software in specific fields, thereby suggesting a need for more diverse evaluation metrics that account for these differences. Furthermore, their analysis underscored the prevalent yet under-supported role of data curation, which is crucial for the advancement of open science initiatives. The study also indicated that women are more frequently involved in data curation and other technical tasks. This disparity points to potential gender inequities within academic labor and highlights the necessity for transparent contribution systems.

Drawing on a dataset of over 80,000 *PLOS ONE* articles (2018-2023), Maddi and Teixeira da Silva (2024) collected and analyzed author roles, DOIs, affiliations, and funding acknowledgments. Their investigation revealed that over 9% of publications exhibited at least one instance of inappropriate authorship, affecting approximately 2.5% of the authors. Furthermore, their analysis indicated a higher likelihood of potential monetary authorship among established researchers and those affiliated with companies or nonprofits. While acknowledging the reliance on self-reported author roles and acknowledgments, their findings underscore the necessity for stronger journal oversight and more conscientious author practices to ensure the integrity of academic publishing. They not only outlined authorship eligibility for the 14 CRediT contributors but also identified potential contributorships, such as article processing charge (APC)-paying, clarifying that this could apply even when the individual's sole contribution was APC payment, without involvement in securing funding.

Fitzgerald et al. (2020) observed that the introduction of CRediT, which clearly specifies contributors' roles, has been encouraging, particularly for collaborative research. However, they argued that the 14 roles specified by CRediT may not fully suit social science journals, particularly in LIS, such as *College & Research Libraries*. To address this, they analyzed 49 papers published in *College & Research Libraries* in 2019, specifically investigating the roles of contributors involved in the production of these papers. Their analysis revealed that while the existing taxonomy is useful, roles specific to LIS and social science research should be added. One such role is "literature synthesis," which involves reviewing previous studies, identifying research gaps, and clarifying research questions—processes essential in nearly all social science research. In applied fields such as LIS, interpreting results is often a key aspect of the study, and so "interpretation of the results" should also be included. The role of "instruments" is another consideration that should be considered, as it pertains to survey instruments and interview protocols. Based on the analysis conducted by Fitzgerald et al. (2020), they recommended that LIS research include 12 roles: conceptualization, data analysis, data curation, instruments, investigation, interpretation, literature synthesis, methodology, software, visualization, writing, and editing. Their research appears to have provided valuable insights for journal editors within the LIS field as they navigate the complexities of specifying contributor roles.

What these prior studies lack, however, is a clear data-driven illustration of the active adoption of CROTs within LIS journals. To address this gap, this study investigated this matter by following the methodology described, ultimately helping foster a better understanding of CROTs' integration into LIS journals.

3. METHODS

The selection process for this study began by identifying all 82 journals indexed under the 'Information Science & Library Science' category in the 2023 edition of the Journal Citation Reports. To evaluate their engagement with CROTs, each journal's official website was meticulously analyzed, followed by a manual review of all articles published by these 82 journals throughout 2023. This initial screening revealed that 13 journals had at least one publication incorporating CROTs. Recognizing that isolated instances might not reflect an established editorial policy and aiming to focus on journals with a demonstrably consistent commitment to CROTs adoption, a threshold was set. Based on an assessment of publication frequencies within the field, and to ensure a meaningful representation of sustained adoption, a minimum of 10 articles explicitly mentioning CROTs in 2023 were required for inclusion. This criterion was instrumental in identifying the final cohort of nine journals, which exhibited a robust and complete integration of CROTs.

To understand how journals incorporate contributor roles, this study first pinpointed CROTs information. This information appeared under various headings, including 'contributors,' 'author contributions,' 'CRediT,' and 'author's notes.' Following this, each paper's CROTs section was systematically analyzed to ascertain whether its content delineated the 14 roles defined by CRediT, a prominent CROTs model (refer to Table 1) (Brand et al., 2015). Beyond this, the analysis extended to include three additional roles particularly relevant to LIS research, as suggested by Fitzgerald et al. (2020): literature synthesis, interpretation, and instruments. This broader investigation involved reviewing not only explicit CROTs sections but also the acknowledgments within the publications.

Data was gathered through a manual collection process, where CROTs information was extracted from articles published in the nine selected journals by accessing their official websites. Two doctoral students participated in this collection, with the paper's author performing a final verification of all collected data to ensure accuracy. As this study focused on a limited scope of nine specific journals, no difficulties were encountered in accessing the full text. While the possibility of journal website discrepancies is very low, it cannot be entirely excluded. For methodological clarity, single-authored articles were excluded from the analysis due to their limited relevance for CROTs analysis in non-collaborative research. Review articles were also excluded to sharpen the study's focus on complex collaborative research. This decision was further supported by their limited presence in the collected data, and preliminary data review indicated they typically involve only one or two participant roles, which diverged from the study's objective to investigate the diverse activities of participants in complex collaborative work found in original research papers.

Ultimately, the dataset for this study comprised 955 multi-authored papers. From these, 749 CROTs were identified within a subset of publications from the nine

Table 1. CRediT–Contributor Role Taxonomy

Term	Definition
Conceptualization	Ideas; formulation or evolution of overarching research goals and aims
Methodology	Development or design of methodology; creation of models
Software	Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components
Validation	Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs
Formal analysis	Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data
Investigation	Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection
Resources	Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools
Data curation	Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later reuse
Writing - original draft	Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation)
Writing - review & editing	Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision—including pre-or post-publication stages
Visualization	Preparation, creation and/or presentation of the published work, specifically visualization/data presentation
Supervision	Oversight and leadership responsibility for research activity planning and execution, including mentorship external to the core team
Project administration	Management and coordination responsibility for the research activity planning and execution
Funding acquisition	Acquisition of the financial support for the project leading to this publication

Adapted from Brand et al. (2015).

chosen journals, found specifically in their acknowledgment sections or where such details were pertinent. The subsequent analysis involved descriptive statistics applied to this collected data, and illustrative examples were included to clarify the findings.

4. RESULTS

4.1. Overview of Publishing with CROTs

While it is important to recognize that CROTs were not a standard feature in every paper, this analysis highlights that a focused group of nine journals demonstrated a more complete adoption of this framework. Table 2 shows the specific sections where CROTs were listed in each journal. Most journals included CROTs in the CRediT section, utilized by five journals, and in the author contributions section, utilized by four journals. One journal also listed them in the notes section. Table 2 indicates that Elsevier was the most active publisher in adopting CROTs, with CROTs appearing in five of their journals. Other publishers, including Taylor & Francis, Oxford University Press, PUC-Campinas, and Wiley, each had one journal that included CROTs.

Table 2. Adoption of CROTs across journals and publishers

Title (publisher)	Issues per year	CROTs		
		CRediT	Author Contributions	Notes
Gov Inform Q (E)	4	Y		
Inform Manage-Mster (E)	8	Y		
Inform Organ-UK (E)	4	Y		
Inform Process Manag (E)	6	Y		
Int J Geogr Inf Sci (T)	12		Y	Y
Int J Inform Manage (E)	6	Y		
J Am Med Inform Assn (O)	12		Y	
Learn Publ (W)	4		Y	
Transinformacao (P)	1		Y	

CROTs, Contributor Role Ontologies and Taxonomies; E, Elsevier; T, Taylor & Francis; O, Oxford University Press; W, Wiley; P, PUC-Campinas; Y, yes.

4.2. Distribution of Authorship and Contributorship

Table 3 shows the average number of authors, single-authored papers, multi-authored papers, and papers listing CROTs across the analyzed journals. The average number of authors varied significantly among the journals. For instance, the *Journal of the American Medical Informatics Association* had a high average of 8.21 authors per paper, while *Transinformacao* had a low average of 2.69 authors. On the other hand, the proportion of multi-authored papers was generally high in all journals. The journals with the highest proportion of multi-authored papers were *Information Processing & Management* and the *Journal of the American Medical Informatics Association*, both at 99%. In contrast, *Learned Publishing* had a relatively lower proportion at 82%. Meanwhile, the number of papers listing CROTs varied significantly between journals. The

Table 3. Distribution of authorship and CROTs in papers published 2023

Title	No. of authors (mean)	Type of authorship			CROTs (B)[a]	B/A (%)
		Single-authored	Multi (A) (n [%])	Total		
Gov Inform Q	3.29	7	69 (91)	76	47	68
Inform Manage-Mster	3.39	3	92 (97)	95	68	74
Inform Organ-Uk	3.00	1	19 (95)	20	11	58
Inform Process Manag	4.41	3	349 (99)	352	271	78
Int J Geogr Inf Sci	4.78	4	93 (96)	97	43	46
Int J Inform Manage	4.12	6	75 (93)	81	59	79
J Am Med Inform Assn	8.21	1	176 (99)	177	176	100
Learn Publ	3.79	12	55 (82)	67	50	91
Transinformacao	2.69	2	27 (93)	29	24	89
Total	4.19	39	955	994	749	78

CROTs, Contributor Role Ontologies and Taxonomies.
[a]This number excludes CROTs from single-authored papers.

Table 4. Distribution of CROTs contribution across papers by journal (unit: %)

	GIQ	IM	IO	IPM	IJGIS	IJIM	JAMIA	LP	TI	All journals
Conceptualization	98	96	100	94	93	95	83	58	92	90
Writing - original draft	100	93	91	96	98	92	91	92	50	89
Writing - review & editing	96	96	100	93	72	95	90	70	88	89
Methodology	98	96	73	96	70	90	39	48	92	78
Formal analysis	77	79	82	54	60	71	70	58	79	70
Investigation	66	68	91	69	58	61	51	38	63	63
Supervision	68	63	36	73	33	51	19	8	17	41
Data curation	57	60	27	60	19	58	30	24	8	38
Validation	30	49	36	56	30	42	20	12	8	32
Visualization	45	43	27	45	21	34	10	24	13	29
Project administration	36	41	45	31	16	36	6	30	8	28
Software	28	37	9	68	53	22	19	6	13	28
Funding acquisition	40	34	9	36	28	22	6	14	0	21
Resources	36	31	36	31	14	24	3	4	8	21

CROTs, Contributor Role Ontologies and Taxonomies.

Journal of the American Medical Informatics Association included CROTs in all 176 of its multi-authored papers, representing 100% of the papers. In contrast, the *International Journal of Geographical Information Science* listed CROTs in 43 out of 93 multi-authored papers, accounting for 46%.

Table 4 summarizes the contribution roles of authors listed in the papers of each journal. The most frequently cited role was conceptualization, with an average of 90%. For example, *Information and Organization* indicated conceptualization in 100% of its papers, while *Learned Publishing* did so in only 58%. Across all journals, writing-related roles such as writing - original draft (89%) and writing - review & editing (89%) were commonly specified. In contrast, roles such as funding acquisition (21%), resources (21%), and project administration (28%), which are not inherently substantial contributions to a paper's design, analysis, or writing, were listed less frequently by journals.

4.3. Applicable CROTs in LIS Journals

Table 5 (Fitzgerald et al., 2020; Maddi & Teixeira da Silva, 2024) compares the 14 roles in CRediT with the 12 roles proposed by Fitzgerald et al. (2020) for *College & Research Libraries* CROTs. This study specifically examined the three additional roles suggested by Fitzgerald et al. (2020)—instruments, interpretation, and literature synthesis—and assessed their presence in the papers.

Table 6 (Fitzgerald et al., 2020) shows the frequency with which the three roles were specified in the journals. The "Instruments" role was included in the *Journal of the American Medical Informatics Association*, while the "Interpretation" role was found in six journals, including *Government Information Quarterly*. The "Literature Synthesis" role appeared in seven journals, including *Information & Management*, and was prominently featured in all but one of those. Though literature synthesis is not directly specified as a role within the CRediT taxonomy, its inclusion is notable, as many papers explicitly identified contributors to this task. In contrast, Table 7 (Fitzgerald et al., 2020) highlights the frequency of five roles that Fitzgerald et al. (2020) considered difficult to categorize as authorship. The analysis showed significant variation in the inclusion of roles related to research projects, such as funding acquisition, project administration, and supervision, across the journals. Overall, few journals assigned contributorship to these roles, although *Information Processing & Management* was an exception, with a higher frequency of these roles being specified. This suggests that it may be important to consider the inclusion of these five roles in certain journals like *Information Processing & Management*.

4.4. Diverse Approaches to CROTs Adoption in LIS Publications

The selected journals demonstrated diverse strategies for CROTs implementation and disclosure. Of the nine journals that were published with CROTs, more than half were from Elsevier. These journals include a section titled "CRediT authorship contribution statement" at the end of the paper, where authorship contributions are outlined based on the CRediT model. However, some papers listed contributorship in author statements rather than the official section recommended by the journal (Table 8), which may be due to authors' lack of understanding of the submission guidelines.

The data analysis revealed variations in how CROTs

Table 5. Comparison of contribution roles: CRediT vs. C&RL

Category[a)]	Roles	CRediT	C&RL (Fitzgerald et al. [2020])
Author contribution roles	Conceptualization	Y	Y
	Literature synthesis	N	**Y**
	Methodology	Y	Y
	Data curation	Y	Y
	Formal analysis	Y	Y
	Investigation	Y	Y
	Instruments	N	**Y**
	Interpretation	N	**Y**
	Software	Y	Y
	Visualization	Y	Y
	Writing - original draft	Y	Y
	Writing - review & editing	Y	Y
Non-author contribution roles	Validation	Y	N
	Funding acquisition	Y	N
	Supervision	Y	N
	Project administration	Y	N
	Resources	Y	N

C&RL, *College & Research Libraries*.
a)Adapted from Maddi and Teixeira da Silva (2024), this study features a category.

Table 6. Distribution of roles incorporated by Fitzgerald et al. (2020)

Title	All CROTs	Roles (n[a)])		
		Literature synthesis	Instruments	Interpretation
Gov Inform Q	47			1 (1)
Inform Manage-Mster	68	1 (1)		
Inform Organ-Uk	11	1 (1)		
Inform Process Manag	271			8
Int J Geogr Inf Sci	43	3		4
Int J Inform Manage	59	2		
J Am Med Inform Assn	176	14 (2)	4 (2)	43 (1)
Learn Publ	50	5 (1)		5 (1)
Transinformacao	24	1		9

CROTs, Contributor Role Ontologies and Taxonomies.
a)These numbers count acknowledgements listing the specific role. Thus, this is detailed in the acknowledgements, not CRediT, author contributions, or notes.

Table 7. Distribution of roles not incorporated by Fitzgerald et al. (2020)

Title	All CROTs	Funding acquisition	Project administration	Resources	Supervision	Validation
Gov Inform Q	47	19	17	17	32	14
Inform Manage-Mster	68	23	28	21	43	33
Inform Organ-UK	11	1	5	4	4	4
Inform Process Manag	271	98	85	84	197	152
Int J Geogr Inf Sci	43	12	7	6	14	13
Int J Inform Manage	59	13	21	14	30	25
J Am Med Inform Assn	176	10	11	5	34	36
Learn Publ	50	7	15	2	4	6
Transinformacao	24	0	2	2	4	2

CROTs, Contributor Role Ontologies and Taxonomies.

Table 8. CROTs posting example 1: Author statement

Contributorship	Source
Author statement Xixi Ye: Conceptualization, Methodology, Result Analysis, Writing-Original draft preparation. Yan-Kai Fu: Methodology, Validation. Huabing Wang: Writing-final draft preparation, Supervision. Jinghong Zhou: Methodology, Validation.	Ye, X., Fu, Y. K., Wang, H., & Zhou, J. (2023). Information asymmetry evaluation in hotel e-commerce market: Dynamics and pricing strategy under pandemic. *Information Processing & Management*, 60(1), 103117.

CROTs, Contributor Role Ontologies and Taxonomies.

Table 9. CROTs posting example 2: Integrating biography into CROTs

Contributorship	Source
Notes on contributors *David A. Bennett* is a professor and department chair in the Department of Geographical and Sustainability Sciences. His research interests lie in GIScience with a focus on agent-based modelling, spatial optimization with application to sustainability and the environment. He contributed to the **literature review**, analysis of simulated data, review and editing of this paper.	Xiong, H., Zhou, X., & Bennett, D. A. (2023). Detecting spatiotemporal propagation patterns of traffic congestion from fine-grained vehicle trajectory data. *International Journal of Geographical Information Science*, 37(5), 1157-1179.

CROTs, Contributor Role Ontologies and Taxonomies.

Table 10. CROTs posting example 3: Instruments

Contributorship	Source
Acknowledgements The authors express their gratitude to both Calvin Beidleman and Kevin Peters for their **instrumental assistance** with many technical details, including alert programming and randomization of participants.	Wright, A., Schreiber, R., Bates, D. W., Aaron, S., Ai, A., Cholan, R. A., ... & Sittig, D. F. (2023). A multi-site randomized trial of a clinical decision support intervention to improve problem list completeness. *Journal of the American Medical Informatics Association*, 30(5), 899-906.

CROTs, Contributor Role Ontologies and Taxonomies.

Table 11. CROTs posting example 4: Interpretation

Contributorship	Source
Acknowledgment We thank the following individuals for their support during the collection, analysis, and **interpretation of the data**: Dzmitry Katsiuba, Yiwei Wu, Anna Boos, and Francis Cheneval. We thank all participants of the interpretation workshops. Most prominently, we thank David Wyss for his engagement during the field trial and initial data analysis, which he conducted as part of his master's thesis. We also want to thank the members of the involved police departments for the effective collaboration throughout the project. The study was supported by Swiss National Science Foundation under grants 208013 and 197485.	Dolata, M., & Schwabe, G. (2023). Moving beyond privacy and airspace safety: Guidelines for just drones in policing. *Government Information Quarterly*, 40(4), 101874.

are presented across journals. One journal, the *International Journal of Geographical Information Science*, adopted a unique format by combining contribution details with author biographies (Table 9). The journal provides information such as the author's affiliation and research interests before listing their specific roles.

Some journals also modified the CRediT format by adding additional keywords for clearer contributorship. This suggests that the CRediT taxonomy may not fully capture the diversity of roles in LIS research. Indeed, as proposed by Fitzgerald et al. (2020), there are additional roles, such as instruments, interpretation, and literature synthesis, which could be incorporated into the CRediT framework. This study found examples where these roles were specified, even though they are not part of the official CRediT taxonomy (Tables 9-11).

Additionally, the removal of five roles from the CROTs list, as suggested by Fitzgerald et al. (2020), warrants careful consideration by individual journals. *Information Processing & Management* displayed a notably higher frequency of non-author roles, such as funding acquisition, project administration, and supervision, compared to other journals. Therefore, removing these roles could potentially confuse its contributors. Conversely, journals like *College & Research Libraries*, where research project outcomes are less frequently the focus of publication, might find some detailed roles less relevant, especially the non-author roles within CRediT. In such cases, the editorial board could consider revising the CROTs roles or providing guidelines for contributors to list these roles in the CROTs or acknowledgement section.

5. DISCUSSION AND CONCLUSIONS

The collaborative and interdisciplinary nature of research in the LIS field highlights a growing need for clear contributorship specification. Despite this, research on CROTs in LIS remains extremely scarce. Existing studies, if any, have not analyzed CROTs across various journals, with only a frontier study by Fitzgerald et al. (2020) examining contributor roles in 49 papers from *College & Research Libraries*. In contrast, this study provided a significant data-driven illustration of the CROTs landscape in LIS papers. Data was collected and analyzed for 749 instances of CROTs from nine SSCI-indexed LIS journals, which often demonstrate authors' preferences for submission within the field, to assess their current implementation. This assessment revealed that the adoption of CROTs in LIS journals remains limited. However, some journals, notably Elsevier, proactively implemented CROTs by listing them for co-authors in their papers. Despite these positive efforts, instances occurred where papers were published without adhering to the prescribed CROTs format, which can be considered a significant shortcoming.

Furthermore, directly applying CRediT, the standardized contributorship model prevalent in the medical field, presents challenges for LIS research. Literature reviews, an integral part of LIS research, often span the entire study. While CRediT roles such as conceptualization or investigation may partially cover literature reviews, the taxonomy lacks an explicit role for this critical task. Therefore, it is proposed that roles such as "literature synthesis" or "literature review" be incorporated into the CRediT framework. Indeed, the data analysis revealed instances where these roles were specifically mentioned, underscoring the necessity for their inclusion. This effort aims to foster a broader discussion on ensuring fairer role distribution among research collaborators.

For example, in almost all journals, the phrase "All authors have contributed equally" was occasionally used instead of specifying individual roles. While this study identified this phenomenon, it did not concretely determine its causes. However, it is possible to speculate on potential reasons. One possibility is that authors collaborated so closely that roles overlapped rather than being strictly divided, unlike in fields such as medicine. Another reason could be that authors reluctantly complied with journal requirements for CROTs specification, resulting in vague or boilerplate statements. This may indicate a lack of awareness of CROTs' importance or a failure to reach consensus among authors on how to define their roles. Ultimately, preventing such ambiguous CROTs statements will require enhancing authors' awareness and commitment to clear role definition. Therefore, future research is needed to investigate LIS authors' perceptions of CROTs, ideally through an in-depth interview approach rather than a simple survey, as interviewing corresponding authors from papers that declare equal contribution could offer particularly rich and meaningful insights.

It is important to recognize the limitations of this study. While this research analyzed CROTs from nine SSCI-indexed LIS journals, its findings may not fully represent the diversity across the entire LIS field or other academic disciplines. The sample might be biased due to the exclusion of journals that publish fewer than 10 articles explicitly mentioning CROTs per year, potentially misrepresenting the actual adoption of CROTs in the broader LIS landscape. A limitation to be acknowledged is that while identifying CROTs-related sections (such as CRediT, Author Contributions, or Notes) is a straightforward task, the manual nature of the data collection process introduces a possibility of human error. Consequently, interpretations of these results should be considered within this context. Future research employing a broader scope and automated methods would be beneficial to validate these findings.

Consequently, future research should focus on developing next-generation CROTs standards that address the needs of academic publishing stakeholders. A crucial step involves conducting surveys or interviews with authors, editors, and publishing organizations to determine the most appropriate CROTs standards for specific fields. For LIS, once a field-specific CROTs standard is established, it should be incorporated into online submission systems to ensure each contributor's role is clearly specified. A potential solution is to make it mandatory for authors to specify their co-authors' roles, similar to how keywords are provided for research content.

Successful implementation of such initiatives could prevent non-contributors from being listed as authors and alleviate dissatisfaction among early-career researchers due to unclear role attribution. It would also ensure proper acknowledgment for researchers contributing innovative methodologies, such as software developers (implementing code, supporting algorithms, conducting code testing) and those involved in statistical analysis. This fosters professional pride and potentially opens new funding opportunities (Holcombe, 2019). Ultimately, establishing clear and fair contributorship practices will not only enhance collaboration and build trust among research teams, but also advance the transparency and maturity of academic publishing cultures.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

Allen, L., O'Connell, A., & Kiermer, V. (2019). How can we ensure visibility and diversity in research contributions? How the Contributor Role Taxonomy (CRediT) is helping the shift from authorship to contributorship. *Learned Publishing*, 32(1), 71-74. https://doi.org/10.1002/leap.1210

Allen, L., Scott, J., Brand, A., Hlava, M., & Altman, M. (2014). Publishing: Credit where credit is due. *Nature*, 508(7496), 312-313. https://doi.org/10.1038/508312a

Alpi, K. M., & Akers, K. G. (2021). CRediT for authors of articles published in the *Journal of the Medical Library Association*. *Journal of the Medical Library Association,* 109(3), 362-364. https://doi.org/10.5195/jmla.2021.1294

Brand, A., Allen, L., Altman, M., Hlava, M., & Scott, J. (2015). Beyond authorship: Attribution, contribution, collaboration, and credit. *Learned Publishing*, 28(2), 151-155. https://doi.org/10.1087/20150211

Chang, Y. W. (2018). Examining interdisciplinarity of library and information science (LIS) based on LIS articles contributed by non-LIS authors. *Scientometrics*, 116(3), 1589–1613. https://doi.org/10.1007/s11192-018-2822-7

Chang, Y. W. (2021). Characteristics of high research performance authors in the field of library and information science and those of their articles. *Scientometrics*, 126(4), 3373–3391. https://doi.org/10.1007/s11192-021-03898-y

Ding, J., Liu, C., Zheng, Q., & Cai, W. (2021). A new method of co-author credit allocation based on contributor roles taxonomy: Proof of concept and evaluation using papers published in *PLOS ONE*. *Scientometrics*, 126(9), 7561-7581. https://doi.org/10.1007/s11192-021-04075-x

Fitzgerald, S., Budd, J., Beile, P., & Kaspar, W. (2020). Modeling transparency in roles: Moving from authorship to contributorship. *College & Research Libraries*, 81(7), 1056-1060. https://doi.org/10.5860/crl.81.7.1056

Hardjosantoso, H. C., Dahi, Y., Verhemel, A., Dahi, I., & Gadjradj, P. S. (2020). Honorary authorships in the ophthalmological literature. *Journal of Current Ophthalmology*, 32(2), 199-202. https://doi.org/10.4103/JOCO.JOCO_104_20

Herbert, B. E., & Kaspar, W. A. (2019). Authorship and the consideration of alternatives. *College & Research Libraries,* 80(1), 2-7. *https://doi.org/10.5860/crl.80.1.2*

Holcombe, A. O. (2019). Contributorship, not authorship: Use CRediT to indicate who did what. *Publications*, 7(3), 48. https://doi.org/10.3390/publications7030048

Holcombe, A. O., Kovacs, M., Aust, F., & Aczel, B. (2021). Documenting contributions to scholarly articles using CRediT and *tenzing*. *PLOS ONE*, 15(12), e0244611. https://doi.org/10.1371/journal.pone.0244611

Hosseini, M., Colomb, J., Holcombe, A. O., Kern, B., Vasilevsky, N. A., & Holmes, K. L. (2023). Evolution and adoption of contributor role ontologies and taxonomies. *Learned Publishing*, 36(2), 275-284. https://doi.org/10.1002/leap.1496

Hosseini, M., Gordijn, B., Wafford, Q. E., & Holmes, K. L. (2024). A systematic scoping review of the ethics of Contributor Role Ontologies and Taxonomies. *Accountability in Research*, 31(6), 678-705. https://doi.org/10.1080/08989621.2022.2161049

Jabbehdari, S., & Walsh, J. P. (2017). Authorship norms and project structures in science. *Science, Technology, & Human Values*, 42(5), 872-900. https://doi.org/10.1177/0162243917697192

Jokić, M. (2020). Productivity, visibility, authorship, and collaboration in library and information science journals: Central and Eastern European authors. *Scientometrics*, 122(2), 1189-1219. https://doi.org/10.1007/s11192-019-03308-4

Larivière, V., Gingras, Y., Sugimoto, C. R., & Tsou, A. (2015). Team size matters: Collaboration and scientific impact since 1900. *Journal of the Association for Information Science and Technology*, 66(7), 1323-1332. https://doi.org/10.1002/asi.23266

Larivière, V., Pontille, D., & Sugimoto, C. R. (2021). Investigating the division of scientific labor using the Contributor Roles Taxonomy (CRediT). *Quantitative Science Studies*, 2(1), 111-128. https://doi.org/10.1162/qss_a_00097

Lee, J., Yang, K., & Oh, D. G. (2020). Factors influencing the choice of a publication venue in library and information science. *Learned Publishing*, 33(3), 323-332. https://doi.org/10.1002/leap.1300

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, 35(5), 673-702. https://doi.org/10.1177/0306312705052359

Maddi, A., & Teixeira da Silva, J. A. (2024). Beyond authorship: Analyzing contributions in *PLOS ONE* and the challenges of appropriate attribution. *Journal of Data and Information Science*, 9(3), 88-115. https://doi.org/10.2478/jdis-2024-0015

Mowatt, G., Shirran, L., Grimshaw, J. M., Rennie, D., Flanagin, A., Yank, V., MacLennan, G., Gøtzsche, P. C., & Bero, L. A. (2002). Prevalence of honorary and ghost authorship in Cochrane reviews. *JAMA*, 287(21), 2769-2771. https://doi.

org/10.1001/jama.287.21.2769

Mwantimwa, K., & Kassim, M. (2023). Why does researchers' collaboration matter in research activities? *Higher Education Quarterly*, 77(4), 741-755. https://doi.org/10.1111/hequ.12431

Nicholas, D., Herman, E., Clark, D., Boukacem-Zeghmouri, C., Rodríguez-Bravo, B., Abrizah, A., Watkinson, A., Xu, J., Sims, D., Serbina, G., Świgoń, M., Jamali, H. R., Tenopir, C., & Allard, S. (2022). Choosing the 'right' journal for publication: Perceptions and practices of pandemic-era early career researchers. *Learned Publishing*, 35(4), 605-616. https://doi.org/10.1002/leap.1488

Paul-Hus, A., Mongeon, P., Sainte-Marie, M., & Larivière, V. (2017). The sum of it all: Revealing collaboration patterns by combining authorship and acknowledgements. *Journal of Informetrics*, 11(1), 80-87. https://doi.org/10.1016/j.joi.2016.11.005

Petras, V. (2023). The identity of information science. *Journal of Documentation*, 80(3), 579-596. https://doi.org/10.1108/JD-04-2023-0074

Prebor, G. (2010). Analysis of the interdisciplinary nature of library and information science. *Journal of Librarianship and Information Science*, 42(4), 256-267. https://doi.org/10.1177/0961000610380820

Reller, T. (2019). *Elsevier expands CRediT approach to authorship*. https://www.prnewswire.com/news-releases/elsevier-expands-credit-approach-to-authorship-300972468.html

Rowley, J., Sbaffi, L., Sugden, M., & Gilbert, A. (2022). Factors influencing researchers' journal selection decisions. *Journal of Information Science*, 48(3), 321-335. https://doi.org/10.1177/0165551520958591

Teixeira da Silva, J. A., Tsigaris, P., & Vuong, Q. H. (2023). Acknowledgments in scientific papers. *Publishing Research Quarterly*, 39(3), 280-299. https://doi.org/10.1007/s12109-023-09955-z

Vakkari, P., Järvelin, K., & Chang, Y. W. (2023). The association of disciplinary background with the evolution of topics and methods in library and information science research 1995–2015. *Journal of the Association for Information Science and Technology*, 74(7), 811-827. https://doi.org/10.1002/asi.24757

Vasilevsky, N. A., Hosseini, M., Teplitzky, S., Ilik, V., Mohammadi, E., Schneider, J., Kern, B., Colomb, J., Edmunds, S. C., Gutzman, K., Himmelstein, D. S., White, M., Smith, B., O'Keefe, L., Haendel, M., & Holmes, K. L. (2021). Is authorship sufficient for today's collaborative research? A call for contributor roles. *Accountability in Research*, 28(1), 23-43. https://doi.org/10.1080/08989621.2020.1779591

Ware, M., & Mabe, M. (2015). *The STM report: An overview of scientific and scholarly journal publishing* (4th ed.). STM: International Association of Scientific, Technical and Medical Publishers.

Wislar, J. S., Flanagin, A., Fontanarosa, P. B., & DeAngelis, C. D. (2011). Honorary and ghost authorship in high impact biomedical journals: A cross sectional survey. *BMJ*, 343, d6128. https://doi.org/10.1136/bmj.d6128

JISTaP https://www.jistap.org
Journal of Information Science Theory and Practice

RESEARCH PAPER
J Inf Sci Theory Pract 13(4): 58-68, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.4.4

Citation Metric Misuse, Interpretive Ambiguity, and Interdisciplinary Complexity—The Case in Biomechanics Journals

Duane Knudson*
Department of Health & Human Performance, Texas State University,
San Marcos, TX, USA
E-mail: dknudson@txstate.edu

ABSTRACT

Biomechanics is an interdisciplinary field with varying citation patterns and centrality of journals where researchers publish. The audit culture of scientific publication has driven persistent, inconsistent, and questionable interpretation of citations and metrics in selecting and evaluating journals. This article extends the understanding of bias in citation patterns and journal metrics for eight biomechanics journals. Citations from three databases and several journal metrics were examined from 2019 to 2023. Long-term changes (1999-2023) were documented for the most prestigious journal in the field (*Journal of Biomechanics*). There was large variation and positive skew in citations and citation rates to top cited articles in the biomechanics journals from 2021 in all databases. The skew and variation in citations and subject categories/areas assigned by databases contributed to biased journal usage metrics for biomechanics journals. For the *Journal of Biomechanics*, citation patterns led to opposing changes in overall usage (Journal Impact Factor) and field-normalized usage (Source-Normalized Impact per Paper) journal metrics. Evidence-based interpretations of journal metrics are illustrated and recommendations made to limit bias in planning, publishing, and evaluating research in interdisciplinary fields such as biomechanics.

Keywords: bias, highly cited, impact, meta-science, subject category

Received: June 10, 2025 **Revised:** July 15, 2025
Accepted: August 7, 2025 **Published:** December 30, 2025

***Corresponding Author:** Duane Knudson
https://orcid.org/0000-0003-0809-7970
E-mail: dknudson@txstate.edu

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee or authors' permission. Open Access articles are automatically archived in the Korea Institute of Science and Technology Information (KISTI)'s Open Access repository (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

© 2025 [Duane Knudson] This article is distributed under the Creative Commons Attribution License (CC BY), allowing unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. It is published by the Korea Institute of Science and Technology Information (KISTI).

1. INTRODUCTION

Citations to previous research is an essential standard of science to give credit to scholars and to integrate results to advance knowledge. The development of indexing of scientific publications, particularly computerized electronic database services that emphasize the tracking and analysis of citations to publications, are increasingly influential in modern research. Garfield (2006) summarized the development of systematic tracking of article citations for use with the Journal Impact Factor (JIF) and the *Science Citation Index*, which has become the *Web of Science*™ and *Journal Citation Reports*™ (Clarivate) bibliometric services. The JIF was developed for librarians to manage holdings with limited financial resources but has been co-opted by many as an indicator of vaguely defined journal and article quality.

The speed of knowledge advancement continues to increase with technological advances in computer-assisted electronic publication and indexing available through the Internet. The fast, hypercompetitive world of grants and publications pressure some researchers to cut corners, illogically 'judging books by their covers' in evaluating research reports by the journal where they are published, indexed, or appear in the JIF. This misuse of citation metrics such as the JIF has been conclusively refuted and criticized for several reasons including skewed distributions and disciplinary differences in citation patterns (Adler et al., 2009; Bollen et al., 2006; Declaration on Research Assessment [DORA], n.d.; Hicks et al., 2015; Kurmis, 2003; Seglen, 1997; Smith, 1998).

Nevertheless, there has been a rapid proliferation of scientific journals (Ghasemi et al., 2022; Pandita & Singh, 2023) and metrics to evaluate journals, articles, and authors (Casadevall & Fang, 2014; Roldan-Valadez et al., 2019; Wilsdon et al., 2015). Interest in journal metrics like the JIF for ranking journals has been described as fetishistic, foolish, mindless, insidious misuse, and an obsession (Adler et al., 2009; Hicks et al., 2015; Hussain, 2015; Osterloh & Frey, 2020; Tourish & Willmott, 2015; Willmott, 2011). Ranking journals using citation metrics result in imprecise and contradictory rankings of "core" or "top-tier" journal status (Knudson, 2024a; 2024c; Knudson et al., 2024; Mason & Singh, 2022; Pajić, 2015; Vîiu & Păunescu, 2021; Wilsdon et al., 2015). The numerous metrics, database sources for data, vague and conflicting definitions of publication quality dimensions, and several meta-science fields (bibliometrics, informetrics, and scientometrics) contribute to considerable miscommunication and conflicting results.

Progress in journal citation metric research is hampered by miscommunication because of a lack of agreement on research/scholarly quality constructs and inconsistent use of terminology (Mason & Singh, 2022). Many terms used are not defined (Egghe, 2022) and are used in conflicting and in synonymous fashion (Aksnes et al., 2019; Helmer et al., 2020; Knudson et al., 2024), including activity, contribution, influence, impact, importance, merit, output, progress, popularity, progress, quality, rigor, usage, and visibility. Despite the terminology chaos, the consensus of the meta-science literature is that while journal metrics are positively associated with each other, they tend to serve as imprecise proxies for one of two constructs: usage/impact or prestige (Bollen et al., 2006; 2009; Franceschet, 2010; Leydesdorff et al., 2016; Zhou et al., 2012). Knudson et al. (2024, p. 95) defined these constructs: "Usage is the archive role of journals in providing transitory, shorter-term observations at the knowledge front (sometimes referred to as popularity). Prestige represents the perception of advancement and codification of knowledge." Many scholars are not familiar with this two-factor structure and evidence-based interpretation of citation metrics.

It is well known from analysis of citations content and context that citation behavior is multidimensional, varies in function and polarity, and differs across disciplines (Tahamtan & Bornmann, 2019). Disciplines also have widely varying numbers and timings of citations (Hayashi & Fujigaki, 1999; Seglen, 1997; Waltman, 2016). Nuanced use and interpretation of citation data allow study of knowledge flow and connections between disciplines (van Leeuwen & Tijssen, 2000; Wang et al., 2025). Misuse of citation metrics, however, often disadvantage interdisciplinary fields, journals, and research given their diversity and variation in citation patterns.

This issue can be particularly problematic in interdisciplinary fields such as kinesiology and biomechanics with widely varying citation patterns (Knudson, 2023; 2024b; Knudson et al., 2024). For example, biomechanics scholars may be disadvantaged in a tenure and promotion review by publishing in a specialized biomechanics journal with a low JIF, incorrectly perceived by faculty as having low quality and prestige, when this metric is aligned with usage that is relatively low because of journal size and specialization, despite the journal being the most relevant and respected in that research area. Grant or institutional evaluations focusing incorrectly on journal usage metrics like JIF may disadvantage some interdisciplinary journals and research if the topics are not currently popular, attracting many citations in the short (2-year) time window of the predominant JIF.

1.1. Citation Metrics in the Interdisciplinary Field of Biomechanics

Biomechanics is a diverse, interdisciplinary field with varying citation patterns and centrality of journals in which scholars publish (Knudson, 2023; 2024b). Journal evaluation and selection by authors in the interdisciplinary field of biomechanics is complicated by journal mission and other factors. Biomechanics include the integration of biological and physical sciences, many kinds of animal and plant motion, and numerous areas of application (allied health, engineering, ergonomics, medicine, sports). These factors may contribute to wide variation in journal citation metrics across areas of biomechanics application (Knudson, 2014; 2015a; 2015b; 2023; 2024b). These studies report that the top 10 to 50 cited articles in biomechanics journals have median citations between 40 to 360 in *Google Scholar* and 8 to 158 citations in *Web of Science*™ over two or more years.

This variation in citations to biomechanics journals is consistent with variation in citations to different research subject categories assigned to biomechanics journals (Knudson, 2023). Khademi et al. (2023) reported that word analysis of Iranian biomechanics articles indexed in *Web of Science*™ also favored the injury research categories of tissue biomechanics and orthopedics, over the categories of body movements, computational, occupational and spine, and corneal biomechanics. The inconsistency in biomechanics journal metrics and subject categories agrees with evidence that the American Society of Biomechanics (ASB) members' perceptions of journal quality vary by their own research interest area in the society (Knudson & Chow, 2008).

A recent study documented the change in the centrality of the citation network of top cited biomechanics journals according to citations indexed in *Journal Citation Reports*™ from 2011 to 2021 using the "Journal Citation Relationships" report (Knudson, 2024b). This study found that the number of top journals cited in twelve seed biomechanics journals in *Journal Citation Reports*™ decreased 17% despite a five-fold increase in citations, indicating a considerable change and concentration of top journals cited in the field. Specifically, over the last decade mega-journals increased from being one to eight of the top 25 cited journals in biomechanics. The top ten journals at the center of the biomechanics citation network in 2021 (based on two-year JIF as a proxy for overall academic usage) included five mega-journals (*): *Journal of Biomechanics*, *Gait & Posture*, *Sensors**, *Scientific Reports**, *Journal of the Mechanical Behavior of Biomedical Materials*, *Applied Sciences**, *Frontiers in Bioengineering and Biotechnology**, *PLoS ONE**, *Sports Biomechanics*, and *Journal of Biomechanical Engineering*. Planning and publishing biomechanics research may be adversely influenced by the misuse of citation metrics in indexing, searching for and evaluating research, and in selecting journal outlets for reports.

1.2. Purpose

This study extends knowledge of bias in citation patterns to top cited biomechanics journals related to the kinesiology/exercise science field by documenting journal citation metrics from three database services. It was hypothesized that top cited biomechanics journals would have variable citations and journal citation metrics. Appropriate interpretation of these metrics is illustrated and recommendations made to avoid their misuse in biomechanics and similar interdisciplinary fields.

2. METHODOLOGY

The *Journal of Biomechanics* and seven other top cited biomechanics journals related to kinesiology were selected as a sample. Biomechanics journals related to kinesiology were selected because this field is also interdisciplinary and is one of the five categories (a.k.a. "exercise and sports science") that can be identified by ASB members (ASB, n.d.). The present author, a scholar in both fields, selected these eight journals using the results of a previous study of the top 30 cited biomechanics journals (Knudson, 2024b, Table 3, p. 3261) based on the Journal Citation Relationships feature of *Journal Citation Reports*™. The *Journal of Biomechanics* is particularly relevant given it is the oldest, most respected (Knudson & Chow, 2008) journal in the field, and is most cited and central (Knudson, 2024b). The next top seven cited journals related to kinesiology were selected (Table 1). Journals that may occasionally report research by kinesiology scholars were excluded, primarily multidisciplinary mega-journals (Ioannidis et al., 2023) and journals in other areas of biomechanics (e.g., engineering).

Three databases were searched to identify the articles, citations, and citation rates to the top 20 cited articles from the eight biomechanics journals for the same year (2021) as the Knudson (2024b) study: *Dimensions*, *Google Scholar*, and *Web of Science*™. Citations to top

Table 1. Descriptive statistics of indexed citations to the top 20 articles for 2021 by *Google Scholar* citations for three database services

Journal	(A) *Google Scholar*				(B) *Web of Science*™				(C) *Dimensions*			
	γ	M_e	*Max*	*Min*	γ	M_e	*Max*	*Min*	γ	M_e	*Max*	*Min*
J Biomech	2.1	70	243	46	2.4	45	164	31	2.0	52	179	37
Gait Posture	1.5	49	127	36	1.8	33	83	16	2.2	37	107	21
Sports Biomech	1.2	63	117	46	1.3	35	76	21	0.9	40	77	25
Clin Biomech	2.1	22	54	18	0.9	16	35	3	1.9	18	47	7
J App Biomech	1.2	11	22	10	1.9	8	20	4	1.6	9	19	5
Hum Mov Sci	1.4	33	88	13	1.3	19	49	6	1.3	23	62	5
J EMG Kines	2.8	22	85	13	1.7	15	50	5	3.0	16	81	8
J Stren Cond Res	1.8	94	242	76	2.2	47	129	33	2.2	56	143	45
Mean	1.7	46	122	33	1.7	27	76	15	1.9	31	77	19

Google Scholar totals were acquired March 21, 2025.
γ, skew; M_e, median; *Max*, maximum score; *Min*, minimum score.

percentiles (75 to 95%) of articles published by journals are the focus of meta-science studies because of the strong positive skew and the many uncited articles (Bornmann & Marx, 2014; 2015). *Google Scholar* totals were acquired March 21, 2025.

Four common journal citation metrics for the *Journal of Biomechanics* were also collected for the last five available years (2021-2023) to illustrate short-term, annual variation. These metrics usually align best with the scholarly usage/popularity construct and are based on the two most prestigious databases: *Scopus*® (Elsevier, Amsterdam, Netherlands) and *Web of Science*™ (Clarivate™, London, UK) (Pranckutė, 2021). Journal metrics from *Scopus*® data include the CiteScore (CS_4), SCImago Journal Rank (SJR_3), and Source-Normalized Impact per Paper ($SNIP_3$). The metric based on *Web of Science*™ data was the two-year Journal Impact Factor (JIF_2). Subscripts in metric abbreviations indicate the time window used to analyze citations. Additional historical data (1999-2023) of the JIF_2 and $SNIP_3$ for the *Journal of Biomechanics* were also collected to illustrate long-term trends. This allowed the illustration of differences in general and field-normalized usage metrics. These two metrics had values for this longer timeframe and would show similar trends given the high correlation between all journal usage metrics.

Data were analyzed with JMP Pro 17.2.0 (SAS Institute, Cary, NC, USA) software for statistical analysis. Continuous data were plotted and tested for normality with the Shapiro-Wilk test and calculation of skew. Appropriate descriptive statistics were calculated and reported. The limitation of inferential statistical analysis across time or journals was necessary given the small number of journals in this specialization of biomechanics and their widely varying size.

3. RESULTS AND DISCUSSION

3.1. Citation Variation and Usage

Citations to top cited articles in the biomechanics journals in all databases had large positive skews, with median citations varying between 11 and 94 (*Google Scholar*), 8 to 47 (*Web of Science*™), and 9 to 56 (*Dimensions*). The top values, from *Google Scholar* and *Web of Science*™, were within the 8 to 360 citations reported for biomechanics journals (Knudson, 2014; 2015a; 2015b; 2023; 2024b). Converting the citations in Table 1 to citation rates (citations/3 years) shows variation from 2.6 to 3.7 citations per year to the top 20 articles in *Journal of Applied Biomechanics*, up to 15 to 23.3 citations per year for the *Journal of Biomechanics*. This confirms past reports of large variation of citation rates (3 to 79 citations/year) across biomechanics journals (Knudson, 2014; 2015a; 2015b; 2023; 2024b). Citation variation across biomechanics journals from different subject categories is also large, calling into question the ability to field normalize research in biomechanics (Knudson, 2023). Strong citations and citation rates to some biomechanics journals is impressive scholarly usage given that medical and sports medicine journals tend to have larger numbers of references, earlier citation, and higher citation rates than most other disciplines like biomechanics (Adler et al., 2009; Dorta-González & Dorta-González, 2013; Finardi, 2014; Knudson, 2013; 2015b; 2022; Patience et al., 2017).

The *Journal of Biomechanics* consistently had high median citations (45 to 70) compared to other biomechanics journals but was nominally lower in median citations than the *Journal of Strength and Conditioning Research* (47 to 94) in all three databases (Table 1). This contributed to a higher subsequent 2021 JIF_2 for the *Journal of Strength and Conditioning Research* (4.4) than the *Journal of Biomechanics* (2.8). A minority of the top 20 cited articles in the *Journal of Strength and Conditioning Research*, however, relate to biomechanics theory, methods, or variables. The common inference of higher likely usage of a biomechanics article in the *Journal of Strength and Conditioning Research* from its citations alone from any database or the JIF_2 would be misleading given the skew in citations and fewer biomechanics articles in the journal. The *Journal of Biomechanics* was ranked 65th of 98 journals by JIF_2 in the engineering, biomedical 'category' while the *Journal of Strength and Conditioning Research* was ranked 18th of 88 journals by JIF in the sports sciences category in 2021. The *Journal Citation Reports*™ currently classifies the *Journal of Strength and Conditioning Research* into the Sports Sciences (n=127) category and the *Journal of Biomechanics* into the Biophysics (n=77) and Engineering, Biomedical (n=123) categories. Comparing these journals would not be appropriate with the JIF_2 from this database given the journals cover different disciplines of science and the journals ranked vary over time.

Table 2. Recent variation in four major journal citation metrics for the *Journal of Biomechanics*

	CS_4	SJR_3	$SNIP_3$	JIF_2
Year				
2019	4.7	1.0	1.3	2.3
2020	4.3	0.8	1.3	2.7
2021	4.4	0.8	1.3	2.8
2022	4.9	0.7	1.2	2.4
2023	5.1	0.7	1.1	2.4
Value				
Mean	4.68	0.80	1.24	2.52
SD	0.33	0.12	0.09	0.22
CV (%)	7.15	15.52	7.21	8.60
Upper 95% CI	5.10	0.95	1.35	2.79
Lower 95% CI	4.26	0.65	1.13	2.25

Metrics based on *Scopus* data were the CiteScore (CS_4), SCImago Journal Rank (SJR_3), Source-Normalized and Impact per Paper ($SNIP_3$), while the Journal Impact Factor (JIF_2) was based on data from *Web of Science*™.

Table 2 lists above average short-term (2019-2023) usage values for the *Journal of Biomechanics* based on four journal citation metrics. The journal currently ranks 133rd of 400 journals by CS_4 and 101st out of 326 by SJR_3 in the biomedical engineering 'subject area' of *Scopus*® for 2023. The short-term relative variation (coefficient of variation [*CV*] between 7 and 16%) of this top journal was slightly more consistent than the five-year variation (*CV*=12-21%) reported for five metrics from fourteen biomechanics journals (Knudson & Quimby, 2023). Year-to-year variation greater than 10% to 30% (Amin & Mabe, 2003; Haghdoost et al., 2014; Knudson & Quimby, 2023; Ogden & Bartley, 2008) and statistical tests (Bornmann, 2017; Schubert & Glänzel, 1983) have been advocated as a standard for meaningful change in most journal citation metrics. A scholar interpreting these recent metrics should note that the journal has slightly above (1.1 to 1.4) average field-normalized usage (1.0) according to $SNIP_3$, with more average general usage based on CS_4, JIF_2, and SJR_3. This should be interpreted as confirmatory evidence of likely usage of articles published in the journal. Qualifications include that these metrics are biased from the positive skew of citations to articles in journals, and so median values are most appropriate for comparison, and these are tentative proxies for scholarly usage, not rigor or prestige.

3.2. Long-Term Usage and Evidence-Based Interpretation

Fig. 1 illustrates long-term, opposing trends of the $SNIP_3$ and JIF_2 usage metrics for the *Journal of Biomechanics*.

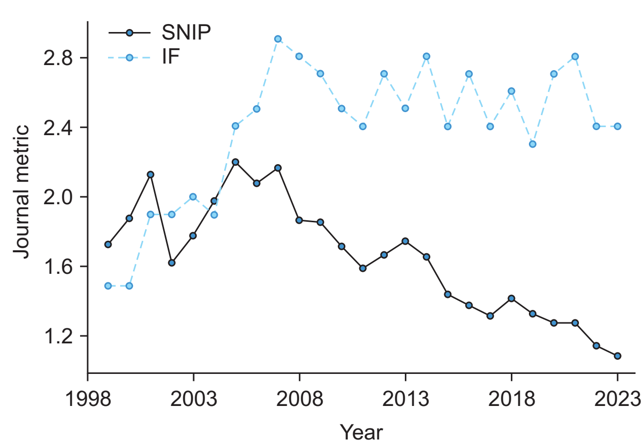

Fig. 1. Long-term (1999-2023) trends in two usage metrics for the *Journal of Biomechanics*.

chanics, beginning in the early 2000s. This example illustrates the different information that can be obtained between a general journal usage metric and field-normalized metric. How should these usage results be interpreted by biomechanists interested in this journal as a publication outlet? Biomechanists should first review meta-science research to know that the average citation rate represented by JIF_2 has more bias (positive skew of citations, short time window, and no normalization for disciplinary citation patterns) than the $SNIP_3$, and so it is hard to interpret the magnitude of the JIF_2 proxy estimate of overall scholarly usage based on the *Web of Science*™ database service. The $SNIP_3$ has a longer time window and field normalization, and so it may serve as a proxy estimate of field (albeit interdisciplinary) usage from a larger database (*Scopus*®). Second, any change in these metrics should be interpreted based on year-to-year variation (Table 2) and the 20-30% change needed in most metrics to be meaningful for biomechanics journals (Knudson & Quimby, 2023). Interpreting the size of journal metric values also requires consideration of mediating factors. For example, like many journal metrics, one cannot calculate $SNIP_3$ on one's own because the *Scopus*® subject area used is unclear, given the journal currently has four (biomedical engineering, biophysics, orthopedics and sports medicine, and rehabilitation). The database currently classifies the journal into two subject areas similar to *Journal Citation Reports*™ research areas but while covering more journals: Biophysics (n=225) and Engineering, Biomedical (n=400).

The $SNIP_3$ data in Fig. 1 indicates there has been a gradual decline in 'field' usage of average articles in the *Journal of Biomechanics* from about twice the usage to about average (1.1). Given the size of this drop is about 50%, there may be a meaningful decrease in field-specific usage of the *Journal of Biomechanics* articles beginning after 2007 based on $SNIP_3$. This result is consistent with the change in citation density in biomechanics journals to include several more mega-journals than in the past (Knudson, 2024b). The apparent decline of the *Journal of Biomechanics* in field, however, provides no information on journal or article quality and could be the result of indexing bias, subject area variation, an artifact of shifting citation patterns in biomechanics (Knudson, 2024b), and the recent rapid growth of biomedical articles published in multi-disciplinary and specialized mega-journals (Ioannidis et al., 2023; Siler et al., 2020). Further research is needed in improving transparency and consistency of subject categories and field-normalization for disciplinary and interdisciplinary fields and journals.

The overall usage of average *Journal of Biomechanics* articles based on JIF_2 and *Journal Citation Reports*™, however, has increased from 1.5 to 2.5, approaching earlier values (Fig. 1). Zadpoor and Nikooyan (2011) reported increases in the JIF_2 of 14 biomechanics journals from 2003 until 2010. They also noted, however, that the rate of increase in JIF_2 in biomechanics journals was less than for journals in other *Journal Citation Reports*™ subject categories (e.g., biomedical, biomaterials). While this increase in overall usage beginning in the early 2000s is larger than the 30% standard, are there other factors to consider? One is a long history of increases in the JIF (Althouse et al., 2009) with recent increases from 1.1 in 1997 to 2.2 in 2016 (Fischer & Steiger, 2018). This indicates that the increase in typical overall scholarly usage of *Journal of Biomechanics* articles results from both the research topics published in the journal, growth in articles published by the journal (521-613 articles published from 1999 until 2005 to consistently over 1,200 after 2008), and general JIF temporal inflation. The scholarly audit culture that overemphasizes the JIF likely leads to the metrics' inflation from increasing numbers of journals, the articles they publish, the number of references cited per article, and unethical manipulation (Ioannidis & Thombs, 2019).

4. RECOMMENDATIONS

Given the results of this study and the well documented limitations of numerous citation metrics and databases cited in this report, biomechanics researchers and scholars in similar interdisciplinary fields should consider four strategies to limit bias in planning and publishing their research. Consistent use of these strategies by scholars in biomechanics could contribute to other research quality initiatives (Glasziou et al., 2014; Ioannidis, 2005; Knudson, 2017; Vagenas et al., 2018) for more meaningful results and faster advancement of knowledge. DORA (2024) and other sources (DORA, n.d.; Hicks et al., 2015; Walters, 2017; Wilsdon et al., 2015) also provide helpful recommendations on avoiding misuse of citation metrics in research.

4.1. Recognize and Avoid Misuse of Journal Citation Metrics

Biomechanics researchers and journal editors should recognize, avoid, and not reward misuse of citation metrics at all levels: article, author, and journal. They should recognize there is compelling evidence that single, journal-level citation metrics are strongly biased, inaccurate for individual articles, and associate poorly with expert, peer evaluation of research quality (Haddawy et al., 2016; Mahmood, 2017; Thelwall et al., 2023). They should avoid using any single citation metric as a precise, ratio-level measure of any specific journal quality construct. They should evaluate research reports on their own merits and not reward misuse of citation metrics, assuming the reports are better or worse based on unreliable ranks or tiers from citation metrics. In addition, it is important to consider a more public stand by considering personal and support organizational (journal, society) endorsement of the Declaration on Research Assessment or DORA (n.d.).

4.2. Recognize Citation Metric Bias in Biomechanics and Interdisciplinary Fields

Biomechanists should recognize that there is bias in all journal citation metrics, and that some metrics disadvantage interdisciplinary journals (e.g., biomechanics) relative to journals that are larger, multidisciplinary, from faster moving disciplines (e.g., biomedical, computer science, genetics), or mega-journals. High-quality interdisciplinary research is difficult and takes time, so relatively lower journal citation metrics in some biomechanics journals is expected and may not be related to lower quality, usage/impact, or prestige. Biomechanics research is interdisciplinary and diverse, and so citation patterns vary significantly in different research areas or database subject categories assigned, leading to additional bias in journal citation metrics. Interdisciplinary researchers should normally seek first to publish in journals relevant to their research topic to reach the scholarly audience, and not chase false visibility and usage in presumed high-impact journals.

4.3. Use Meta-Science for Evidence-Based Interpretation of Citation Metrics

Biomechanists should strive to use evidence-based interpretation of multiple citation metrics based on the consensus of meta-science research. Citation metrics should be used primarily at article and author level, not the journal level. If journal metrics are required by some institutional policy, it is best to strive to use multiple citation metrics only as confirmatory evidence about the likely usage or prestige expected for a particular journal in the context of that metric and databases' limitations. One should strive to use field normalized citation metrics when possible and interpret them carefully, especially given interdisciplinary fields like biomechanics. This is particularly important given the variability in citation patterns across disciplines and within biomechanics, and because different databases use different subject categories and lack a specific biomechanics category. Importantly, do not infer and imply in a communication that any journal citation metric is a proxy for the quality of a research report appearing in a journal (Dougherty & Horne, 2022; Haddawy et al., 2016; Mahmood, 2017; Starbuck, 2005; Thelwall et al., 2023).

4.4. Use Multiple and Thorough Database Searches

Biomechanics and all researchers should strive to use numerous, careful searches of multiple databases to adequately find relevant research. Given the diversity and interdisciplinary nature of biomechanics, research in the field appears in a wide variety of journals. Biomechanics knowledge advances the most when all relevant research is synthesized in planning original research and systematic review reports. The many biases introduced by large variation in citation patterns in disciplines increasingly influence database coverage, search engines, and relevance ranking of results (Bade, 2007; Delgado-Quirós & Ortega, 2024; 2025; Drivas, 2025; Gusenbauer & Haddaway, 2020; Kiester & Turp, 2022; Knudson, 2022; 2023; Mongeon & Paul-Hus, 2016; Pozsgai et al., 2021).

5. CONCLUSIONS

Eight top cited biomechanics journals related to kinesiology had strong positive skews of top cited articles indexed in three major database services. Short- and long-term longitudinal data for the most prestigious journal in the field, the *Journal of Biomechanics* (Knudson, 2024a; Knudson & Chow, 2008) illustrated the need for evidence-based interpretation of journal citation metrics for correct interpretation, avoiding biases that adversely affect diverse, slower, interdisciplinary fields. Variation in four journal usage metrics for the *Journal of Biomechanics* indicated different and meaningful long-term trends over time between overall (JIF_2) and field-normalized ($SNIP_3$) metrics. The high prestige of the journal is not reflected in general and commonly misused JIF_2 usage metrics. This is a good example of bias in journal metrics, particularly for interdisciplinary journals, and of the resulting problems from oversimplification and misuse of journal metrics and their lack of evidence-based interpretation in disciplinary and interdisciplinary contexts.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

Adler, R., Ewing, J., & Taylor, P. (2009). Citation: Statistics a report from the International Mathematical Union (IMU) in cooperation with the International Council of Industrial and Applied Mathematics (ICIAM) and the Institute of Mathematical Statistics (IMS). *Statistical Science*, 24(1), 1-14. https://doi.org/10.1214/09-STS285

Aksnes, D. W., Langfeldt, L., & Wouters, P. (2019). Citations, citation indicators, and research quality: An overview of basic concepts and theories. *SAGE Open*, 9(1), 2158244019829575. https://doi.org/10.1177/2158244019829575

Althouse, B. M., West, J. D., Bergstrom, C. T., & Bergstrom, T. (2009). Differences in impact factor across fields and over time. *Journal of the American Society for Information Science and Technology*, 60(1), 27-34. https://doi.org/10.1002/asi.20936

American Society of Biomechanics (ASB). (n.d.). *About*. https://asbweb.org/about/

Amin, M., & Mabe, M. A. (2003). Impact factors: Use and abuse. *Medicina (B Aires)*, 63(4), 347-354.

Bade, D. (2007). Relevance ranking is not relevance ranking or, when the user is not the user, the search results are not search results. *Online Information Review*, 31(6), 831-844. https://doi.org/10.1108/14684520710841793

Bollen, J., Rodriquez, M. A., & Van de Sompel, H. (2006). Journal status. *Scientometrics*, 69(3), 669-687. https://doi.org/10.1007/s11192-006-0176-z

Bollen, J., Van de Sompel, H., Hagberg, A., & Chute, R. (2009). A principal component analysis of 39 scientific impact measures. *PLoS ONE*, 4(6), e6022. https://doi.org/10.1371/journal.pone.0006022

Bornmann, L. (2017). Confidence intervals for Journal Impact Factors. *Scientometrics*, 111(3), 1869-1871. https://doi.org/10.1007/s11192-017-2365-3

Bornmann, L., & Marx, W. (2014). How to evaluate individual researchers working in the natural and life sciences meaningfully? A proposal of methods based on percentiles of citations. *Scientometrics*, 98(1), 487-509. https://doi.org/10.1007/s11192-013-1161-y

Bornmann, L., & Marx, W. (2015). Methods for the generation of normalized citation impact scores in bibliometrics: Which method best reflects the judgements of experts? *Journal of Informetrics*, 9(2), 408-418. https://doi.org/10.1016/j.joi.2015.01.006

Casadevall, A., & Fang, F. C. (2014). Causes for the persistence of impact factor mania. *mBio*, 5(2). https://doi.org/10.1128/mbio.00064-14

Declaration on Research Assessment (DORA). (2024). *Guidance on the responsible use of quantitative indicators in research assessment*. DORA. https://doi.org/10.5281/zenodo.10979644

Declaration on Research Assessment (DORA). (n.d.). *San Francisco declaration on research assessment*. https://sfdora.org/read/

Delgado-Quirós, L., & Ortega, J. L. (2024). Completeness degree of publication metadata in eight free-access scholarly databases. *Quantitative Science Studies*, 5(1), 31-49. https://doi.org/10.1162/qss_a_00286

Delgado-Quirós, L., & Ortega, J. L. (2025). Citation counts and inclusion of references in seven free-access scholarly databases: A comparative analysis. *Journal of Informetrics*, 19(1), 101618. https://doi.org/10.1016/j.joi.2024.101618

Dorta-González, P., & Dorta-González, M. I. (2013). Comparing journals from different fields of science and social science through a JCR subject categories normalized impact factor. *Scientometrics*, 95(2), 645-672. https://doi.org/10.1007/s11192-012-0929-9

Dougherty, M. R., & Horne, Z. (2022). Citation counts and Journal Impact Factors do not capture some indicators of research quality in the behavioural and brain sciences. *Royal Society Open Science*, 9(8), 220334. https://doi.org/10.1098/rsos.220334

Drivas, K. (2025). The role of online search platforms in scientific diffusion. *Journal of the Association for Information Science and Technology*, 76(3), 580-603. https://doi.org/10.1002/asi.24959

Egghe, L. (2022). Impact measures: What are they? *Scientometrics*, 127(1), 385-406. https://doi.org/10.1007/s11192-021-04053-3

Finardi, U. (2014). On the time evolution of received citations, in different scientific fields: An empirical study. *Journal of Informetrics*, 8(1), 13-24. https://doi.org/10.1016/j.joi.2013.10.003

Fischer, I., & Steiger, H. J. (2018). Dynamics of Journal Impact Factors and limits to their inflation. *Journal of Scholarly Publishing*, 50(1), 26-36. https://doi.org/10.3138/jsp.50.1.06

Franceschet, M. (2010). The difference between popularity and prestige in the sciences and in the social sciences: A bibliometric analysis. *Journal of Informetrics*, 4(1), 55-63. https://doi.org/10.1016/j.joi.2009.08.001

Garfield, E. (2006). The history and meaning of the Journal Impact Factor. *JAMA*, 295(1), 90-93. https://doi.org/10.1001/

jama.295.1.90
Ghasemi, A., Mirmiran, P., Kashfi, K., & Bahadoran, Z. (2022). Scientific publishing in biomedicine: A brief history of scientific journals. *International Journal of Endocrinology and Metabolism*, 21(1), e131812. https://doi.org/10.5812/ijem-131812
Glasziou, P., Altman, D. G., Bossuyt, P., Boutron, I., Clarke, M., Julious, S., Michie, S., Moher, D., & Wager, E. (2014). Reducing waste from incomplete or unusable reports of biomedical research. *The Lancet*, 383(9913), 267-276. https://doi.org/10.1016/S0140-6736(13)62228-X
Gusenbauer, M., & Haddaway, N. R. (2020). Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, 11(2), 181-217. https://doi.org/10.1002/jrsm.1378
Haddawy, P., Hassan, S. U., Asghar, A., & Amin, S. (2016). A comprehensive examination of the relation of three citation-based journal metrics to expert judgment of journal quality. *Journal of Informetrics*, 10(1), 162-173. https://doi.org/10.1016/j.joi.2015.12.005
Haghdoost, A., Zare, M., & Bazrafshan, A. (2014). How variable are the journal impact measures? *Online Information Review*, 38(6), 723-737. https://doi.org/10.1108/oir-05-2014-0102
Hayashi, T., & Fujigaki, Y. (1999). Differences in knowledge production between disciplines based on analysis of paper styles and citation patterns. *Scientometrics*, 46(1), 73-86. https://doi.org/10.1007/BF02766296
Helmer, S., Blumenthal, D. B., & Paschen, K. (2020). What is meaningful research and how should we measure it? *Scientometrics*, 125(1), 153-169. https://doi.org/10.1007/s11192-020-03649-5
Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429-431. https://doi.org/10.1038/520429a
Hussain, S. (2015). Journal list fetishism and the 'sign of 4' in the ABS guide: A question of trust? *Organization*, 22(1), 119-138. https://doi.org/10.1177/1350508413506763
Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLOS Medicine*, 2(8), e124. https://doi.org/10.1371/journal.pmed.0020124
Ioannidis, J. P. A., Pezzullo, A. M., & Boccia, S. (2023). The rapid growth of mega-journals: Threats and opportunities. *JAMA*, 329(15), 1253-1254. https://doi.org/10.1001/jama.2023.3212
Ioannidis, J. P. A., & Thombs, B. D. (2019). A user's guide to inflated and manipulated impact factors. *European Journal of Clinical Investigation*, 49(9), e13151. https://doi.org/10.1111/eci.13151
Khademi, R., Serati Shirazi, M., & Haghpanah, S. A. (2023). Analyzing and mapping the scientific structure of Iran's biomechanics documents using the co-occurrence analysis. *Caspian Journal of Scientometrics*, 10(2), 15. https://doi.org/10.22088/cjs.10.2.15
Kiester, L., & Turp, C. (2022). Artificial intelligence behind the scenes: PubMed's best match algorithm. *Journal of the Medical Library Association*, 110(1), 15-22. https://doi.org/10.5195/jmla.2022.1236
Knudson, D. (2013). Impact and prestige of kinesiology-related journals. *Comprehensive Psychology*, 2. https://doi.org/10.2466/50.17.Cp.2.13
Knudson, D. (2014). Citation rates for highly-cited papers from different sub-disciplinary areas within kinesiology. *Chronicle of Kinesiology in Higher Education*, 25, 9-17.
Knudson, D. (2015a). Citation rate of highly-cited papers in 100 kinesiology-related journals. *Measurement in Physical Education and Exercise Science*, 19(1), 44-50. https://doi.org/10.1080/1091367X.2014.988336
Knudson, D. (2015b). Evidence of citation bias in kinesiology-related journals. The Chronicle of Kinesiology in Higher Education, 26(1), 5-12.
Knudson, D. (2017). Confidence crisis of results in biomechanics research. *Sports Biomechanics*, 16(4), 425-433. https://doi.org/10.1080/14763141.2016.1246603
Knudson, D. (2022). What kinesiology research is most visible to the academic world? *Quest*, 74(3), 285-298. https://doi.org/10.1080/00336297.2022.2092880
Knudson, D. (2023). Publication metrics and subject categories of biomechanics journals. *Journal of Information Science Theory and Practice*, 11(4), 40-50. https://doi.org/10.1633/JISTaP.2023.11.4.3
Knudson, D. (2024a). Are there meaningful prestige metrics of kinesiology-related journals? *Measurement in Physical Education and Exercise Science*, 28(4), 316-337. https://doi.org/10.1080/1091367X.2024.2341849
Knudson, D. (2024b). Center of mass change in the biomechanics citation network. *Sports Biomechanics*, 23(12), 3257-3267. https://doi.org/10.1080/14763141.2023.2287030
Knudson, D. (2024c). Scopus citation metrics for top and bottom quintile kinesiology-related journals. *International Journal of Kinesiology in Higher Education*, 8(4), 343-354. https://doi.org/10.1080/24711616.2024.2416184
Knudson, D., Cardinal, B. J., & McCullagh, P. (2024). Synthesis of publication metrics in kinesiology-related journals: Proxies for rigor, usage, and prestige. *Quest*, 76(1), 93-112. https://doi.org/10.1080/00336297.2023.2237150

Knudson, D., & Chow, J. W. (2008). North American perception of the prestige of biomechanics serials. *Gait & Posture*, 27(4), 559-563. https://doi.org/10.1016/j.gaitpost.2007.07.005

Knudson, D., & Quimby, S. (2023). Variation in recent citation metrics for biomechanics journals. *ISBS Proceedings Archive*, 41(1), Article 64.

Kurmis, A. P. (2003). Understanding the limitations of the Journal Impact Factor. *Journal of Bone and Joint Surgery*, 85(12), 2449-2454.

Leydesdorff, L., Bornmann, L., Comins, J. A., & Milojević, S. (2016). Citations: Indicators of quality? The impact fallacy. *Frontiers in Research Metrics and Analytics*, 1, Article 1. https://doi.org/10.3389/frma.2016.00001

Mahmood, K. (2017). Correlation between perception-based journal rankings and the Journal Impact Factor (JIF): A systematic review and meta-analysis. *Serials Review*, 43(2), 120-129. https://doi.org/10.1080/00987913.2017.1290483

Mason, S., & Singh, L. (2022). When a journal is both at the 'top' and the 'bottom': The illogicality of conflating citation-based metrics with quality. *Scientometrics*, 127(6), 3683-3694. https://doi.org/10.1007/s11192-022-04402-w

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics*, 106(1), 213-228. https://doi.org/10.1007/s11192-015-1765-5

Ogden, T. L., & Bartley, D. L. (2008). The ups and downs of Journal Impact Factors. *The Annals of Occupational Hygiene*, 52(2), 73-82. https://doi.org/10.1093/annhyg/men002

Osterloh, M., & Frey, B. S. (2020). How to avoid borrowed plumes in academia. *Research Policy*, 49(1), 103831. https://doi.org/10.1016/j.respol.2019.103831

Pajić, D. (2015). On the stability of citation-based journal rankings. *Journal of Informetrics*, 9(4), 990-1006. https://doi.org/10.1016/j.joi.2015.08.005

Pandita, R., & Singh, S. (2023). Growth and distribution of research journals across the world. *Journal of Information Science Theory and Practice*, 11(2), 22-36. https://doi.org/10.1633/JISTaP.2023.11.2.3

Patience, G. S., Patience, C. A., Blais, B., & Bertrand, F. (2017). Citation analysis of scientific categories. *Heliyon*, 3(5). https://doi.org/10.1016/j.heliyon.2017.e00300

Pozsgai, G., Lövei, G. L., Vasseur, L., Gurr, G., Batáry, P., Korponai, J., Littlewood, N. A., Liu, J., Móra, A., Obrycki, J., Reynolds, O., Stockan, J. A., VanVolkenburg, H., Zhang, J., Zhou, W., & You, M. (2021). Irreproducibility in searches of scientific literature: A comparative analysis. *Ecology and Evolution*, 11(21), 14658-14668. https://doi.org/10.1002/ece3.8154

Pranckutė, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications*, 9(1), 12. https://doi.org/10.3390/publications9010012

Roldan-Valadez, E., Salazar-Ruiz, S. Y., Ibarra-Contreras, R., & Rios, C. (2019). Current concepts on bibliometrics: A brief review about impact factor, Eigenfactor score, CiteScore, SCImago Journal Rank, Source-Normalised Impact per Paper, H-index, and alternative metrics. *Irish Journal of Medical Science (1971-)*, 188(3), 939-951. https://doi.org/10.1007/s11845-018-1936-5

Schubert, A., & Glänzel, W. (1983). Statistical reliability of comparisons based on the citation impact of scientific publications. *Scientometrics*, 5(1), 59-73. https://doi.org/10.1007/bf02097178

Seglen, P. O. (1997). Why the impact factor of journals should not be used for evaluating research. *BMJ*, 314(7079), 497. https://doi.org/10.1136/bmj.314.7079.497

Siler, K., Larivière, V., & Sugimoto, C. R. (2020). The diverse niches of megajournals: Specialism within generalism. *Journal of the Association for Information Science and Technology*, 71(7), 800-816. https://doi.org/10.1002/asi.24299

Smith, R. (1998). Unscientific practice flourishes in science. *British Medical Journal*, 316(7137), 1036. https://doi.org/10.1136/bmj.316.7137.1036

Starbuck, W. H. (2005). How much better are the most-prestigious journals? The statistics of academic publication. *Organization Science*, 16(2), 180-200. https://doi.org/10.1287/orsc.1040.0107

Tahamtan, I., & Bornmann, L. (2019). What do citation counts measure? An updated review of studies on citations in scientific documents published between 2006 and 2018. *Scientometrics*, 121(3), 1635-1684. https://doi.org/10.1007/s11192-019-03243-4

Thelwall, M., Kousha, K., Makita, M., Abdoli, M., Stuart, E., Wilson, P., & Levitt, J. (2023). In which fields do higher impact journals publish higher quality articles? *Scientometrics*, 128(7), 3915-3933. https://doi.org/10.1007/s11192-023-04735-0

Tourish, D., & Willmott, H. (2015). In defiance of folly: Journal rankings, mindless measures and the ABS Guide. *Critical Perspectives on Accounting*, 26, 37-46. https://doi.org/10.1016/j.cpa.2014.02.004

Vagenas, G., Palaiothodorou, D., & Knudson, D. (2018). Thirty-year trends of study design and statistics in applied sports and exercise biomechanics research. *International Journal of Exercise Science*, 11(1), 239-259. https://doi.org/10.70252/YWWI1479 van Leeuwen, T., & Tijssen, R. (2000). Interdisciplinary dynamics of modern science: analysis of cross-disciplinary citation flows. *Research Evaluation*, 9(3), 183-187. https://doi.org/10.3152/147154400781777241

Vîiu, G. A., & Păunescu, M. (2021). The lack of meaningful boundary differences between Journal Impact Factor quartiles undermines their independent use in research evaluation. *Scientometrics*, 126(2), 1495-1525. https://doi.org/10.1007/s11192-020-03801-1

Walters, W. H. (2017). Citation-based journal rankings: Key questions, metrics, and data sources. *IEEE Access*, 5, 22036-22053. https://doi.org/10.1109/ACCESS.2017.2761400

Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365-391. https://doi.org/10.1016/j.joi.2016.02.007

Wang, X., Li, D., Liu, X., & Wang, Z. (2025). Measuring knowledge flow in the interdisciplinary field of biosecurity: Full counting method or fractional counting method? *Scientometrics*, 130(2), 1101-1128. https://doi.org/10.1007/s11192-025-05249-7

Willmott, H. (2011). Journal list fetishism and the perversion of scholarship: Reactivity and the ABS list. *Organization*, 18(4), 429-442. https://doi.org/10.1177/1350508411403532

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I., Wouters, P., Hill, J., & Johnson, B. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. https://doi.org/10.13140/RG.2.1.4929.1363

Zadpoor, A. A., & Nikooyan, A. A. (2011). Publication and citation in biomechanics: A comparison with closely related fields (2003–2010). *Journal of Mechanics in Medicine and Biology*, 11(04), 705-711. https://doi.org/10.1142/S0219519411004836

Zhou, Y. B., Lü, L., & Li, M. (2012). Quantifying the influence of scientists and their publications: Distinguishing between prestige and popularity. *New Journal of Physics*, 14(3), 033033. https://doi.org/10.1088/1367-2630/14/3/033033

JISTaP https://www.jistap.org
Journal of Information Science Theory and Practice

RESEARCH PAPER
J Inf Sci Theory Pract 13(4): 69-82, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.4.5

From Intention to Impact: Predictors and Learning Outcomes of AI Tool Adoption Among Saudi University Students

Mohammed Alarefi*
Division of Business Administration, Department of Management, College of Business Administration, Prince Sultan University, Riyadh, Saudi Arabia
E-mail: malarefi@psu.edu.sa

Naif Nawaf Alharbi
Alhummam Logistics Solutions, Al-Qassim, Saudi Arabia
E-mail: naifnawaf09@gmail.com

ABSTRACT

This study examines the predictors and consequences of artificial intelligence (AI) tool adoption among university students in Saudi Arabia, integrating the technology acceptance model (TAM), unified theory of acceptance and use of technology (UTAUT), as well as Cognitive Social Theory with AI-specific trust and ethical concern. Using a cross-sectional survey of 317 students from four geographically distributed public universities, we employed partial least square-structural equation modeling to test ten hypotheses. The results reveal that perceived usefulness, trust in AI, and social influence significantly predict AI adoption, while perceived ease of use and ethical concerns show no significant effects. Self-efficacy partially mediates the relationships between perceived usefulness and perceived ease of use with adoption, but not for trust. Notably, AI adoption significantly enhances academic performance and cognitive skill development. The findings contribute to technology acceptance literature by validating the combination of TAM, UTAUT, and social cognitive theory in AI adoption contexts, highlighting trust as a critical factor beyond traditional TAM constructs, and demonstrating AI's positive educational outcomes. Practically, universities should emphasize AI's usefulness, foster trust through transparency, and integrate self-efficacy training to maximize adoption benefits.

Keywords: artificial intelligence adoption, technology acceptance model, academic performance, cognitive skills, Saudi Arabia

Received: July 4, 2025
Accepted: October 17, 2025
Revised: September 12, 2025
Published: December 30, 2025

***Corresponding Author:** Mohammed Alarefi
https://orcid.org/0009-0009-6749-5689
E-mail: malarefi@psu.edu.sa

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee or authors' permission. Open Access articles are automatically archived in the Korea Institute of Science and Technology Information (KISTI)'s Open Access repository (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

© 2025 [Mohammed Alarefi, Naif Nawaf Alharbi] This article is distributed under the Creative Commons Attribution License (CC BY), allowing unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. It is published by the Korea Institute of Science and Technology Information (KISTI).

1. INTRODUCTION

Recently, artificial intelligence (AI) in education has revolutionised teaching techniques and opened new possibilities for individualised and adaptable learning (Alarefi, 2024). Educational institutions are increasingly using AI-driven solutions such as intelligent tutoring systems, automated grading software, and data-informed learning analytics to boost student engagement, instructional effectiveness, and learning pathways (Cukurova, 2025). AI's potential benefits in education are well known, but the factors determining students' acceptance of these technologies, and their academic and cognitive outcomes are examined in limited studies. Understanding these elements is crucial to optimising AI in education while reducing risks like overreliance on automation and ethical problems including data privacy and algorithmic prejudice (Samala et al., 2024).

Previous well established models and theories such as The technology acceptance model (TAM) by Davis (1989) and the unified theory of acceptance and use of technology (UTAUT) by Venkatesh et al. (2003) show that perceived usefulness (PU), ease of use, and social influence strongly influence students' acceptance of digital learning tools. AI has drawbacks, including algorithmic decision-making, data security concerns, and potential consequences on critical thinking (Mouta et al., 2024). Overreliance on AI-driven suggestions may reduce students' problem-solving autonomy and creativity (Roll & Wylie, 2016), although other research show that real-time feedback and personalised information might increase learning efficiency (Chen et al., 2020). AI implementation's psychological and social effects on student motivation, classroom dynamics, and teacher-student relationships require further study.

Despite growing global interest in AI in education, most of the research has focused on Western and East Asian contexts, leaving many cultural and educational systems unprepared for new technologies (Assaf et al., 2024). This study fills this gap by examining student AI use and its implications in Saudi Arabia, which is undergoing rapid digital change as part of Vision 2030. As part of Vision 2030, Saudi Arabia is investing in AI and digital education to diversify the economy and build a knowledge-based economy (Alarefi, 2023a). The Saudi Ministry of Education has implemented intelligent learning platforms, AI-driven tutoring systems, and digital literacy programs (Alarefi, 2023b). These programs' success depends on students' willingness to use AI technologies and their impact on educational performance and cognitive skill development.

In the context of Saudi Arabia, several factors can contribute to the usage of AI, and these based on TAM and UTAUT can be related to the usefulness of the usage as well as the ease of use along with the influence of others, such as other students, teachers, and schools, as well as the community. Students' self-efficacy is a critical factor that can determine the usage (Ali & Allawi, 2024). However, there are limited studies that have examined the AI usage by students. In Saudi Arabia, despite strong Internet penetration and a tech-savvy youth (Alanzi et al., 2023), legacy educational methods and reluctance to change may limit AI-driven usage for educational purposes (Ghanem & Alshahrani, 2021). Ethical concerns regarding data privacy and AI decision-making may also affect students' trust in these technologies (AlMulhim, 2023). Understanding these contextual factors is essential for creating AI-driven educational tools that match local values and learning styles.

This research provides Saudi Arabian policymakers and educators with regional views on AI in education to improve the global discussion. This study will identify Saudi students' AI adoption motivations and assess its academic and cognitive effects, offering evidence-based recommendations for improving AI integration in Saudi classrooms. The findings contribute to ethical and pedagogical discussions about AI in education, ensuring that technology innovation supports sustainable and fair learning. The next sections discuss the literature review followed by the research methodology, findings, discussion, and conclusion.

2. LITERATURE REVIEW

2.1. Theoretical Framework

The theoretical framework of this study is based on three theories, TAM, UTAUT, and social cognitive theory by Bandura (1986). TAM, which was developed by Davis (1989), emphasizes PU and perceived ease of use (PEOU) as key variables in technology acceptance. These include learners' rational assessments of AI technology's academic advantages and the justification of the efforts required to use the technology. TAM acknowledges that adoption decisions entail psychological and social elements as well as value judgements. The present study also include the UTAUT model by Venkatesh et al. (2003), which is similar to TAM in focusing on usefulness (performance expectancy) and effort needed (effort expectancy). In addition, UTAUT adds social influence, which is the understanding that peer and instructor expectations drive educational adoption decisions (Simango et al., 2024). A robust theoretical framework for studying AI in education offers several benefits. It integrates AI-related challenges while maintaining strong psychological assumptions. The framework can explain adoption decisions and predict educational outcomes. It also offers practical advice for educators and policymakers using AI tools to maximise educational benefits while reducing students' fears and psychological barriers to adoption.

2.2. Artificial Intelligence Adoption

Significant academic study has examined the deployment of AI in education and its implications on learning outcomes. This study incorporates existing research to focus on Saudi Arabia's educational revolution's unique environment by examining technology adoption theory, AI-specific characteristics, and educational results. TAM, which emphasises PU and PEOU, has dominated research on educational technology adoption (Davis, 1989). Recent meta-analytic study confirms these variables' continuing importance, accounting for 40% of adoption intentions across educational contexts (Scherer et al., 2019). Integrating social cognitive theory, especially self-efficacy as a mediator, has improved the model. Research shows that students' technical competence confidence increases PU by 22-35% across instructional technologies (Taherdoost, 2018). In higher education, institutional support systems and peer networks influence up to 28% of adoption decisions (Al-Rahmi et al., 2021). These studies show that technology adoption is multifaceted, incorporating cognitive, emotional, and social factors.

AI technology's unique traits necessitate new adoption frameworks. Open algorithmic operations increase adoption likelihood by 31% compared to opaque systems, hence trust is important (Araujo et al., 2020). Since 68% of students worry about data safety in educational AI applications (Bond & Bedenlier, 2019), ethical issues are crucial. Collectivist cultures have 19% more social effect on adoption decisions than individualist ones (Nguyen et al., 2023). Adapting AI adoption frameworks to the technology's unique traits and cultural contexts is necessary.

Empirical study on AI and education shows promise and challenges. Well-deployed AI systems improve learning efficiency by 18-22% over time (Luckin et al., 2016). Meta-analytic evidence indicates particular benefits for personalized learning, with adaptive tutoring systems demonstrating effect sizes of 0.47 on standardized achievement measures (Chen et al., 2022). However, concerning findings have emerged regarding cognitive impacts, with experimental studies showing 12-15% declines in critical thinking metrics among frequent AI tool users (Penglong et al., 2024). Perhaps most significantly, research identifies substantial variation in outcomes based on students' self-efficacy levels, with high-efficacy students showing 27% greater learning gains compared to their low-efficacy counterparts when using the same AI systems (Alam, 2021). These differential effects underscore the importance of considering individual differences in AI adoption strategies.

Saudi Arabia's Vision 2030 initiative has positioned the nation as a particularly informative case study of large-scale educational AI implementation. Government investments exceeding $20 billion have created one of the world's most comprehensive smart education infrastructures (Abulibdeh et al., 2024). Early adoption studies in Saudi universities report promising integration rates (68%) for basic AI tools, though more advanced applications show substantially lower adoption (Abulibdeh et al., 2024). Qualitative research highlights the cultural specificity of adoption patterns, with Islamic values influencing perceptions of appropriate AI use in 43% of student respondents (AlMulhim, 2023). Despite these advancements, significant research gaps persist, particularly regarding: (1) the mediating mechanisms through which self-efficacy influences AI adoption, (2) the long-term cognitive consequences of educational AI use, and (3) the development of culturally appropriate implementation frameworks. The current study addresses these gaps through its integrated theoretical approach and empirical investigation of AI adoption in Saudi classrooms.

2.3. Conceptual Framework

This study uses known theoretical views and contemporary technological variables to rigorously examine the integration of AI technologies in higher education. The framework organises key aspects into three dimensions: antecedent causes, mediating processes, and educational results, providing a comprehensive perspective on AI integration in education. Based on TAM, PU is students' assessment of AI technologies' potential to increase learning and academic performance. These technologies' cognitive efforts and PEOU are strongly related. Based on UTAUT, the social influence element recognises the impact of institutional restrictions, teacher endorsements, and peer behaviours on technology adoption patterns. Two AI-specific components complete the antecedent dimensions:

trust in AI systems, which indicates algorithmic reliability and fairness, and ethical concerns, which encompass data protection and AI biases.

According to the social cognitive theory by Bandura (1986), self-efficacy is the main mediator of antecedent components into adoption behaviour. This mediation uses three interwoven channels. The competence route suggests that self-efficacy boosts students' confidence in their AI product use, which increases PU and uptake. Increased self-efficacy reduces technology-related uneasiness, reducing usability concerns. By increasing AI system readiness, self-efficacy boosts trust and adoption. Therefore, self-efficacy is proposed to mediate the paths of PU, PEOU, and trust in AI and adoption of AI.

The current study also proposed a direct effect between AI adoption and academic outcomes such as academic performance and cognitive skill development. Course grades and memory improve with academic performance. Cognitive skill development improves critical thinking, problem-solving, and creativity. AI integration in education has immediate (academic performance) and long-term consequences (cognitive skill development).

2.4. Perceived Usefulness and Artificial Intelligence Adoption

Based on TAM, PU is a critical factor for driving technology adoption (Davis, 1989). Scherer et al. (2019) found that PU explained 34% of learning technology adoption intentions in a meta-analysis of 161 studies. Alam (2021) found that students who thought AI writing tools improved essay quality were 2.3 times more likely to use them. In Middle Eastern universities, Al-Emran et al. (2020) found that perceived learning enhancement was the strongest predictor of educational technology adoption. Therefore, PU is anticipated to affect positively AI tool adoption.

H1: PU will increase AI tool adoption.

2.5. Perceived Ease of Use and Artificial Intelligence Adoption

Based on TAM and UTAUT, ease of use is an essential factor for promoting technology adoption. Empirical evidence supports the enduring relationship between ease of use and technology adoption, particularly for novel technologies. UTAUT research found ease of use effects were strongest during initial implementation phases (β=0.35) (Venkatesh et al., 2003). In AI-specific research, Araujo et al. (2020) demonstrated that reducing interface complexity increased adoption rates by 28% in their experimental study. These findings are particularly relevant given the technical nature of AI tools. A study by Chen et al. (2020) showed students rated ease of use as their primary concern (72% of respondents) when first encountering AI tutoring systems. Therefore, PEOU is expected to have a positive effect on AI adoption.

H2: PEOU will positively influence AI tool adoption.

2.6. Trust in Artificial Intelligence and Artificial Intelligence Adoption

Trust has emerged as a critical factor in AI adoption due to algorithmic opacity concerns. Warkentin et al. (2002)'s foundational work on technology trust found it explained an additional 18% of variance in adoption beyond TAM factors. In educational AI, Bond and Bedenlier (2019)'s multinational study showed trust was the second strongest predictor (β=0.33) after PU. Experimental research by Araujo et al. (2020) demonstrated that increasing algorithmic transparency boosted trust scores by 41% and subsequent adoption rates by 29%. Therefore, in this study, trust in AI is expected to affect positively the AI adoption.

H3: Trust in AI systems will positively predict adoption of AI tools.

2.7. Social Influence and Artificial Intelligence Adoption

The social dimension of technology adoption has shown particular strength in educational settings. Al-Rahmi et al. (2021)'s study of 783 university students found social influence accounted for 27% of variance in learning technology adoption. Bellaaj et al. (2015) reported social influence effects were 19% stronger than in individualist cultures. Recent AI-specific research by Alangari et al. (2022) in Saudi universities found peer recommendations increased AI tool trial rates by 43%, while instructor endorsements increased sustained usage by 31%.

H4: Social influence will be positively associated with AI tool adoption.

2.8. Ethical Concerns and Artificial Intelligence Adoption

Growing research highlights ethical considerations as adoption barriers for AI in education. A survey by Bond and Bedenlier (2019) found 68% of students expressed

 https://doi.org/10.1633/JISTaP.2025.13.4.5

significant data privacy concerns with educational AI. A study by Chiu et al. (2023) showed institutions with poor ethical transparency had 37% lower AI adoption rates. In Saudi contexts, AlMulhim (2023)'s qualitative research revealed Islamic ethical considerations influenced 43% of students' AI tool evaluations, often creating adoption hesitancy. In this study, ethical concern are expected to have a negative effect on AI adoption.

H5: Ethical concerns will negatively affect AI tool adoption.

2.9. Self-Efficacy As a Mediator

Social cognitive theory establishes self-efficacy as critical for converting beliefs into action. A meta-analysis study by Taherdoost (2018) found self-efficacy mediated 22-35% of PU effects across technology studies. In AI education research, the study of Alam (2021) showed that self-efficacy accounted for 28% of the total effect of PU on adoption. This mediation appears particularly strong for complex technologies (Roll & Wylie, 2016). Therefore, this study proposed that self-efficacy mediates the effect of PU on AI adoption.

H6: Self-efficacy will mediate the positive relationship between PU and AI tool adoption.

Research consistently demonstrates self-efficacy's role in overcoming usability challenges. Compeau and Higgins (1995)'s seminal work showed self-efficacy mediated 41% of ease of use effects. Recent AI studies reveal that this mediation is particularly crucial during initial adoption phases. Chen et al. (2020) found high self-efficacy students overcame early usability barriers 2.1 times faster than low self-efficacy peers. This aligns with the findings of Al-Emran et al. (2020) that self-efficacy mediated 33% of the ease-adoption relationship in Middle Eastern e-learning contexts. Accordingly, the following is proposed:

H7: Self-efficacy will mediate the positive relationship between PEOU and AI tool adoption.

Emerging research identifies self-efficacy as key for translating trust into adoption behaviour. A study by Tapalova and Zhiyenbayeva (2022) demonstrated students with high self-efficacy were 3.2 times more likely to act on their trust in AI systems. Path analysis by Alam (2021) revealed self-efficacy mediated 24% of trust's total effect on AI adoption. Research by Araujo et al. (2020) showed self-efficacy mediated the effect of trust on technology adoption. Accordingly, the following is proposed.

H8: Self-efficacy will mediate the positive relationship between trust in AI and AI tool adoption.

2.10. Artificial Intelligence Adoption and Educational Outcome

Empirical evidence links educational technology adoption to measurable learning gains. Luckin et al. (2016)'s meta-analysis found proper AI tool use increased learning efficiency by 18-22%. Specific to academic outcomes, Chen et al. (2020) reported a 0.47 effect size on course grades from adaptive tutoring systems. In Saudi contexts, the study of Abulibdeh et al. (2024) showed that AI-adopting classrooms outperformed controls by 11% on standardized exams. Scherer et al. (2019)'s research found adopting students demonstrated 27% higher assignment completion rates. Therefore, it is proposed that using AI will increase the academic performance of students.

H9: Adoption of AI tools will positively influence academic performance.

Research suggests AI tools can enhance higher-order thinking when properly implemented. Huang et al. (2021)'s taxonomy-informed studies show certain AI designs boost analysis and evaluation skills. However, effects depend on usage patterns. Roll and Wylie (2016) found students using AI as a complement for thinking showed 19% greater critical thinking gains. Recent work by Alam (2021) identified an optimal usage threshold (3-5 hours weekly) beyond which cognitive benefits plateaued or declined, suggesting nuanced relationships requiring further investigation. Thus, this study examines the effect of AI adoption on cognitive skill development and proposed a positive relationship.

H10: Adoption of AI tools will positively affect cognitive skill development.

3. RESEARCH METHODOLOGY

This quantitative study employed a cross-sectional survey design to examine the predictors and consequences of AI tool adoption among university students in Saudi Arabia. The target population comprised undergraduate students enrolled in public universities across Saudi Arabia. A multi-stage sampling approach was implemented to ensure representative geographical coverage and disciplinary diversity. First, four universities were purposively selected from different regions. These institutions were chosen based on their regional distribution, student population size, and established e-learning infrastructure. Within each university, stratified random sampling was employed to select participants proportionally across academic disciplines (humanities, health sciences, and business) and year of study (1st to 4th year). The sample size was determined using Cochran's formula for finite populations, targeting a 95% confidence level with a 5% margin of error. This approach yielded a required sample of 1,200 participants, with 300 students recruited from each university.

The study utilized a structured questionnaire developed through an extensive four-phase process. The questionnaire items were adapted from validated instruments in technology acceptance literature while incorporating AI-specific constructs. The survey comprised six sections. The first is the demographic information and it includes three questions regarding age, gender, and major. The second section include the variables of the study. PU was measured using six items, and PEOU was measured using five items adopted from Davis (1989). Social influence includes four items and was adopted from Venkatesh et al. (2003). For trust in AI, it was adopted from Lian (2015) while measurement of ethical concern was adopted from Bond and Bedenlier (2019). Further, the adoption of AI was adopted from Venkatesh et al. (2003). Measurement of self-efficacy was adopted from Alam (2021) while measurement of academic performance was adopted from Chen et al. (2022) and measurement of cognitive skill development (five items) was adopted from Kayali and Alaaraj (2020).

The second stage included the validation of the instrument. The initial English version was reviewed by a panel of five experts. The content validity index was calculated, with all scales exceeding the recommended threshold of 0.80. Experts provided feedback on item clarity, relevance, and cultural appropriateness. The third stage included translation of the questionnaire. The questionnaire underwent rigorous translation into Arabic following the forward-backward translation protocol. Two bilingual academics independently translated the instrument, followed by reconciliation of discrepancies. The fourth state was the pilot testing. The Arabic questionnaire was pilot tested with 40 students (10 from each university). Reliability analysis showed excellent internal consistency (CA ranging from 0.83 to 0.92 across scales). Minor adjustments were made to improve item clarity based on pilot feedback.

Data collection occurred over an 8-week period (between November 2024 and January 2025) through an online survey platform. The researcher coordinated with university administrators to distribute the survey through official channels. A link was sent to the students, and they were reminded weekly. The final sample included 357 completed responses, representing a 29.7% response rate. The sample characteristics mirrored the target population in terms of gender distribution, academic level, and discipline representation.

A comprehensive data cleaning protocol was implemented. Cases with more than 20% missing items (n=18) were excluded. For remaining missing values, multiple imputation using chained equations was employed. Little's missing completely at random (MCAR) test indicated the data was MCAR (χ^2=32.18, p=0.14). The outliers was examined, and this resulted in removing 22 responses. Multivariate outliers were identified using Mahalanobis distance (p<0.001). After verifying these were legitimate responses, extreme values were winsorized to the 95th percentile. All variables showed acceptable skewness (<|1.0|) and kurtosis (<|1.0|). Variance inflation factors ranged from 1.12 to 2.87, well below the threshold of 5, indicating no concerning multicollinearity among predictors. Comparison of early (n=192) and late respondents (n=125) revealed no significant differences in demographic characteristics or scale scores (all p>0.05), suggesting minimal non-response bias. In total 317 responses were valid and complete.

4. FINDINGS

4.1. Descriptive Information

The anticipated samples of 317 students from four Saudi universities are projected to show balanced demographic distribution, with 48% female representation and proportional distribution across academic years and majors. All the respondents were under 25 years of age, and this is in line with this study which investigates undergraduate students. Initial variable analysis suggested moderate-to-high scores for PU (mean [M]=3.82, standard deviation [SD]=0.73), PEOU (M=3.71, SD=0.81), and social influence (M=3.62, SD=0.81), with trust in AI systems showing more modest levels (M=3.33, SD=0.93). Ethical concerns demonstrated significant variability (M=3.51, SD=1.0), reflecting divergent student attitudes toward AI adoption. The adoption variable showed a mean score of 3.2 (M=3.22, SD=0.83), indicating most students engage

Table 1. Factor loading

Construct	Item code	Item statement	Source	Status	Factor loading
Perceived usefulness	PU1	Using AI improves my academic performance	Davis (1989)	Retained	0.75
	PU2	AI helps me complete academic tasks more efficiently		Retained	0.91
	PU3	AI tools enhance the quality of my academic work		Retained	0.85
	PU4	Using AI is beneficial for my learning outcomes		Retained	0.81
	PU5	AI usage contributes positively to my academic success		Retained	0.74
	PU6	AI tools are a valuable asset in my education		Retained	0.79
Perceived ease of use	PEOU1	AI tools are easy to use	Davis (1989)	Retained	0.73
	PEOU2	Learning to use AI tools is easy for me		Retained	0.88
	PEOU3	I find AI tools to be user-friendly		Retained	0.81
	PEOU4	Interacting with AI tools does not require a lot of effort		Retained	0.84
	PEOU5	It is easy for me to become skilful at using AI tools		Retained	0.76
Social influence	SI1	People important to me think I should use AI tools	Venkatesh et al. (2003)	Retained	0.91
	SI2	Instructors expect me to use AI tools in learning		Retained	0.87
	SI3	My peers think using AI tools is beneficial		Retained	0.71
	SI4	I feel social pressure to use AI tools		Retained	0.72
Trust in AI	TAI1	I trust the decisions made by AI tools in academic settings	Lian (2015)	Retained	0.73
	TAI2	AI systems used in education are reliable		Retained	0.73
	TAI3	I believe AI tools make fair academic suggestions		Retained	0.79
	TAI4	I am confident in the accuracy of AI tools		Retained	0.77
Ethical concerns	EC1	I worry AI might misuse my data	Bond and Bedenlier (2019)	Deleted	0.53
	EC2	AI tools might breach student privacy		Retained	0.82
	EC3	AI in education may cause ethical issues		Retained	0.79
	EC4	The use of AI might result in unfair grading		Retained	0.73
Self-efficacy	SE1	I am confident in my ability to use AI tools effectively	Alam (2021)	Retained	0.75
	SE2	I can complete tasks using AI without help		Retained	0.77
	SE3	I feel confident troubleshooting AI tools when issues arise		Deleted	0.56
	SE4	I am capable of learning new AI features		Retained	0.78
	SE5	I feel well-prepared to use AI tools in my studies		Retained	0.79
Academic performance	AP1	My academic results have improved since using AI tools	Chen et al. (2022)	Retained	0.81
	AP2	AI tools help me complete academic assignments more effectively		Retained	0.76
	AP3	I score better on tests when I use AI tools		Retained	0.81
	AP4	I meet academic deadlines better with the help of AI tools		Retained	0.77

Table 1. Continued

Construct	Item code	Item statement	Source	Status	Factor loading
Cognitive skill development	CSD1	AI tools help improve my problem-solving skills	Kayali & Alaaraj (2020)	Retained	0.79
	CSD2	I use AI tools to enhance my critical thinking		Retained	0.91
	CSD3	AI helps enhance my analytical and critical thinking skills		Deleted	0.51
	CSD4	I use AI to explore complex ideas in my field		Retained	0.87
	CSD5	AI contributes to my creative thinking in academic work		Retained	0.73
Adoption of AI	AI1	I regularly use AI tools for my studies	Venkatesh et al. (2003)	Retained	0.87
	AI2	I integrate AI tools into my academic routines		Retained	0.91
	AI3	I prefer to use AI tools when available		Retained	0.88
	AI4	AI tools are a regular part of my learning process		Retained	0.89
	AI5	I rely on AI tools to complete academic tasks		Retained	0.86

AI, artificial intelligence.

with AI tools occasionally rather than extensively. In addition, the self-efficacy showed a moderate level of 3.1 (M=3.1, SD=0.5). Similarly, the educational outcome such as academic performance showed a moderate level of 3.11 (M=3.11, SD=0.31) and 3.19 for cognitive skills development (M=3.19, SD=0.53).

4.2. Measurement Model

Based on the recommendation of Hair et al. (2023), the first stage in evaluating the measurement model is to assess the factor loading. All the factor loading showed acceptable levels except for item 1 from ethical concerns, and item 3 from cognitive skill development, as well as item 3 from self-efficacy. As shown in Table 1 (Alam, 2021; Bond & Bedenlier, 2019; Chen et al., 2022; Davis, 1989; Kayali & Alaaraj, 2020; Lian, 2015; Venkatesh et al., 2003), the factor loading of all items including the deleted items is given.

As shown in Table 2, measurement model results confirm the robustness of the theoretical constructs. All scales demonstrate strong reliability by having Cronbach's alpha (CA) and composite reliability above 0.70 with particularly good performance from the self-efficacy measures. The convergent validity was achieved because the value of average variance extracted is above 0.50, as shown in Table 2.

Table 2. Reliability and validity

Variable	Cronbach's alpha	Composite reliability	Average variance extracted
Academic performance	0.840	0.874	0.585
Adoption of AI	0.859	0.860	0.842
Cognitive skill development	0.851	0.858	0.841
Ethical concern	0.731	0.747	0.531
Perceived ease of use	0.851	0.852	0.872
Perceived usefulness	0.838	0.841	0.802
Self-efficacy	0.926	0.935	0.599
Social Influence	0.916	0.920	0.749
Trust in AI	0.883	0.883	0.881

AI, artificial intelligence.

The discriminant validity was assessed using heterotrait-monotrait ratio of correlations (HTMT)'s correlation in which the correlation among variables was less than 0.85, indicating that there is no issues of discriminant validity. These psychometric results give confidence in the quality of the measurement approach and the meaningfulness of subsequent hypothesis tests.

4.3. Structural Model

In accordance with recent partial least square-structural equation modeling (PLS-SEM) evaluation standards (Hair et al., 2023), the model fit was assessed using several global fit indices. The standardized root mean square residual value was 0.056, which is below the recommended threshold of 0.08, indicating a good overall model fit. The normed fit index was 0.91, which also meets the minimum criterion of 0.90, further supporting model adequacy. Additionally, the discrepancy measures d_ULS and d_G were 0.225 and 0.193, respectively, and fall within acceptable limits for PLS-SEM analysis, indicating no significant model misfit. Although Chi-square (χ^2=1,985.56) is reported, it is primarily descriptive in the context of PLS-SEM and not interpreted in the same way as in covariance based-structural equation modeling. These indicators collectively demonstrate that the structural model exhibits an acceptable fit with the empirical data, supporting the robustness of the hypothesized relationships. Table 3 (Hair et al., 2023) shows the results of testing the model fit for the structural model.

The structural model was assessed by examining the R-square (R2) and F-square (F2) as well as the path coefficient. As can be seen in Table 4, the R-square of AI adoption is 0.695, indicating that the five predictors can explain 69.5% of the variation in AI adoption. AI adoption can explain 29% of academic performance and 24.2% of cognitive skill development. After including the mediator self-efficacy, the R-square of self-efficacy as shown in Fig. 1 accounted to 63.4%. The R-square of the adoption of AI increased to 70.9% while the academic performance remained the same at 29% and increased to 24.3% for the cognitive skill development. The F-square for all significant paths are above 0.02 while for non-significant paths, they are less than 0.02.

Based on the direct effect model and the mediating effect model, Table 4 shows the results of testing the hypotheses. H is the hypothesis, B is the coefficient, Std. is the standard deviation, T is the T-value, *p* is the *p*-value, F2 is the effect size or F-square, and R2 is the R-square.

Our analysis supports seven of the ten hypotheses. PU has a strong positive effect on AI adoption (B=0.308, $p<0.05$) indicating that H1 is supported. The second hypothesis is rejected (H2), which is related to the effect of PEOU on adoption of AI. Trust in AI has a strong effect on the adoption of AI (B=0.298, $p<0.05$). Thus, H3 is sup-

Table 3. Model fit indices for the structural model

Model fit index	Value	Recommended threshold	Reference	Interpretation
Standardized root mean square residual	0.056	<0.08	Hair et al. (2023)	Good model fit
Normed fit index	0.91	>0.90	Hair et al. (2023)	Acceptable fit
d_ULS (unweighted least squares discrepancy)	0.225	As low as possible	Hair et al. (2023)	Acceptable
d_G (geodesic discrepancy)	0.193	As low as possible	Hair et al. (2023)	Acceptable
Chi-square (χ^2)	1,985.56	Not significant preferred	Hair et al. (2023)	Descriptive only in PLS-SEM

PLS-SEM, partial least square-structural equation modeling.

Table 4. Result of hypotheses testing

Hypotheses	Path	B	Std.	T	*p*-value	F2	R2
H1	PU→Adoption of AI	0.308	0.049	6.276	<0.05	0.12	0.695
H2	PEOU→Adoption of AI	0.062	0.050	1.249	0.212	0.01	
H3	Trust in AI→Adoption of AI	0.298	0.066	4.508	<0.05	0.11	
H4	Social influence→Adoption of AI	0.287	0.062	4.588	<0.05	0.10	
H5	Ethical concern→Adoption of AI	0.014	0.030	0.460	0.645	0.00	
H6	PU→Self-efficacy→Adoption of AI	0.088	0.025	3.488	<0.001	0.03	0.709
H7	PEOU→Self-efficacy→Adoption of AI	0.057	0.020	2.895	0.004	0.02	
H8	Trust in AI→Self-efficacy→Adoption of AI	0.035	0.018	1.931	0.054	0.01	
H9	Adoption of AI→Academic performance	0.539	0.028	19.482	<0.05	0.21	0.290
H10	Adoption of AI→Cognitive skill development	0.492	0.052	9.548	<0.05	0.19	0.242

B, coefficient; Std., standard deviation; T, T-value; F2, F-square; R2, R-square.

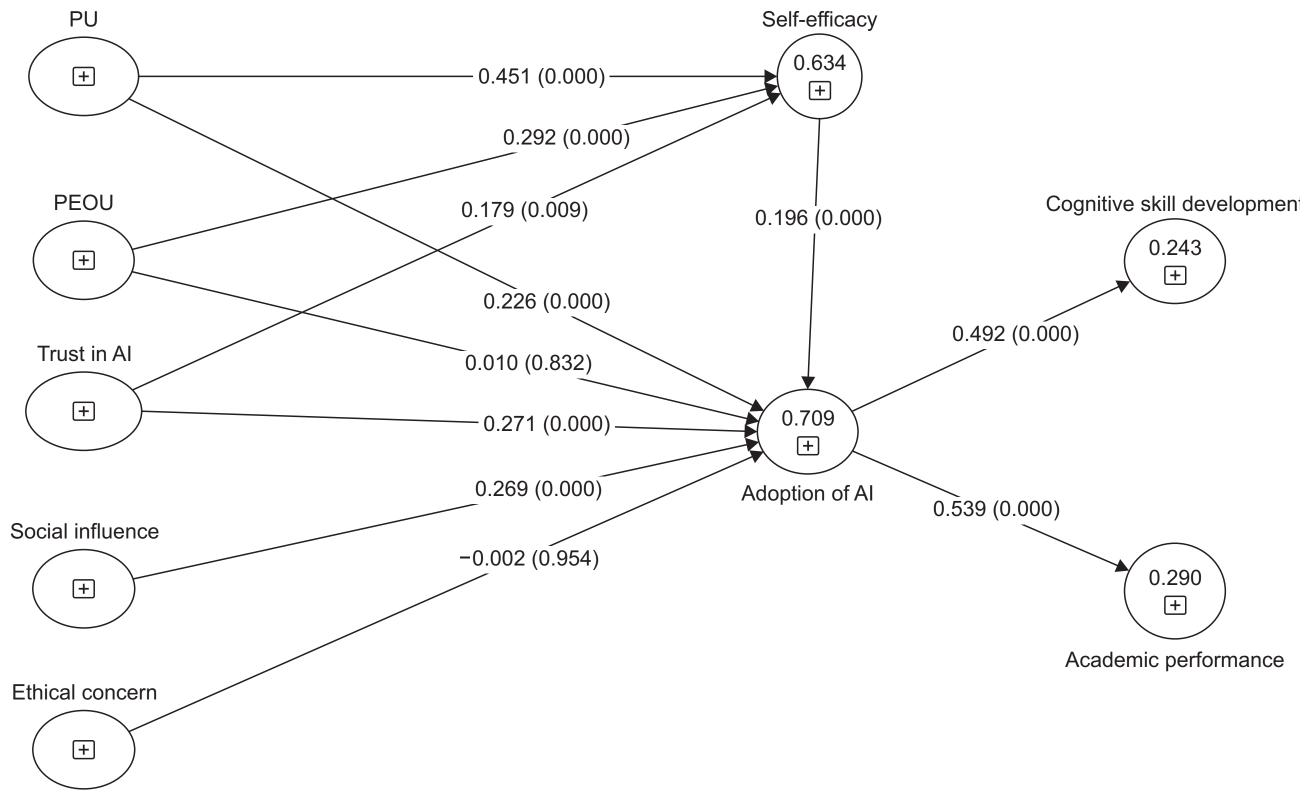

Fig. 1. Structural model. PU, perceived usefulness; PEOU, perceived ease of use; AI, artificial intelligence.

ported. For H4, it is supported because the social influence significantly predicted the adoption of AI (B=0.287, $p<0.05$). For H5, it is rejected. Ethical concerns have no effect on adoption of AI. The non-significant effect of ethical concerns suggests students may compartmentalize these worries when making practical adoption decisions, prioritizing immediate academic benefits over longer-term ethical considerations.

In terms of the mediation role of self-efficacy, it was confirmed with PU and PEOU but not with trust in AI. This indicates that part of the relationship between PU and PEOU can be explained by self-efficacy while the self-efficacy has no role between trust in AI and adoption of AI. The mediation effects of self-efficacy are confirmed for PU and ease of use, highlighting students' confidence as a key mechanism translating beliefs into adoption behaviour. Therefore, H6 and H7 are supported while H8 is rejected. The lack of mediation for trust through self-efficacy indicates these constructs operate independently in this sample. Students' trust in AI systems appears unrelated to their confidence in using them. For the effect of AI adoption on educational outcome, it has a significant effect on academic performance (B=0.539, $p<0.05$) and cognitive skill development (B=0.492, $p<0.05$). Therefore, H9 and H10 are supported.

5. DISCUSSION

This study's findings offer important theoretical contributions to technology acceptance literature in several ways. The supported hypotheses generally validate the applicability of TAM in AI adoption contexts within Saudi higher education. The strong predictive power of PU aligns with Davis (1989)'s original propositions, suggesting these core TAM constructs remain relevant for emerging educational technologies. In addition, social influence as a significant predictor coordinates with conventional wisdom from UTAUT (Venkatesh et al., 2003). This finding suggests that social norms may operate similarly in Saudi academic environments compared to Western contexts where UTAUT was developed.

The rejection of hypotheses related to PEOU and ethical concerns warrants deeper theoretical reflection. The non-significant effect of PEOU suggests that, in the Saudi higher education context, students may already possess sufficient digital literacy and exposure to technology, which reduces the salience of usability issues. This aligns with findings from studies in digitally advanced societies, where once baseline technological competence is achieved, ease of use becomes less critical than PU or social endorsement in predicting adoption. Regarding ethical concerns, the absence of significant influence may reflect a combination of cultural and contextual factors. Students may prioritize the immediate academic benefits of AI tools over abstract ethical risks, particularly when trust in state institutions and regulatory frameworks is high. Additionally, cultural and religious values, such as reliance on institutional authority and Islamic ethical orientations, may mitigate individual apprehensions about privacy and algorithmic fairness. These insights suggest that while global AI ethics debates are prominent, in practice, students in this context tend to compartmentalize such concerns, giving precedence to pragmatic educational gains.

The dismissal of ethical concerns by students, despite the prominence of global debates on AI ethics, requires a nuanced interpretation. In the Saudi context, several interrelated factors may account for this pattern. First, cultural norms and the collectivist orientation of society may lead students to defer to institutional and governmental oversight, assuming that AI systems deployed in universities are already aligned with ethical and legal standards. Second, religious and moral frameworks, particularly Islamic digital ethics, emphasize trust in authority and compliance with institutional rules, which may reduce the salience of individual-level ethical questioning. Third, the relatively low awareness of technical risks, such as data privacy breaches or algorithmic bias, may limit students' ability to critically evaluate AI systems. Finally, the pragmatic orientation of students, who are more concerned with immediate academic benefits, further explains why ethical issues were deprioritized. These findings suggest that ethical awareness campaigns, integrated into digital literacy and AI literacy programs, could help students better engage with the ethical dimensions of AI in education while maintaining trust in institutional safeguards.

The significant mediating role of self-efficacy (H6, H7) extends Bandura (1986)'s social cognitive theory by demonstrating its relevance for AI technology adoption. However, the lack of mediation for trust (H8) suggests that confidence in AI systems may function differently than trust in human instructors or conventional technologies, highlighting a need for theoretical refinement regarding trust mechanisms in AI-enabled learning environments. The non-significant effect of ethical concerns contradicts much contemporary AI ethics literature, possibly indicating a disconnect between abstract ethical considerations and practical adoption decisions among students.

6. IMPLICATIONS

This study improves theoretical understanding of educational technology adoption. It initially strengthens TAM by confirming its core concepts and defining key AI adoption boundaries. The strong predictive power of PU supports TAM, proving its relevance for advanced technology such as AI. The non-significant effect of PEOU challenges the TAM assumptions, suggesting that adoption decisions may be less influenced by ease of use as users gain technical proficiency.

The present research improves theoretical understanding by including AI trust as a major predictor, surpassing traditional TAM components. This supports earlier work calling for models that account for AI systems' unique traits, notably their opaque decision-making processes. Human-AI interactions require innovative theoretical frameworks to explain trust formation since users must rely on system outputs without understanding the algorithms. The significant impact of social influence supports UTAUT theory in collectivist education. This finding suggests that social norms may similarly affect technology adoption in numerous cultures, even though reference groups (e.g., peers versus instructors) and their proportionate effect may differ. Ethics' limited impact weakens assumptions in AI ethics literature, demonstrating a disjunction between theoretical ethical norms and implementation choices that requires additional theoretical research.

Beyond validating established frameworks such as TAM, UTAUT, and social cognitive theory, this study highlights the need for incorporating additional constructs that capture the unique challenges of AI adoption in educational contexts. Future research could integrate variables such as techno-stress, which reflects the psychological strain arising from overexposure to AI-driven platforms, and AI literacy, which measures students' ability to critically understand and evaluate AI systems. Both constructs are increasingly relevant as universities adopt advanced digital infrastructures. Similarly, resistance to change may offer explanatory power in contexts where traditional learning methods remain dominant and cultural norms shape technology acceptance. Embedding these constructs into future models would not only extend theoretical novelty but also create a culturally adaptive framework tailored to Middle East and North Africa (MENA) education systems, thereby advancing the literature beyond the boundaries of classical adoption theories.

The findings reveal many AI adoption measures for institutions. Institutions must emphasise academic benefits while using AI technology since PU affects acceptance. This may include discipline-specific AI application presentations, peer demonstration programs where early adopters discuss their discoveries, and practical demonstrations of how AI tools may improve learning. AI systems need trust, therefore universities must build it. Transparency standards for institutional AI tools, clear descriptions of AI system capabilities and constraints, and outlets for student algorithmic concerns are practical ways. These methods may help pupils trust AI in education. The partial mediation of self-efficacy underlines the necessity for AI technology instruction to increase student confidence. Institutions should offer incremental skill-building programs like practical workshops and supervised practice instead of only technical instruction. Customised programs must focus on how AI may improve academic learning. Although ethical concerns did not significantly affect adoption decisions, student answers show that institutions should not ignore ethical education. Institutions should incorporate AI ethics into digital literacy courses, create academic AI use norms, and host ethical discussions on AI in education. AI adoption positively impacts academic achievement and cognitive skill development, suggesting that institutions should proactively integrate AI into curriculum to optimise learning. This may entail revising exams to reflect AI-enhanced skills, giving faculty development on AI pedagogy, and helping teachers integrate AI technology to maximise educational advantages.

7. CONCLUSION

This study advances our understanding of AI adoption in higher education by validating core technology acceptance principles, while identifying important cultural specificities in the Saudi context. The findings suggest that while traditional adoption factors remain relevant, the unique characteristics of AI technologies and local educational cultures require theoretical refinement and tailored implementation approaches. By addressing the identified limitations in future research, scholars can develop more comprehensive models of educational technology adoption in the AI era. For practitioners, the results highlight both opportunities and challenges in leveraging AI tools to enhance teaching and learning outcomes. The findings of this study have several limitations. The focus on the Saudi context limits the findings to Saudi Arabia and undergraduate students.

Another key limitation of this study lies in the measurement of academic performance and cognitive skills, which relied exclusively on self-reported data. While self-assessments provide valuable insights into students' perceptions, they are inherently subjective and may be influenced by social desirability bias or limited self-awareness. The absence of objective performance indicators, such as grade point average records, standardized test scores, or instructor evaluations, reduces the robustness of outcome assessment. Future research should aim to triangulate self-reports with objective academic data and faculty assessments, thereby enhancing measurement validity and allowing for a more comprehensive evaluation of the educational impact of AI adoption.

A further limitation concerns the scope of the sample, which was restricted to undergraduate students in public universities. While this focus ensured a clear and context-specific understanding of AI adoption in Saudi higher education, it limits the generalizability of the findings. Patterns of adoption may differ in private universities, where resources and institutional policies may be more flexible, or among graduate students, who often engage with AI in more specialized and research-intensive ways. Moreover, the exclusive national focus does not capture potential cross-cultural variations across MENA or other global contexts. Future studies should therefore broaden the sampling frame to include private institutions, graduate cohorts, and cross-national comparisons in order to validate and extend the present findings.

Future studies should examine whether these patterns hold in different national contexts. This study adopted a cross-sectional design. Therefore, it cannot capture how adoption patterns evolve over time. Further studies are recommended to conduct longitudinal research because it could reveal important developmental trajectories in AI tool usage. In term of the measurements of the variables, academic performance can be measured objectively or subjectively. Similarly, measures of cognitive skill development relied on self-reports rather than objective assessments. Future work should incorporate standardized academic and cognitive tests. The focus on public universities may not reflect patterns at private institutions. Additionally, graduate students were not included in the sample. Future research should investigate why ease of use is less impactful in this context compared to Western studies. In addition, future studies can examine the disconnect between ethical concerns and adoption behaviour. Furthermore, future studies should explore instructor perspectives on AI adoption to complement student data. Additional variables can be added as moderators or mediators. These include the AI literacy, perceived risk, technostress, and IT knowledge of students and governmental policies to support the adoption of AI and its impact on educational outcomes.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

Abulibdeh, A., Zaidan, E., & Abulibdeh, R. (2024). Navigating the confluence of artificial intelligence and education for sustainable development in the era of industry 4.0: Challenges, opportunities, and ethical dimensions. *Journal of Cleaner Production*, 437, 140527. https://doi.org/10.1016/j.jclepro.2023.140527

Al-Emran, M., Arpaci, I., & Salloum, S. A. (2020). An empirical examination of continuous intention to use m-learning: An integrated model. *Education and Information Technologies*, 25(4), 2899-2918. https://doi.org/10.1007/s10639-019-10094-2

Al-Rahmi, A. M., Al-Rahmi, W. M., Alturki, U., Aldraiweesh, A., Almutairy, S., & Al-Adwan, A. S. (2021). Exploring the factors affecting mobile learning for sustainability in higher education. *Sustainability*, 13(14), 7893. https://doi.org/10.3390/su13147893

Alam, A. (2021). *Possibilities and apprehensions in the landscape of artificial intelligence in education*. Paper presented at 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA), Nagpur, India. https://doi.org/10.1109/ICCICA52458.2021.9697272

Alangari, S., Alshahrani, S. M., Khan, N. A., Alghamdi, A. A., Almalki, J., & Al Shehri, W. (2022). Developing a blockchain-based digitally secured model for the educational sector in Saudi Arabia toward digital transformation. *PeerJ Computer Science*, 8, e1120. https://doi.org/10.7717/peerj-cs.1120

Alanzi, T., Alhajri, A., Almulhim, S., Alharbi, S., Alfaifi, S., Almarhoun, E., Mulla, R., Alasafra, Z. O., Alalwan, Z., Alnasser, F., Almukhtar, F., Al Ghadeer, F., Amro, S., Alodhayb, I., & Alanzi, N. (2023). Artificial intelligence and patient autonomy in obesity treatment decisions: An empirical study of the challenges. *Cureus*, 15(11), e49725. https://doi.org/10.7759/cureus.49725

Alarefi, M. (2023a). Adoption of IoT by telecommunication companies in GCC: The role of blockchain. *Decision Science Letters*, 12(1), 55-68. https://doi.org/10.5267/j.dsl.2022.10.006

Alarefi, M. (2023b). Cloud computing usage by governmental organizations in Saudi Arabia based on Vision 2030. *Uncertain Supply Chain Management*, 11(1), 169-178. https://doi.org/10.5267/j.uscm.2022.10.010

Alarefi, M. (2024). The impact of artificial intelligence on business performance in Saudi Arabia: The role of technological readiness and data quality. *Engineering, Technology & Applied Science Research*, 14(5), 16802-16807. https://doi.org/10.48084/etasr.7871

Ali, M. I., & Allawi, A. A. (2024). An artificial neural network prediction model of GFRP residual tensile strength. *Engineering, Technology & Applied Science Research*, 14(6), 18277-18282. https://doi.org/10.48084/etasr.9107

AlMulhim, A. F. (2023). The impact of administrative management and information technology on e-government success: The mediating role of knowledge management practices. *Cogent Business & Management*, 10(1), 2202030. https://doi.org/10.1080/23311975.2023.2202030

Araujo, T., Helberger, N., Kruikemeier, S., & de Vreese, C. H. (2020). In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI & SOCIETY*, 35(3), 611-623. https://doi.org/10.1007/s00146-019-00931-w

Assaf, R., Omar, M., Saleh, Y., Attar, H., Alaqra, N. T., & Kanan, M. (2024). Assessing the acceptance for implementing artificial intelligence technologies in the governmental sector: An empirical study. *Engineering, Technology & Applied Science Research*, 14(6), 18160-18170. https://doi.org/10.48084/etasr.8711

Bandura, A. (1986). *Social foundations of thought and action : A social cognitive theory*. Prentice-Hall.

Bellaaj, M., Zekri, D. I., & Albugami, D. M. (2015). The continued use of e-learning system : An empirical investigation using UTAUT model at the University of Tabuk. *Journal of Theoretical and Applied Information Technology,* 72(3), 464-474.

Bond, M., & Bedenlier, S. (2019). Facilitating student engagement through educational technology: Towards a conceptual framework. *Journal of Interactive Media in Education*, 1(11), 1-14. https://doi.org/10.5334/jime.528

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264-75278. https://doi.org/10.1109/ACCESS.2020.2988510

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology & Society*, 25(1), 28-47.

Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. https://doi.org/10.1016/j.caeai.2022.100118

Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. *MIS Quarterly,* 19(2), 189-211. https://doi.org/10.2307/249688

Cukurova, M. (2025). The interplay of learning, analytics and artificial intelligence in education: A vision for hybrid intelligence. *British Journal of Educational Technology*, 56(2), 469-488. https://doi.org/10.1111/bjet.13514

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. https://doi.org/10.2307/249008

Ghanem, S., & Alshahrani, S. (2021). The effect of cloud computing adoption on organizational performance of SMEs in Saudi Arabia. *International Journal of Contemporary Management and Information Technology*, 1(2), 1-6.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. (2023). *Advanced issues in partial least squares structural equation modeling* (2nd ed.). SAGE.

Huang, J., Saleh, S., & Liu, Y. (2021). A review on artificial intelligence in education. *Academic Journal of Interdisciplinary Studies*, 10(3), 206. https://doi.org/10.36941/ajis-2021-0077

Kayali, M., & Alaaraj, S. (2020). Adoption of cloud based e-learning in developing countries: A combination of DOI, TAM and UTAUT. *International Journal of Contemporary Management and Information Technology,* 1(1), 1-7.

Lian, J. W. (2015). Critical factors for cloud based e-invoice service adoption in Taiwan: An empirical study. *International Journal of Information Management*, 35(1), 98-109. https://doi.org/10.1016/j.ijinfomgt.2014.10.005

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2024). Uncovering blind spots in education ethics: Insights from a systematic literature review on artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 34(3), 1166-1205. https://doi.org/10.1007/s40593-023-00384-9

Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, 28(4), 4221-4241. https://doi.org/10.1007/s10639-022-11316-w

Penglong, G., Mustapa, F. D., Xue, Z., & Yangyang, Z. (2024). Impact of artificial intelligence usage and technology competence on competitive advantage with mediating role of effective information management system. *Profesional de la información*, 33(5), e330501. https://doi.org/10.3145/epi.2024.ene.0501

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 26(2), 582-599. https://doi.org/10.1007/s40593-016-0110-3

Samala, A. D., Rawas, S., Wang, T., Reed, J. M., Kim, J., Howard, N. J., & Ertz, M. (2024). Unveiling the landscape of generative artificial intelligence in education: A comprehensive taxonomy of applications, challenges, and future prospects. *Education and Information Technologies*, 30(3), 3239–3278. https://doi.org/10.1007/s10639-024-12936-0

Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13-35. https://doi.org/10.1016/j.compedu.2018.09.009

Simango, A., Mananze, S., & Bila, J. (2024). Application of artificial intelligence in the identification of Banana Bunch Top Virus (BBTV) in Mozambique. *Engineering, Technology & Applied Science Research*, 14(6), 18734-18740. https://doi.org/10.48084/etasr.7442

Taherdoost, H. (2018). A review of technology acceptance and adoption models and theories. *Procedia Manufacturing*, 22, 960-967. https://doi.org/10.1016/j.promfg.2018.03.137

Tapalova, O., & Zhiyenbayeva, N. (2022). Artificial intelligence in education: AIEd for personalised learning pathways. *Electronic Journal of e-Learning*, 20(5), 639-653. https://doi.org/10.34190/ejel.20.5.2597

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478.

Warkentin, M., Gefen, D., Pavlou, P. A., & Rose, G. M. (2002). Encouraging citizen adoption of e-government by building trust. *Electronic Markets*, 12(3), 157-162. https://doi.org/10.1080/101967802320245929

JISTaP https://www.jistap.org
Journal of Information Science Theory and Practice

RESEARCH PAPER
J Inf Sci Theory Pract 13(4): 83-99, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.4.6

When Does Data Integration Enhance Predictive Performance? An Empirical Analysis of Open Government Data

Junyoung Jeong*
National Information Society Agency, Daegu, Korea
E-mail: j.y.jeong@skku.edu

ABSTRACT

While open government data (OGD) is increasingly recognized as a critical resource for economic growth and data-driven innovation, methods for proactively evaluating the potential utilization of these datasets remain underdeveloped. This study addresses this gap by investigating two key methodological questions: first, whether automated machine learning (AutoML) is an appropriate tool for measuring and evaluating OGD utilization, and second, how the composition of training data affects the performance of models designed to predict such utilization. This research specifically compares the efficacy of two distinct data strategies: models trained on integrated datasets spanning multiple domains versus those trained on domain-specific datasets. Using metadata from the South Korean government's extensive OGD portal, this study employs AutoML to systematically build and evaluate predictive models under these different training conditions. The findings reveal that the training data strategy is a critical determinant of predictive accuracy, with the integrated-domain approach frequently yielding superior performance over domain-specific models. This research provides empirical evidence on the impact of data integration strategies in this context and establishes a methodological framework for the prospective assessment of OGD value, offering a more robust alternative to traditional retrospective evaluation metrics.

Keywords: data integration strategy, prediction model, predictive performance, automated machine learning, open government data, open government data utilization

Received: July 10, 2025 **Revised:** October 20, 2025
Accepted: October 31, 2025 **Published:** December 30, 2025

***Corresponding Author:** Junyoung Jeong
https://orcid.org/0009-0005-2147-3290
E-mail: j.y.jeong@skku.edu

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee or authors' permission. Open Access articles are automatically archived in the Korea Institute of Science and Technology Information (KISTI)'s Open Access repository (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

© 2025 [Junyoung Jeong] This article is distributed under the Creative Commons Attribution License (CC BY), allowing unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. It is published by the Korea Institute of Science and Technology Information (KISTI).

1. INTRODUCTION

Since the introduction of OpenAI's ChatGPT service, open government data (OGD) has increasingly been recognized as a critical resource driving digital transformation, particularly as this transformation accelerates across various sectors of society (Mhlanga, 2023). Within this context, OGD is frequently highlighted as a sustainable resource for value creation, with the potential to significantly contribute to both social and economic advancement (Hamari et al., 2016). Moreover, in the realm of corporate sustainability and environment, social, governance (ESG) management, Helbig (2021) posits that leveraging OGD offers substantial strategic value, enhancing operational efficiency and fostering innovation (Helbig et al., 2021). Governments have actively promoted open government policies to enhance the accessibility and utilization of OGD as a substantial source of information. Since the launch of open government initiatives in the United Kingdom and the United States in 2009, numerous countries, including South Korea, have pursued efforts to drive innovation and increase transparency through the use of OGD (O'Hara, 2012; Peled, 2011). In short, developed countries have been quantitatively expanding open data since the beginning of their policies, and this has created considerable added value based on OGD.

However, the translation of quantitative outputs of OGD into qualitative outcomes remains limited (Hong, 2014). The fundamental value of OGD lies not merely in its availability but in its applicability to real-world business and societal contexts. In other words, the true potential of OGD is realized when end users can effectively leverage it (Janssen et al., 2012). Given that the utility derived from OGD represents the core of its value (Weerakkody et al., 2017), proactive measures should be taken to evaluate and enhance data usability from the perspective of the end user. Research on the utilization of OGD has explored various aspects but has often taken a reactive approach, focusing on indirect factors rather than the primary purpose of OGD utilization. For instance, studies on OGD legislation and governance have highlighted the importance of clarifying the interplay between related laws and regulations, cautioning against the indiscriminate enactment and amendment of OGD-related legislation in an era of rapid societal change (Kim, 2023; Kim & Lee, 2020). These studies emphasize the need for a robust and coherent legal framework (Devins et al., 2017). Regarding OGD utilization policies, some research has compared the OGD policies of different Korean administrations, initially emphasizing the principle of openness (Kim et al., 2023). Other studies, based on analyses of OGD policies, have recommended identifying data closely tied to everyday life to enhance utilization and designing policies from a user-centric perspective (Yun & Hyun, 2019). Additionally, research on OGD technology has provided technical recommendations for improving the infrastructure of platforms that host OGD. These studies have also offered insights into tools and technical standards for data analysis (Jetzek et al., 2013; Lee, 2020; Máchová et al., 2018; Osagie et al., 2017).

The field of automated machine learning (AutoML) has witnessed remarkable progress over the past decade, achieving significant breakthroughs across multiple core domains, including model selection (Feurer et al., 2015; Thornton et al., 2013), hyperparameter optimization (Bergstra & Bengio, 2012; Snoek et al., 2012), and neural architecture search (Real et al., 2019; Zoph & Le, 2017). These foundational advances have established AutoML as a transformative paradigm that efficienates machine learning by automating traditionally labor-intensive optimization processes. Contemporary AutoML research has predominantly concentrated on conventional machine learning tasks, with substantial emphasis on image classification (Elsken et al., 2019), natural language processing (So et al., 2019), and time series forecasting (Kang et al., 2017). This body of work has primarily pursued performance optimization objectives using curated benchmark datasets, reflecting the field's emphasis on algorithmic sophistication and computational efficiency.

However, existing AutoML research exhibits several critical limitations that constrain its practical applicability. First, the predominant focus on algorithmic performance optimization has overshadowed considerations of data quality and structural characteristics, thereby limiting the predictive utility of these systems when deployed on heterogeneous and incomplete datasets typical of public sector environments (Zöller & Huber, 2021). This algorithmic myopia represents a fundamental misalignment between research priorities and real-world data complexity. Second, current evaluation frameworks rely heavily on static performance metrics such as accuracy and F1-score, failing to capture the dynamic and contextual factors that characterize actual OGD data utilization patterns and user requirements (Hutter et al., 2019). This narrow evaluation paradigm inadequately reflects the multifaceted nature of public sector AI deployment scenarios, where performance must be assessed within broader operational and stakeholder contexts. Third, the application

 https://doi.org/10.1633/JISTaP.2025.13.4.6

domain of AutoML research has been disproportionately concentrated on commercial and academic benchmarking tasks, consequently neglecting the unique challenges and opportunities presented by public sector data governance and policy-oriented applications. This bias has resulted in a systematic underexploration of AutoML's potential for generating social value and supporting evidence-based policymaking processes.

The AutoML approach employed in this study presents a novel methodological framework for systematically analyzing the complex stakeholder networks and innovative civic engagement processes that have long captivated information science researchers within OGD ecosystems. The contemporary emphasis on dynamic development approaches for OGD platform ecosystem construction, which prioritizes early-stage user engagement, finds strong methodological support in the AutoML framework. By leveraging metadata accumulated through user participation, this approach provides empirical foundations for continuous system improvement processes. The integration of user-generated metadata with automated learning algorithms creates feedback loops that can inform adaptive platform design and enhance stakeholder engagement strategies over time.

The core methodological approach of this research is grounded in Information Triangulation Theory, a well-established principle in the social sciences. Triangulation refers to the deliberate use of multiple data sources, methods, or theoretical perspectives to overcome the inherent biases and limitations of a single-method or single-observer study (Denzin, 2009). The central premise of this theory is that by cross-validating a phenomenon from diverse standpoints, a study can achieve a higher degree of credibility and validity in its findings (Patton, 2014). When information from multiple, distinct sources converges on a consistent conclusion, that conclusion is considered more robust and trustworthy than one derived from any single source alone.

This principle provides a compelling theoretical rationale for the central comparison in this study: integrated versus domain-specific training. The use of an integrated data training strategy can be conceptualized as a computational analogue to data triangulation. By combining heterogeneous metadata from disparate government domains into a single dataset, the predictive models are exposed to a more holistic view of the information ecosystem, allowing them to learn beyond the limited perspective and potential biases of any one domain. This research, therefore, hypothesizes that the triangulation of information afforded by multi-domain data will enable the construction of more generalizable and accurate predictive models than those trained on siloed, domain-specific data. This study empirically tests whether this data integration strategy, grounded in triangulation theory, can provide a more reliable understanding of the complex phenomenon of OGD utilization.

While prior studies have largely proposed reactive and prescriptive measures for OGD utilization, they have been limited in proactively addressing utilization challenges. Few studies have offered forward-looking solutions in this regard. Against this backdrop, the present study seeks to address the following research questions: (1) Is AutoML an appropriate tool for measuring and evaluating the utilization of OGD? (2) How does the composition of the training data affect the performance of models designed to predict OGD utilization? These questions serve as the foundation for the analysis presented in this study. In other words, these research questions were formulated by identifying research gaps through a review of the literature. The results of the research questions explored in this study can contribute to the advancement of the data business and enhance the products and services associated with it.

2. RELATED RESEARCH

2.1. Research on Open Government Data

OGD has increasingly been recognized as a resource for innovation, capable of generating added value through its inherent characteristics (Blakemore & Craglia, 2006; Charalabidis et al., 2018). The benefits of actively leveraging OGD have been identified across diverse stakeholders—such as researchers, businesses, and policymakers—who use it to develop new business opportunities or enhance decision-making processes (Zhang et al., 2005). Notably, several studies have demonstrated that utilizing OGD fosters innovation in the products and services of private companies across various sectors (Kitsios et al., 2017). The use of OGD is further emphasized as a means to facilitate data-driven business integration, enhance decision support, and improve overall business quality (Janssen & Zuiderwijk, 2014). Legal and governance research on OGD utilization highlights its contributions to achieving diverse social and political objectives, such as increasing transparency, fostering public participation, enhancing democratic accountability, promoting economic growth and innovation, and improving public sector efficiency while reducing costs. However, these studies also underscore the potential risks, such as threats to privacy and related rights due to the disclosure of government data containing personal information (Borgesius et al., 2015).

In terms of policy research, Janssen and Zuiderwijk (2014) noted that governments seek to encourage the disclosure and use of OGD to derive significant benefits. Drawing on an analysis of Dutch government agency policies, the study proposes a framework for evaluating OGD policies. This framework identifies four critical elements for policy assessment: the policy environment and context, the content of the policy, performance indicators, and the public value created through data utilization. Through these prior studies, it is evident that research from diverse perspectives has been actively conducted to enhance the utilization of OGD. However, much of the existing research has taken an indirect and reactive approach to assessing the value of OGD utilization. In response, this study seeks to propose a differentiated framework to proactively measure and evaluate the intangible value of OGD utilization. Specifically, the study focuses on data provided by the Korean government, aiming to develop a more direct and forward-looking approach to assess its potential value.

2.2. Research on Information Triangulation Theory and Prediction Models

The superior performance of models trained on integrated datasets can be theoretically grounded in the principles of Information Triangulation. Originating as a methodological concept in the social sciences, triangulation refers to the practice of using multiple data sources, methods, or theories to obtain a more robust and validated understanding of a phenomenon (Denzin, 2009). The core tenet is that synthesizing information from diverse perspectives can mitigate the biases inherent in any single source, thereby leading to a more comprehensive and credible conclusion (Patton, 2014). In the context of this study, the integrated training approach serves as a computational analogue to data triangulation. By combining metadata from disparate government domains, the model gains a more holistic view of the data ecosystem, enabling it to learn more generalizable patterns than would be possible from a single, siloed domain.

This study employs machine learning as a tool for the quantitative evaluation of OGD utilization. Machine learning has already been widely applied across various domains closely tied to daily life, including transportation (Bak & Kim, 2023), healthcare (Kruppa et al., 2012), information and communication (Lee et al., 2021), and environmental studies (Ahmed et al., 2019). Its introduction has significantly expanded the scope of insights and knowledge that can be applied within these fields.

Numerous studies have leveraged machine learning to predict patent counts, which serve as indicators of utilization in the patent domain. For instance, classification algorithms such as Self-Organizing Maps (SOM), Kernel Principal Component Analysis (KPCA), and support vector machines (SVM) have been applied to classify patent utilization (Wu et al., 2016). Similarly, XGBoost, a boosting-based classification algorithm, has been employed to predict patent counts (Cho & Lee, 2018). Moreover, Erdogan et al. (2024) explored high-utilization patents, enabling companies to make informed R&D investment decisions (Erdogan et al., 2024). Collectively, these studies highlight the validity and practicality of using machine learning tools to measure and evaluate the intangible value of patent utilization proactively. These findings affirm the suitability of machine learning methods for assessing and predicting utilization metrics in related contexts.

A growing body of research has explored predictions using AutoML. For example, Feurer et al. (2015) demonstrated that AutoML can automate machine learning pipelines, significantly reducing the time required to construct models while enabling rapid iterative testing of models and hyperparameters (Feurer et al., 2015). Their findings indicate that AutoML often produces higher-performing predictive models compared to traditional manual methods. More recently, Ferreira (2024) evaluated the effectiveness of AutoML as a tool for supervised learning. A range of AutoML tools—including Auto-Keras, Auto-PyTorch, Auto-sklearn, AutoGluon, H2O AutoML, rminer, TPOT, and TransmogrifAI—was analyzed and their performance compared, underscoring the versatility and potential of AutoML technologies.

Regarding training data for machine learning, prior research has provided valuable insights. Studies on data characteristics suggest that tree-based algorithms are particularly well-suited for datasets containing a significant amount of categorical data. Additionally, with respect to data size, existing research has confirmed that larger training datasets typically lead to better classification performance compared to smaller datasets (Foody et al., 2006). Building on these prior studies, the present research adopts the AutoML approach to investigate the impact of training dataset composition on model performance. Specifically, it compares the performance of models trained on an aggregated dataset, which combines data from multiple disciplines, with models trained on datasets specific to each of the 16 disciplines. This comparison aims to determine whether differences in model performance are influenced by variations in the size of training data across these disciplines.

3. MATERIALS AND METHODS

3.1. Data

This study focuses on structured data provided by the Korean government through the OGD portal (data.go.kr) from 2012 to 2022. The metadata was collected and analyzed by categorizing the data into two utilization methods: download-based File Data and application programming interface (API)-based data. Metadata, which provides detailed descriptive information about data, was collected with a focus on classification systems, data descriptions, and other relevant attributes. Variables deemed unusable, such as contact names and phone numbers, were excluded from the analysis. To account for variations in data volume across the 16 administrative sectors defined by the scope of the Korean government's tasks, the training datasets were divided into two categories: an integrated dataset, where data from all sectors were combined without categorization, and sector-specific datasets, where data were categorized according to the 16 sectors. For this study, the data was sourced from original datasets held within information systems operated by the South Korean government. From these sources, 16 domain-specific datasets were comprehensively pooled for use. The entire set of metadata items, as designated by the South Korean government's administrative guidelines, was utilized for the analysis. Furthermore, in accordance with the research design, any records with missing metadata were consequently removed prior to conducting the analysis.

To ensure the reliability and robustness of the proposed predictive models, a systematic data cleansing and preprocessing pipeline was applied to the raw datasets. For the File Data, the initial dataset consisted of 88,027 records. A multi-stage filtering process was implemented to remove data unsuitable for model training and to guarantee data quality and consistency. First, records with missing values in the input variables were identified and excluded. Second, datasets that were discarded or rejected at the request of the providing OGD institution were removed. Finally, as this study's scope is confined to OGD provided by the South Korean government up to the year 2022, all data registered after this cutoff were filtered out. This process resulted in a final analytical dataset of 44,648 records. As a specific example, the data for the Public Administration domain was refined from an initial 8,220 records to 6,051, with 2,169 records being removed. The API Data underwent an identical cleansing pipeline. The initial dataset of 13,444 records was preprocessed to yield a final analytical set of 6,677 records. For the Public Administration domain within the API Data, the record count was reduced from 1,787 to 876, with 911 records being excluded. Ultimately, approximately 50.7% of the original raw File Data and 49.7% of the original raw API Data were utilized for the final analysis.

3.1.1. Input Variables

Given the distinct metadata configurations inherent to different data access paradigms, this research constructed separate training datasets for File Data and API Data to accommodate their respective utilization patterns. File Data represents a download-based access model, while API Data enables real-time data retrieval through programmatic interfaces, necessitating differentiated feature engineering approaches. The File Data dataset comprises 37 variables, consisting of 23 continuous variables and 14 categorical variables. In contrast, the API Data dataset incorporates 42 variables, encompassing 23 continuous variables and 19 categorical variables. This asymmetric variable distribution reflects the additional metadata complexity associated with API-based data access mechanisms. Several variables are common to both data access modalities. The shared categorical variables include classification taxonomy, provider institution, usage license scope, research publication status, and fee structure. Common continuous variables encompass file size, data description length, keyword count within data descriptions, and null value frequency. These shared features capture fundamental data characteristics that transcend access methodology. API Data exhibits unique characteristics that necessitate additional variables specific to this access paradigm. API-specific categorical variables include detailed API type classification, API category designation, and open data center API provision status. These variables capture the technical and institutional dimensions unique to programmatic data access, reflecting the enhanced complexity of real-time data integration workflows compared to static file-based access patterns.

3.1.2. Target Variables

To establish target variables that quantitatively represent the utilization of OGD, indicators were designed to follow a normal distribution, ensuring better model performance (Hong et al., 2003; West et al., 1995). The normality of the distribution of each target variable was validated (Kwon et al., 2019). For File Data, the target variable was adjusted to account for the number of downloads over the provision period, incorporating corrections for cases where multiple files were provided as attachments. This adjustment reflects the concept of exploitation half-life, commonly used in the patent field, to account for the duration of data availability, as shown in Fig. 1A.

For API data, the target variables included the number of API calls and the number of applications utilizing the API. Similar to File Data, these variables were adjusted for the provision period to ensure consistency in measuring utilization. The normality of these variables was also checked to confirm their suitability for constructing indicators, as shown in Fig. 1B.

3.2. Proposed Methods

3.2.1. Automated Machine Learning Method

AutoML refers to a suite of tools designed to streamline the development of machine learning models by automating various stages of the process. These stages include data preprocessing, model selection, hyperparameter optimization, model evaluation, and deployment, minimizing the need for manual intervention. AutoML offers several advantages, such as accelerating the model development process by simplifying complex workflows and improving performance through automated hyperparameter tuning (Elshoky et al., 2022; Feurer et al., 2022).

In this study, the Auto-sklearn library was utilized, an AutoML framework built on scikit-learn. Auto-sklearn systematically explores a range of machine learning algorithms—including SVM, Random Forest (RF), and k-Nearest Neighbors—to identify the model that best fits the data. Additionally, it employs Bayesian optimization, a probabilistic technique leveraging Gaussian Processes, to optimize hyperparameters efficiently and achieve superior performance with fewer iterations. Auto-sklearn also supports the construction of ensemble models, which combine predictions from multiple individual models to enhance overall performance (Feurer et al., 2022). Finally, a comparative analysis was conducted against a standalone RF model to directly evaluate the AutoML approach, followed by a SHapley Additive exPlanations (SHAP) value analysis to investigate the model's feature importances.

Park (2022) employed Auto-sklearn to develop a machine learning model for predicting pollutant concentrations in rivers. The study demonstrated the potential of AutoML as an effective tool for forecasting pollution sources in environmental contexts through a comparative performance analysis (Park, 2022). Similarly, Angarita-Zapata et al. (2020) utilized Auto-sklearn to address the challenge of traffic prediction, demonstrating its effectiveness in accurately modeling traffic patterns and assessing the performance of network equipment. The mechanism of Auto-sklearn utilized in the literature and in this study is shown in Fig. 2.

A

$$\textit{Target variable of File Data} = \log\left(\frac{\textit{Download counts}}{\textit{Provision period} \times \textit{number of attachments}}\right)$$

B

$$\textit{Target variable of API data} = \log\left(\frac{\textit{API call counts} \times \textit{API request counts}}{\textit{Service period}}\right)$$

Fig. 1. Target variable of File Data (A) and API Data (B). API, application programming interface.

3.2.2. Model Performance Evaluation

To assess model performance, this research utilized the mean squared error (MSE) and root mean squared error (RMSE) metrics. These metrics measure the magnitude of prediction errors, defined as the difference between predicted and actual values. While MSE is sensitive to outliers due to its squaring of errors, RMSE, which incorporates the square RMSE, provides a more intuitive interpretation of the average error magnitude as it retains the same units as the original data (Almalaq & Edwards, 2017; Aswin et al., 2018).

MSE, defined as the average of the squared differences between predicted and actual values, is particularly sensitive to large prediction errors. This sensitivity makes it

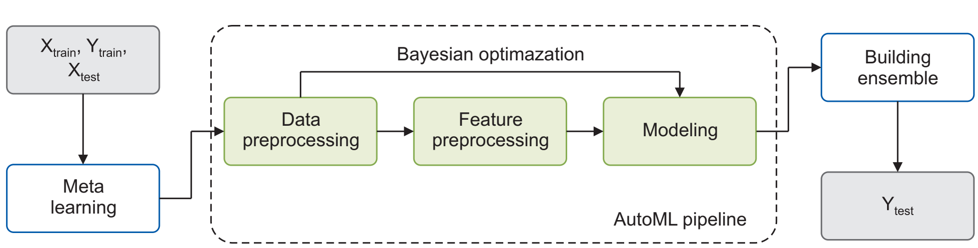

Fig. 2. Mechanism of the automated machine learning used in this study. AutoML, automated machine learning.

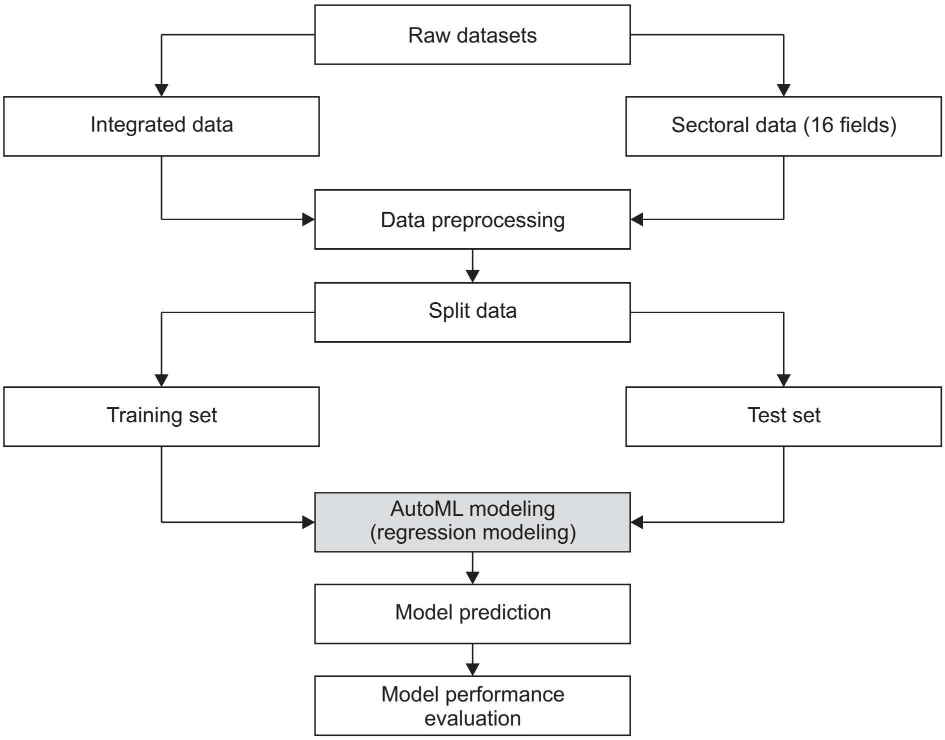

Fig. 3. Workflow for predicting the usability of open government data using automated machine learning model. AutoML, automated machine learning.

useful for identifying models where significant outliers may exist. RMSE, on the other hand, is a scale-dependent error measure derived by taking the square root of MSE, making it less influenced by extreme values and more interpretable when evaluating average error size. RMSE tends to increase when predicting larger magnitude values and decrease for smaller magnitude predictions, offering a comprehensive understanding of model performance (Aswin et al., 2018).

Both MSE and RMSE are widely employed in machine learning regression tasks to evaluate and train predictive models. The workflow for performance evaluation, including the calculation of MSE and RMSE, is presented in Fig. 3. A lower MSE or RMSE indicates superior predictive performance, and these metrics were integral to analysis of model accuracy and reliability (Almalaq & Edwards, 2017). Finally, the formulas for RMSE and MSE are shown in Fig. 4.

A

$$MSE=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \quad (y_i\text{=predicted value, }\hat{y}_i\text{=actual value})$$

B

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \quad (y_i\text{=predicted value, }\hat{y}_i\text{=actual value})$$

Fig. 4. Formula of MSE (A) and RMSE (B). MSE, mean squared error; RMSE, root mean squared error.

4. RESULTS

4.1. File Data

An analysis of the study results across File Data for 16 major taxonomies revealed notable differences in model performance between integrated and sector-specific training data. Analyzing the results in Table 1, for the education sector, models trained on integrated data achieved an RMSE of 0.684 and an MSE of 0.468, compared to an RMSE of 0.705 and an MSE of 0.497 for sector-specific data. These findings indicate better model performance with integrated data, as evidenced by lower error metrics and more accurate predictions. Similarly, 11 other sectors, including transportation and logistics (RMSE 0.695, MSE 0.484), agriculture and fisheries (RMSE 0.674, MSE 0.454), culture and tourism (RMSE 0.688, MSE 0.473), healthcare (RMSE 0.606, MSE 0.367), social welfare (RMSE 0.672, MSE 0.451), industry and employment (RMSE 0.685, MSE 0.469), disaster safety (RMSE 0.642, MSE 0.412), finance (RMSE 0.583, MSE 0.340), unification and diplomacy (RMSE 0.666, MSE 0.443), and environment and meteorology (RMSE 0.713, MSE 0.508), exhibited superior performance on integrated data.

Table 1. Model performance by 16 sectors (File Data)

Item				
Field	Training data			
	Integrated data		Field data	
	Evaluation indicator			
	RMSE	MSE	RMSE	MSE
Public administration	0.727	0.529	0.708	0.501
Science and technology	0.931	0.866	0.737	0.543
Education	0.684	0.468	0.705	0.497
Transportation and logistics	0.695	0.484	0.782	0.612
Land management	0.729	0.531	0.703	0.494
Agriculture and fisheries	0.674	0.454	0.687	0.472
Culture and tourism	0.688	0.473	0.732	0.535
Law	0.614	0.377	0.572	0.327
Healthcare	0.606	0.367	0.709	0.502
Social welfare	0.672	0.451	0.694	0.482
Industry and employment	0.685	0.469	0.702	0.492
Food and health	0.703	0.494	0.663	0.440
Disaster safety	0.642	0.412	0.805	0.648
Finance	0.583	0.340	0.585	0.342
Unification and diplomacy	0.666	0.443	0.674	0.455
Environment and meteorology	0.713	0.508	0.780	0.609

RMSE, root mean squared error; MSE, mean squared error.

Conversely, in the public administration sector, the model performed better with sector-specific data (RMSE 0.708, MSE 0.501) than with integrated data (RMSE 0.727, MSE 0.529). A similar trend was observed in the science and technology sector, where sector-specific data led to improved performance (RMSE 0.737, MSE 0.543) compared to integrated data (RMSE 0.931, MSE 0.866). In addition to public administration and science and technology, five other sectors—including land management (RMSE 0.703, MSE 0.494), law (RMSE 0.572, MSE 0.327), and food and health (RMSE 0.663, MSE 0.440) showed better performance with sector-specific training data. Overall, RMSE values ranged from 0.583 to 0.934 for integrated data, reflecting a 1.60-fold variation in performance, and from 0.572 to 0.805 for sector-specific data, indicating a 1.41-fold difference. These results highlight a general trend where integrated data provides superior performance in certain domains, while sector-specific data yields better results in others. Additionally, the comparative analysis of MSE and RMSE—both of which are sensitive to error—reinforced these findings, further underscoring the influence of training data granularity on model performance.

4.2. API Data

An analysis of the study results using API Data across 16 major taxonomies revealed distinct differences in model performance between integrated and domain-specific training data. Analyzing the results in Table 2, for the science and technology sector, the model trained on integrated data achieved an RMSE of 4.569 and an MSE of 20.874, integrated to an RMSE of 4.801 and an MSE of 23.050 for domain-specific data. These results indicate that the integrated data yielded slightly lower errors and more accurate predictions. Similar trends were observed in 10 other sectors where integrated data outperformed domain-specific data, including education (RMSE 2.581, MSE 6.661), transportation and logistics (RMSE 3.932,

Table 2. Model performance by 16 sectors (API Data)

Item				
Field	Training data			
	Integrated data		Field data	
	Evaluation indicator			
	RMSE	MSE	RMSE	MSE
Public administration	3.629	13.168	3.214	10.330
Science and technology	4.569	20.874	4.801	23.050
Education	2.581	6.661	2.587	6.693
Transportation and logistics	3.932	15.460	4.084	16.683
Land management	3.788	14.350	3.018	9.108
Agriculture and fisheries	2.911	8.475	3.215	10.336
Culture and tourism	2.447	5.989	2.637	6.956
Law	2.011	4.046	4.136	17.105
Healthcare	2.332	5.437	2.567	6.589
Social welfare	2.772	7.686	2.262	5.116
Industry and employment	3.082	9.497	3.420	11.694
Food and health	2.825	7.979	2.782	7.738
Disaster safety	2.899	8.403	2.398	5.749
Finance	4.679	21.895	3.233	10.451
Unification and diplomacy	2.871	8.243	3.122	9.748
Environment and meteorology	2.602	6.770	2.985	8.913

API, application programming interface; RMSE, root mean squared error; MSE, mean squared error.

MSE 15.460), agriculture and fisheries (RMSE 2.911, MSE 8.475), culture and tourism (RMSE 2.447, MSE 5.989), law (RMSE 2.011, MSE 4.046), healthcare (RMSE 2.332, MSE 5.437), industry and employment (RMSE 3.082, MSE 9.497), unification and diplomacy (RMSE 2.871, MSE 8.243), and environment and meteorology (RMSE 2.602, MSE 6.770).

Conversely, in the public administration sector, domain-specific data produced better results, with an RMSE of 3.214 and an MSE of 10.330 compared to an RMSE of 3.629 and an MSE of 13.168 for integrated data. A similar pattern was observed in the land management sector, where domain-specific data outperformed integrated data with an RMSE of 3.018 and an MSE of 9.108, compared to an RMSE of 3.788 and an MSE of 14.350. In addition to public administration and land management, six other sectors—social welfare (RMSE 2.262, MSE 5.116), food and health (RMSE 2.782, MSE 7.738), disaster safety (RMSE 2.398, MSE 5.749), and finance (RMSE 3.233, MSE 10.451) demonstrated better performance with domain-specific training data. Overall, RMSE values for models trained on integrated data ranged from 2.011 to 4.569, reflecting a performance difference of approximately 2.27. For domain-specific data, RMSE values ranged from 2.262 to 4.801, indicating a performance difference of approximately 2.12. A similar trend was observed with MSE and RMSE metrics, which are particularly sensitive to error, further reinforcing the performance variations between integrated and domain-specific training approaches.

4.3. Comparative Performance Analysis of AutoML and Random Forest Model (File Data)

To evaluate the suitability of an AutoML-based approach, this study conducts a comparative performance analysis against a standard RF model. The evaluation utilizes RMSE and MSE as performance metrics, with results measured separately for models trained on integrated and domain-specific (field) data. The experimental

Table 3. Random forest model performance by 16 sectors (File Data)

Item				
Field	Training data			
	Integrated data		Field data	
	Evaluation indicator			
	RMSE	MSE	RMSE	MSE
Public administration	0.745	0.556	0.735	0.541
Science and technology	0.891	0.793	0.769	0.592
Education	0.692	0.478	0.727	0.528
Transportation and logistics	0.733	0.538	0.794	0.631
Land management	0.731	0.534	0.708	0.501
Agriculture and fisheries	0.694	0.482	0.700	0.491
Culture and tourism	0.732	0.535	0.762	0.581
Law	0.620	0.385	0.611	0.374
Healthcare	0.660	0.435	0.723	0.523
Social welfare	0.697	0.486	0.695	0.483
Industry and employment	0.705	0.497	0.717	0.515
Food and health	0.738	0.544	0.720	0.518
Disaster safety	0.672	0.451	0.814	0.663
Finance	0.600	0.360	0.592	0.350
Unification and diplomacy	0.689	0.475	0.682	0.465
Environment and meteorology	0.727	0.528	0.776	0.601

RMSE, root mean squared error; MSE, mean squared error.

results for the RF model are presented in Table 3, and the performance delta between the two approaches is summarized in Table 4. The empirical results indicate that the AutoML-based models generally outperformed the standalone RF models, achieving lower prediction errors (RMSE and MSE) across most test cases. For instance, in the Public administration domain, the AutoML model trained on domain-specific data achieved an RMSE of 0.708, a notable improvement over the RF model's RMSE of 0.735, corresponding to a performance gain of approximately 0.027. This suggests that AutoML's automated optimization process enables a more precise capture of complex patterns inherent in the domain-specific data.

However, the superiority of AutoML was not universal. An exceptional case was observed in the science and technology domain, where the RF model trained on integrated data yielded a lower RMSE (0.891) than its AutoML counterpart (0.931). Nevertheless, for this specific domain, the highest performance was achieved when models were trained on domain-specific data, a scenario in which the AutoML approach demonstrated superior results. Therefore, the application of AutoML is still considered appropriate and beneficial in this context. In aggregate, the experimental findings provide empirical evidence that an AutoML approach, which automates the complex processes of model selection and hyperparameter tuning, represents a more effective and robust alternative to employing a single, manually configured model for the task of predicting public data utilization.

4.4. Feature Importance

4.4.1. File Data

An analysis of feature contributions to the target variable, conducted via SHAP values, reveals that data size exerts a decisive influence on the model's predictions. The distribution of SHAP values for this feature was observed to be asymmetric; it is relatively narrow and concentrated

Table 4. Performance comparison of AutoML and random forest model (File Data)

Item				
Field	Training data			
	Integrated data		Field data	
	Evaluation indicator			
	RMSE	MSE	RMSE	MSE
Public administration	-0.01849	-0.02676	-0.02746	-0.03990
Science and technology	0.04036	0.07276	-0.03232	-0.04885
Education	-0.00760	-0.01031	-0.02193	-0.03143
Transportation and logistics	-0.03839	-0.05386	-0.01249	-0.01921
Land management	-0.00185	-0.00314	-0.00476	-0.00693
Agriculture and fisheries	-0.02018	-0.02789	-0.01341	-0.01857
Culture and tourism	-0.04355	-0.06217	-0.03015	-0.04587
Law	-0.00648	-0.00800	-0.03949	-0.04692
Healthcare	-0.05387	-0.06843	-0.01414	-0.02093
Social welfare	-0.02504	-0.03487	-0.00110	-0.00117
Industry and employment	-0.01969	-0.02759	-0.01541	-0.02267
Food and health	-0.03475	-0.05027	-0.05671	-0.07798
Disaster safety	-0.02954	-0.03897	-0.00899	-0.01459
Finance	-0.01700	-0.02000	-0.00660	-0.00799
Unification and diplomacy	-0.02315	-0.03192	-0.00813	-0.01030
Environment and meteorology	-0.01368	-0.02006	0.00448	0.00756

AutoML, automated machine learning; RMSE, root mean squared error; MSE, mean squared error.

in the positive region, while being markedly wide and dispersed in the negative region. This characteristic implies that an increase in data size does not proportionally induce a sharp rise in the predicted value. Instead, the analysis empirically confirms a clear and consistent negative relationship between File Data size and its utilization. The prevalence of blue points (low feature values) in the negative SHAP region indicates that as data size diminishes, the predicted utilization correspondingly decreases. This pattern can be interpreted to mean that a minimum threshold of data size is necessary to ensure a baseline level of practical value.

For the open data provider, interpretation must account for its nature as a categorical variable. Unlike with continuous variables, the color gradient across points does not represent a numerical or ordinal relationship but rather reflects the qualitative differences and unique characteristics of each distinct provider. The wide dispersion of SHAP values across both positive and negative domains clearly indicates that the impact of a provider's unique organizational characteristics on the target variable is highly heterogeneous. This suggests that while some providers act as positive catalysts for data utilization, others may exert a limited or even detrimental influence.

The data core keyword count feature exhibits a pattern consistent with theoretical expectations, demonstrating a clear positive impact on the predicted outcome. The dense concentration of red points (high feature values) in the positive SHAP region empirically substantiates the existence of a robust and consistent positive correlation between an increase in the number of core keywords and a corresponding rise in the predicted utilization value. A detailed visualization of this SHAP value analysis for all input variables is presented in Fig. 5, which illustrates the relative importance and directional contribution of each feature.

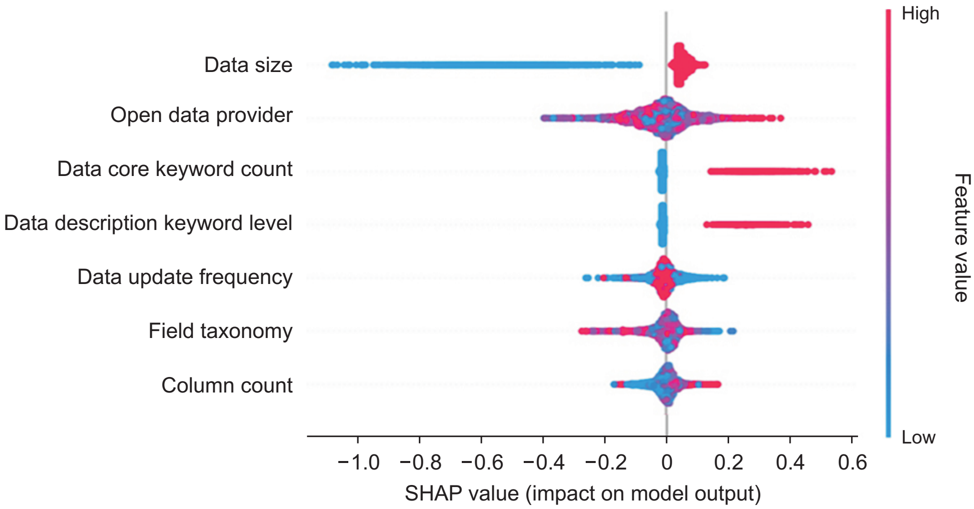

Fig. 5. Summary plot illustrating impacts of features (File Data). SHAP, SHapley Additive exPlanations.

Fig. 6. Summary plot illustrating impacts of features (API Data). API, application programming interface; SHAP, SHapley Additive exPlanations.

4.4.2. API Data

An analysis of the SHAP values for the API Data model identified the Scope of data use permission as the most powerful and decisive predictive feature. This indicates that the utilization of API data is strongly contingent on the permissiveness of its usage rights. The SHAP value distribution reveals a distinct pattern: Red points, representing a broader scope of permission, are concentrated in the positive region, while blue points, representing a more restrictive scope, are concentrated in the negative region. This color-based distribution confirms that as the scope of data use expands, the predicted utilization value shows a statistically significant increase. Conversely, a more restrictive scope corresponds to a substantial decrease in the prediction.

Consistent with the analysis of File Data, the open data provider feature must be interpreted as a categorical variable, where color variations reflect qualitative differences between providers rather than a numerical scale. The wide dispersion of SHAP values across both positive and negative domains suggests that the influence of each provider is highly heterogeneous. In effect, some organizations act as positive catalysts, promoting the use of their API Data, while others have a more limited or even a detrimental impact on utilization.

The analysis of the open data center API status feature revealed a clear tendency for APIs provided directly by the central open data center to be associated with relatively high prediction values. This can be interpreted as evidence that the standardized governance framework and consistent service quality of a central authority contribute to higher user confidence and a greater willingness to utilize the data.

The Data description null count feature was found to have a bidirectional and complex influence on the model, with its SHAP values distributed across a wide range in both positive and negative territories. A noteworthy inverse relationship was observed: Red points, indicating a high number of null values, were concentrated in the negative SHAP region, while blue points, indicating fewer null values, were predominant in the positive region. This pattern suggests that an increase in missing values within the data description significantly decreases the predicted utilization. Conversely, a lower null count, signifying greater data completeness, leads to a corresponding increase in the prediction. The summary plot for this SHAP value analysis is presented in Fig. 6.

5. DISCUSSION

As an alternative approach to proactively evaluate the utilization of OGD, this study developed a prediction model using an AutoML technique. The model was trained using metadata from accumulated OGD, with performance compared based on the classification of training data. The training data was categorized into integrated data and domain-specific data, considering the attributes of OGD (such as field and utilization method) and classified across 16 taxonomy systems. The results demonstrated that the models and methods yielding the best performance varied according to data attributes, aligning with previous research emphasizing the importance of considering data properties, which can significantly influence performance outcomes (Si et al., 2024). In terms of the implications for different utilization methods of OGD (File Data, API Data), API Data exhibited a broader range of performance indicators due to the wider distribution of target variables. This finding supports prior studies indicating that model performance is contingent on the characteristics of the independent and dependent variables in the data (Satoła & Satoła, 2024; Zhang & Zhu, 2022). Conversely, File Data outperformed API Data in terms of RMSE and MSE metrics, suggesting that File Data offers superior overall performance in predictive tasks.

While the integrated training approach demonstrated superior performance across a majority of domains, the exceptional result where domain-specific models outperformed in particular sectors warrants a deeper analysis. This anomalous phenomenon, where a model trained on a smaller, more focused dataset surpassed one trained on a larger, more diverse dataset, suggests that the principle of more data is always better is not universally applicable. This phenomenon appears to be attributable to the highly idiosyncratic characteristics present in these domains, a condition that can be hypothesized as negative transfer in machine learning.

Negative transfer refers to a scenario where the learning process on integrated data is adversely affected (Pan & Yang, 2010). In such cases, the integrated model, while learning general patterns from the aggregate data, may also learn patterns and noise that directly conflict with the specific logic of these particular domains. The domain-specific model, by contrast, is not exposed to this irrelevant or contradictory information, allowing it to optimize exclusively for the distinct patterns of its own domain (Weiss et al., 2016).

Furthermore, the success of domain-specific models in these particular sectors empirically demonstrates that the effectiveness of a data integration strategy is conditional upon the degree of underlying similarity between the domains being combined. Future research should involve a deeper feature-level analysis of these exceptional sectors to precisely identify the unique characteristics that cause this negative transfer effect, thereby enabling the development of more adaptive and selective data integration strategies.

The superior performance of models trained on integrated datasets compared to domain-specific models can be understood through the theoretical framework of transfer learning. Pan and Yang (2010) demonstrated that multi-domain trained models can enhance performance within individual domains by learning shared representations across disparate data sources. These results suggest that the rich feature space provided by integrated datasets can compensate for the limited characteristics inherent to individual domains, thereby enabling the construction of more robust predictive models. The effectiveness of integrated datasets is fundamentally linked to the critical importance of data diversity in machine learning systems. Torralba and Efros (2011) established that models trained on large-scale, diverse datasets exhibit superior generalization performance even within specific domains, a finding that aligns with empirical results. Within the open data context, integration across multiple domains appears to mitigate domain-specific biases while facilitating more balanced feature learning processes. From the perspective of feature representation learning in deep learning and AutoML frameworks, integrated datasets provide an environment conducive to learning richer, multi-layered feature representations. As demonstrated by Bengio et al. (2013), diverse domain data promotes hierarchical feature learning, enabling effective acquisition of both low-level common features and high-level domain-specific characteristics. The convergence of these findings across disparate application domains underscores the fundamental value of data integration strategies in advancing machine learning model performance and generalizability.

This study's findings contribute not only to the field of machine learning but also offer significant insights into the theoretical understanding of information utilization in digital government contexts. The consistently superior performance of models trained on integrated datasets can be framed within the established principles of Information Triangulation Theory. This theory posits that synthesizing information from multiple, diverse sources leads to a more robust and reliable understanding than relying on a single source. This research provides an empirical validation of this concept in a computational setting, demonstrating that the triangulation of data from various government domains enhances the predictive accuracy of OGD utilization. Rather than viewing information diversity purely as a computational optimization problem, these results suggest it is a fundamental principle for understanding and predicting information value in complex ecosystems.

The empirical evidence aligns with this theoretical framing. For File Data, models trained on integrated data demonstrated superior performance across most domains, with exceptions noted in highly specialized fields like public administration and science and technology. A similar trend was observed for API Data, where the integrated approach proved more effective in a majority of fields. These results suggest that while domain-specific characteristics are important, the broader context provided by an integrated dataset often contains richer signals that lead to more generalizable and robust predictive models. This finding directly supports the core tenet of Information Triangulation: that an aggregation of diverse perspectives, or in this case, data sources yields a more comprehensive and accurate assessment.

From a machine learning perspective, this phenomenon is consistent with the principles of transfer learning, where knowledge gained from multiple domains enhances performance on a specific task. As established by Pan and Yang (2010), models trained on diverse data sources can learn shared representations that compensate for the limited information within any single domain. The integrated datasets in this study likely facilitate a more balanced feature learning process, mitigating domain-specific biases and allowing the models to identify cross-domain patterns that are crucial for prediction. This provides a computational mechanism that validates the principles of Information Triangulation, bridging information science theory with machine learning practice.

Therefore, this research challenges the conventional focus of AutoML research on single-dataset performance optimization. The findings suggest that the process of data integration and selection should be considered a core component of the AutoML pipeline itself, a perspective that expands on the framework articulated by Hutter et al. (2019). By demonstrating that cross-domain interactions can be more consequential for predictive accuracy than single-domain features, this work highlights a critical, and often overlooked, phase in the AutoML workflow.

6. CONCLUSION

The academic significance of this study is its empirical validation that AutoML can serve as an effective tool for measuring and evaluating the intangible value of OGD utilization. By building and comparing the performance of AutoML models, this study demonstrated an alternative assessment approach not previously explored, addressing limitations in prior research. Furthermore, the results established the importance of tailoring the application of training data based on the attributes of OGD, such as field and utilization method. From a practical perspective, this study's framework enables OGD providers to proactively diagnose and evaluate data usage, allowing them to more effectively align data supply with user demand.

This study confirmed the technical feasibility of applying AutoML to predict OGD utilization and presented a systematic approach to assessing the potential value of public datasets. The primary finding is that the training data strategy—specifically, the choice between an integrated or domain-specific approach—is a critical factor influencing predictive performance. Models trained on integrated data sources outperformed domain-specific models in 11 of 16 File Data sectors and 10 of 16 API Data sectors. This result confirms that moving beyond domain-specific silos enhances predictive accuracy, a finding that aligns with established principles in machine learning regarding the benefits of diverse training sets.

From a practical standpoint, this research provides a framework that OGD providers can use for the proactive evaluation of data assets, enabling a more effective alignment of data supply with user demand. However, this study is subject to certain limitations.

Furthermore, it is important to acknowledge the inherent assumptions of the proposed approach. The use of download and API call metrics serves as a proxy for utilization, which may not fully capture the more nuanced concepts of true societal value or impact. The findings are also derived exclusively from the South Korean OGD context, and their generalizability to other governmental data ecosystems requires further investigation. Future research would be significantly strengthened by incorporating qualitative validation with end-users to provide a richer understanding of the factors that drive meaningful data adoption beyond simple access metrics.

Ultimately, this work advances the application of computational methods within information science. It presents a viable, data-driven methodology for prospectively evaluating information systems, shifting the focus from traditional retrospective metrics toward a forward-looking assessment of a dataset's potential value. By establishing an empirical basis for data strategy and governance, the findings offer a concrete methodology for policymakers and data scientists seeking to maximize the impact of open data initiatives.

CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES

Ahmed, I., Seadawy, A. R., & Lu, D. (2019). M-shaped rational solitons and their interaction with kink waves in the Fokas–Lenells equation. *Physica Scripta*, 94(5), 055205. https://doi.org/10.1088/1402-4896/ab0455

Almalaq, A., & Edwards, G. (2017, December 18-21). A review of deep learning methods applied on load forecasting. In X. Chen, B. Luo, F. Luo, V. Palade, & M. A. Wani (Eds.), *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)* (pp. 511-516). IEEE. https://doi.org/10.1109/ICMLA.2017.0-110

Angarita-Zapata, J. S., Masegosa, A. D., & Triguero, I. (2020, June 15–19). General-purpose automated machine learning for transportation: A case study of Auto-sklearn for traffic forecasting. In M. J. Lesot, S. Vieira, M. Z. Reformat, J. P. Carvalho, A. Wilbik, B. Bouchon-Meunier, & R. R. Yager (Eds.), *Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2020)* (pp. 728-744). Springer Cham. https://doi.org/10.1007/978-3-030-50143-3_57

Aswin, S., Geetha, P., & Vinayakumar, R. (2018). *Deep learning models for the prediction of rainfall*. Paper presented at 2018 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India. https://doi.org/10.1109/ICCSP.2018.8523829

Bak, S. H., & Kim, K. P. (2023). An AutoML-driven antenna performance prediction model in the autonomous driving radar manufacturing process. *KSII Transactions on Internet and Information Systems*, 17(12), 3330-3344. https://doi.org/10.3837/tiis.2023.12.006

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1798-1828. https://doi.org/10.1109/TPAMI.2013.50

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(1), 281–305.

Blakemore, M., & Craglia, M. (2006). Access to public-sector information in Europe: Policy, rights, and obligations. *The Information Society*, 22(1), 13-24. https://doi.org/10.1080/01972240500388180

Borgesius, F. Z., Eechoud, M. V., & Gray, J. (2015). *Open data, privacy, and fair information principles: Towards a balancing framework*. https://ssrn.com/abstract=2695005

Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., & Ferro, E. (2018). *The world of open data: Concepts, methods, tools and experiences*. Springer International Publishing.

Cho, H. & Lee, H. (2018, April 4-7). Predicting patent quality using machine learning techniques. *Proceedings of the Spring Joint Conference of the Korean Society of Industrial Engineering* (pp. 1343-1350), The Korean Operations Research and Management Science Society.

Denzin, N. K. (2009). *The research act: A theoretical introduction to sociological methods*. Routledge. https://doi.org/10.4324/9781315134543

Devins, C., Felin, T., Kauffman, S., & Koppl, R. (2017). The law and big data. *Cornell Journal of Law and Public Policy*, 27(2), Article 3.

Elshoky, B. R. G., Younis, E. M. G., Ali, A. A., & Ibrahim, O. A. S. (2022). Comparing automated and non-automated machine learning for autism spectrum disorders classification using facial images. *ETRI Journal*, 44(4), 613-623. https://doi.org/10.4218/etrij.2021-0097

Elsken, T., Metzen, J. H., & Hutter, F. (2019). Neural architecture search: A survey. *Journal of Machine Learning Research*, 20, 1-21.

Erdogan, Z., Altuntas, S., & Dereli, T. (2024). Predicting patent quality based on machine learning approach. *IEEE Transactions on Engineering Management*, 71, 3144-3157. https://doi.org/10.1109/TEM.2022.3207376

Ferreira, L. F. F. (2024). *An automated and efficient machine learning framework for one-class classification tasks* (doctoral dissertation). Universidade do Minho, Braga, Portugal.

Feurer, M., Eggensperger, K., Falkner, S., Lindauer, M., & Hutter, F. (2022). Auto-sklearn 2.0: Hands-free AutoML via Meta-Learning. *Journal of Machine Learning Research*, 23, 1-61.

Feurer, M., Klein, A., Eggensperger, K., Springenberg, J., Blum, M., & Hutter, F. (2015, December 7-12). Efficient and robust automated machine learning. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 28 (NIPS 2015)*. NeurIPS.

Foody, G. M., Mathur, A., Sanchez-Hernandez, C., & Boyd, D. S. (2006). Training set size requirements for the classification of a specific class. *Remote Sensing of Environment*, 104(1), 1-14. https://doi.org/10.1016/j.rse.2006.03.004

Hamari, J., Sjöklint, M., & Ukkonen, A. (2016). The sharing economy: Why people participate in collaborative consumption. *Journal of the Association for Information Science and Technology*, 67(9), 2047-2059. https://doi.org/10.1002/asi.23552

Helbig, R., von Höveling, S., Solsbach, A., & Marx Gómez, J. (2021). Strategic analysis of providing corporate sustainability open data. *Intelligent Systems in Accounting, Finance and Management*, 28(3), 195-214. https://doi.org/10.1002/isaf.1501

Hong, S., Malik, M. L., & Lee, M. K. (2003). Testing configural, metric, scalar, and latent mean invariance across genders in sociotropy and autonomy using a non-Western sample. *Educational and Psychological Measurement*, 63(4), 636-654. https://doi.org/10.1177/0013164403251332

Hong, Y. W. (2014). A study on the invigorating strategies for open government data. *Journal of the Korean Data and Information Science Society*, 25(4), 769-777. https://doi.org/10.7465/jkdi.2014.25.4.769

Hutter, F., Kotthoff, L., & Vanschoren, J. (2019). *Automated machine learning: Methods, systems, challenges*. Springer Cham. https://doi.org/10.1007/978-3-030-05318-5

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, 29(4), 258-268. https://doi.org/10.1080/10580530.2012.716740

Janssen, M., & Zuiderwijk, A. (2014). Infomediary business models for connecting open data providers and users. *Social Science Computer Review*, 32(5), 694-711. https://doi.org/10.1177/0894439314525902

Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2013). *Generating value from open government data.* Paper presented at Thirty Fourth International Conference on Information Systems, Milano, Italy.

Kang, Y., Hyndman, R. J., & Smith-Miles, K. (2017). Visualising forecasting algorithm performance using time series instance spaces. *International Journal of Forecasting*, 33(2), 345-358. https://doi.org/10.1016/j.ijforecast.2016.09.004

Kim, E. S. (2023). A study on the improvement of the legal system for the promotion of opening and utilization of open government data - Focusing on cases of refusal to provide -. *Informatization Policy*, 30(2), 46-67. https://doi.org/10.22693/NIAIP.2023.30.2.046

Kim, G. H., Jeong, S. H., Yang, J. D., & We, J. Y. (2023). A decade of open data policy research using big data analytics techniques: Focusing on comparative analysis by regime. *Public Policy Review*, 37(4), 45-67. https://doi.org/10.17327/ippa.2023.37.4.003

Kim, M. H., & Lee, B. O. (2020). Recent trends of the amendment of the public sector information directive in EU and its implications to the Republic of Korea. *SungKyunKwan Law Review*, 32(1), 1-30. https://doi.org/10.17008/skklr.2020.32.1.001

Kitsios, F., Papachristos, N., & Kamariotou, M. (2017, July 24-27). Business models for open data ecosystem: Challenges and motivations for entrepreneurship and innovation. In D. Diaz, Y. Manolopoulos, B. Theodoulidis, & M. Zaki (Eds.), *2017 IEEE 19th Conference on Business Informatics (CBI)* (pp. 398-407). IEEE. https://doi.org/10.1109/CBI.2017.51

Kruppa, J., Ziegler, A., & König, I. R. (2012). Risk estimation and risk prediction using machine-learning methods. *Human Genetics*, 131(10), 1639-1654. https://doi.org/10.1007/s00439-012-1194-y

Kwon, H., Park, J., & Lee, Y. (2019). Stacking ensemble technique for classifying breast cancer. *Healthcare Informatics Research*, 25(4), 283-288. https://doi.org/10.4258/hir.2019.25.4.283

Lee, D. S., Choi, W. I., Nam, Y., & Park, Y. S. (2021). Predicting potential occurrence of pine wilt disease based on environmental factors in South Korea using machine learning algorithms. *Ecological Informatics*, 64, 101378. https://doi.org/10.1016/j.ecoinf.2021.101378

Lee, J. W. (2020). Forecasting open government data demand using keyword network analysis. *Informatization Policy*, 27(4), 24-46. https://doi.org/10.22693/NIAIP.2020.27.4.024

Máchová, R., Volejníková, J., & Lněnička, M. (2018). Impact of e-government development on the level of corruption: Measuring the effects of related indices in time and dimensions. *Review of Economic Perspectives*, 18(2), 99-121. https://doi.org/10.2478/revecp-2018-0006

Mhlanga, D. (2023). Digital transformation education, opportunities, and challenges of the application of ChatGPT to emerging economies. *Education Research International*, 2023(1), 7605075. https://doi.org/10.1155/2023/7605075

O'Hara, K. (2012). *Transparency, open data and trust in government: Shaping the infosphere*. Paper presented at the 4th Annual ACM Web Science Conference, Evanston, Illinois. https://doi.org/10.1145/2380718.2380747

Osagie, E., Waqar, M., Adebayo, S., Stasiewicz, A., Porwol, L., & Ojo, A. (2017). *Usability evaluation of an open data platform*. Paper presented at the 18th Annual International Conference on Digital Government Research, Staten Island, NY, USA. https://doi.org/10.1145/3085228.3085315

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345-1359. https://doi.org/10.1109/TKDE.2009.191

Park, J. (2022). Development of a model to predict water quality using an automated machine learning algorithm. *Journal of the Korean Society of Water and Wastewater*, 36(6), 329-337. https://doi.org/10.11001/jksww.2022.36.6.329

Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). SAGE Publications.

Peled, A. (2011). When transparency and collaboration collide: The USA open data program. *Journal of the American Society for Information Science and Technology*, 62(11), 2085-2094. https://doi.org/10.1002/asi.21622

Real, E., Aggarwal, A., Huang, Y., & Le, Q. V. (2019). *Regularized evolution for image classifier architecture search,* Paper presented at the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, Honolulu, Hawaii, USA. https://doi.org/10.1609/aaai.v33i01.33014780

Satoła, A., & Satoła, K. (2024). Performance comparison of machine learning models used for predicting subclinical mastitis in dairy cows: Bagging, boosting, stacking, and super-learner ensembles versus single machine learning models. *Journal of Dairy Science*, 107(6), 3959-3972. https://doi.org/10.3168/jds.2023-24243

Si, B., Ni, Z., Xu, J., Li, Y., & Liu, F. (2024). Interactive effects of hyperparameter optimization techniques and data characteristics on the performance of machine learning algorithms for building energy metamodeling. *Case Studies in Thermal Engineering*, 55, 104124. https://doi.org/10.1016/j.csite.2024.104124

Snoek, J., Larochelle, H., & Adams, R. P. (2012, December 3-8). Practical Bayesian optimization of machine learning algorithms. In F. Pereira, C. J. Burges, L. Bottou, & K.Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 25 (NIPS 2012)*. NeurIPS.

So, D., Le, Q., & Liang, C. (2019). *The evolved transformer.* Paper presented at the 36th International Conference on Machine Learning, Proceedings of Machine Learning Research. Long Beach, CA, USA.

Thornton, C., Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2013, August 11-14). Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. In R. Ghani, T. E. Senator, P. Bradley, R. Parekh, & J. He (Eds.), *The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 847-855). KDD. https://doi.org/10.1145/2487575.2487629

Torralba, A., & Efros, A. A. (2011, June 20-25). Unbiased look at dataset bias. In F. Durand, Y. Weiss, W. T. Freeman, & A. Levin (Eds.), *Conference on Computer Vision and Pattern Recognition (CVPR) 2011* (pp. 1521-1528). IEEE. https://doi.org/10.1109/CVPR.2011.5995347

Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U., & Dwivedi, Y. K. (2017). Open data and its usability: An empirical view from the citizen's perspective. *Information Systems Frontiers*, 19(2), 285-300. https://doi.org/10.1007/s10796-016-9679-1

Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big Data*, 3(1), 9. https://doi.org/10.1186/s40537-016-0043-6

West, S. G., Curran, P. J., & Finch, J. F. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56-75). SAGE Publications.

Wu, J. L., Chang, P. C., Tsao, C. C., & Fan, C. Y. (2016). A patent quality analysis and classification system using self-organizing maps with support vector machine. *Applied Soft Computing*, 41, 305-316. https://doi.org/10.1016/j.asoc.2016.01.020

Yun, S., & Hyun, J. (2019). An analysis of open data policy in Korea: Focused on national core data in open data portal. *Korean Public Management Review*, 33(1), 219-247. https://doi.org/10.24210/kapm.2019.33.1.010

Zhang, H., & Zhu, T. (2022). Stacking model for photovoltaic-power-generation prediction. *Sustainability*, 14(9), 5669. https://doi.org/10.3390/su14095669

Zhang, J., Dawes, S. S., & Sarkis, J. (2005). Exploring stakeholders' expectations of the benefits and barriers of e-government knowledge sharing. *Journal of Enterprise Information Management*, 18(5), 548-567. https://doi.org/10.1108/17410390510624007

Zöller, M. A., & Huber, M. F. (2021). Benchmark and survey of automated machine learning frameworks. *Journal of Artificial Intelligence Research*, 70, 409-472.

Zoph, B., & Le, Q. V. (2017). *Neural architecture search with reinforcement learning*. https://arxiv.org/abs/1611.01578

Call for Paper

Journal of Information Science Theory and Practice (JISTaP)

We would like to invite you to submit or recommend papers to Journal of Information Science Theory and Practice (JISTaP, eISSN: 2287-4577, pISSN: 2287-9099), a fast track peer-reviewed and no-fee open access academic journal published by Korea Institute of Science and Technology Information (KISTI), which is a government-funded research institute providing STI services to support high-tech R&D for researchers in Korea. JISTaP marks a transition from Journal of Information Management to an English-language international journal in the area of library and information science.

JISTaP aims at publishing original studies, review papers and brief communications on information science theory and practice. The journal provides an international forum for practical as well as theoretical research in the interdisciplinary areas of information science, such as information processing and management, knowledge organization, scholarly communication and bibliometrics.

We welcome materials that reflect a wide range of perspectives and approaches on diverse areas of information science theory, application and practice. Topics covered by the journal include: information processing and management; information policy; library management; knowledge organization; metadata and classification; information seeking; information retrieval; information systems; scientific and technical information service; human-computer interaction; social media design; analytics; scholarly communication and bibliometrics. Above all, we encourage submissions of catalytic nature that explore the question of how theory can be applied to solve real world problems in the broad discipline of information science.

Co-Editors in Chief : Gary Marchionini & Dong-Geun Oh

Please click the "Online Submission" link in the JISTaP website (http://www.jistap.org), which will take you to a login/account creation page. Please consult the "Author's Guide" page to prepare your manuscript according to the JISTaP manuscript guidelines.

Any question? Hyekyong Hwang (managing editor) : jistap@kisti.re.kr

JISTaP
Journal of Information Science Theory and Practice

Information for Authors

The Journal of Information Science Theory and Practice (JISTaP), which is published quarterly by the Korea Institute of Science and Technology (KISTI), welcomes materials that reflect a wide range of perspectives and approaches on diverse areas of information science theory, application and practice. JISTaP is an open access journal operated under an Open Access Policy. See the section on Open Access for detailed information on the Open Access Policy.

A. Originality and Copyright

All submissions must be original, unpublished, and not under consideration for publication elsewhere. Once an article is accepted for publication, all papers are accessible to all users at no cost. If used for other researches, its source should be indicated in an appropriate manner and the content can only be used for uncommercial purpose under Creative Commons license.

B. Peer Review

All submitted manuscripts undergo a single-blind peer review process in which the identities of the reviewers are withheld from the authors.

C. Manuscript Submission

Authors should submit their manuscripts online via AccessON Peer Review Management System Plus (ACOMS^{+}). Online submission facilitates processing and reviewing of submitted articles, thereby substantially shortening the paper lifecycle from submission to publication. After checking the manuscript's compliance to the Manuscript Guidelines, please follow the "Online Submission" hyperlink in the top navigation menu to begin the online manuscript submission process.

D. Open Access and Creative Commons License Statement

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee or authors' permission. Open Access articles are automatically archived in the Korea Institute of Science and Technology Information (KISTI)'s Open Access repository (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

E. Manuscript Guidelines

Manuscripts that do not adhere to the guidelines outlined below will be returned for correction. Please read the guidelines carefully and make sure the manuscript follows the guidelines as specified. We strongly recommend that authors download and use the manuscript template in preparing their submissions.

1. Page Layout :
All articles should be submitted in single column text on standard Letter Size paper (21.59 × 27.94 cm) with normal margins.

2. Length :
Manuscripts should normally be between 4,500 and 9,000 words (10 to 20 pages).

3. File Type :
Articles should be submitted in Microsoft Word format. To facilitate the manuscript preparation process and speed up the publication process, please use the manuscript template.

4. Text Style :
- Use a standard font (e.g., Times New Roman) no smaller than size 10.
- Use single line spacing for paragraphs.
- Use footnotes to provide additional information peripheral to the text. Footnotes to tables should be marked by superscript lowercase letters or asterisks.

5. Title Page :
The title page should start with a concise but descriptive title and the full names of authors along with their affiliations and contact information (i.e., postal and email addresses). An abstract of 150 to 250 words should appear below the title and authors, followed by keywords (4 to 6).

Author1
Affiliation, Postal Address. E-mail
Author2
Affiliation, Postal Address. E-mail

ABSTRACT
A brief summary (150-250 words) of the paper goes here.

Keywords: 4 to 6 Keywords, separated by commas.

6. Numbered Type :

1. INTRODUCTION

All articles should be submitted in single column text on standard letter size paper (21.59 × 27.94 cm) with normal margins[1. Text should be in 11-point standard font (e.g., Times New Roman) with single line spacing. [1 Normal margin dimensions are 3 cm from the top and 2.54 cm from the bottom and sides.

2. SECTIONS

The top-level section heading should be in 14-point bold all uppercase letters.

2.1. Subsection Heading 1

The first-level subsection heading should be in 12-point bold with the first letter of each word capitalized.

2.1.1. Subsection Heading 2

The second-level subsection heading should be in 11-point italic with the first letter of each word capitalized.

7. Figures and Tables :

All figures and tables should be placed at the end of the manuscript after the reference list. To note the placement of figures and tables in text, "Insert Table (or Figure) # here" should be inserted in appropriate places. Please use high resolution graphics whenever possible and make sure figures and tables can be easily resized and moved.

Figure

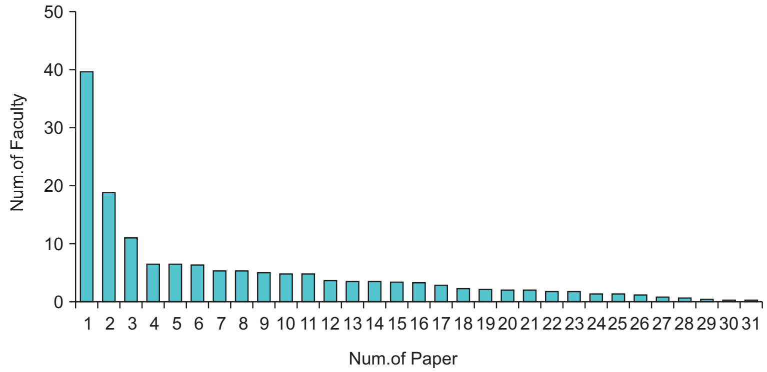

Fig. 1. Distribution of authors over publication count.

Table

Table 1. The title of table goes here

Study	Time period study	Data
Smith Wesson (1996)	1970 - 1995	684 papers in 4 SSCI journals
Reeves [a (2002)	1997 - 2001	597 papers in 3 SSCI journals
Jones Wilson [b (2011)	2000 - 2009	2,166 papers in 4 SSCI journals

[a Table footnote a goes here
[b Table footnote b goes here

8. Acknowledgements :

Acknowledgements should appear in a separate section before the reference list.

9. Conflicts of Interest :

Authors of manuscripts must disclose any potential conflicts of interest at the time of submission. Statements on conflict of interest have no influence on the editorial decision to publish.

10. Citations :

Citations in text should follow the author-date method (authors' surname followed by publication year).
- Several studies found... (Barakat et al., 1995; Garfield, 1955; Meho & Yang, 2007).
- In a recent study (Smith & Jones, 2011)...
- Smith and Jones (2011) investigated...

11. Reference List :

Reference list, formatted in accordance with the American Psychological Association (APA) style, should be alphabetized by the first authors last name.

Journal article
- Author, A., Author, B. & Author, C. (Year). Article title. *Journal Title*, volume(issue), start page-end page.
- Smith, K., Jones, L. J., & Brown, M. (2012). Effect of Asian citation databases on the impact factor. *Journal of Information Science Practice and Theory*, 1(2), 21-34.

Book
- Author, A., & Author, B. (Year). *Book title*. Publisher Name.
- Smith, K., Jones, L. J., & Brown, M. (2012). *Citation patterns of Asian scholars*. Sage.

Book chapter
- Author, A., & Author, B. (Year). Chapter title. In A. Editor, B. Editor, & C. Editor (Eds.), *Book title* (pp. xx-xx). Publisher Name.
- Smith, K. & Brown, M. (2012). Author impact factor by weighted citation counts. In G. Martin (Ed.), *Bibliometric approach to quality assessment* (pp. 101-121). Springer.

Conference paper
- Author, A., & Author, B. (Year, Month Day). Article title. In A. Editor & B. Editor (Eds.), *Conference title* (pp. xx-xx). Publisher Name.
- Lin, H., Sun, B., Wu, J. & Xiong, H. (2016, June 24-26). Topic detection from short text: A term-based consensus clustering method. In B. Yang (Ed.), *2016 13th International Conference on Service Systems and Service Management (ICSSSM 2016)* (pp. 1-6). IEEE.

Online document
- Author, A., & Author, B. (Year). *Article title*. URL.
- Smith, K. & Brown, M. (2010). The future of digital library in Asia. *Digital Libraries*, 7,111-119. http://www.diglib.org/publist.htm.

JISTaP

Journal of Information Science Theory and Practice

KISTI www.kisti.re.kr
66, Hoegi-ro, Dongdaemun-gu, Seoul, Republic of Korea
(T) +82-2-3299-6105 (F) +82-2-3299-6067
https://www.jistap.org