eISSN : 2287-4577
pISSN : 2287-9099
http://www.jistap.org
Vol. 11 No. 1 March 30, 2023

Journal of Information Science Theory and Practice

Journal of Information Science Theory and Practice

JISTaP

01 Information Verification Practices and Perception of Social Media Users on Fact-Checking Services
14 A Comparative Analysis of Research Trends in the Information and Communication Technology Field of South and North Korea Using Data Mining
31 Formulating a Conceptual Model of Digital Service Transformation Based on a Systematic Literature Review
49 Social Media Celebrities in Bangladesh
61 GOMME: A Generic Ontology Modelling Methodology for Epics
79 Development of Ontology for Thai Country Songs

Vol. 11 I No. 1 I March 30, 2023

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 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: Seonhee Lee, 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 and au-thors, permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far 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).

2023 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

Wan-Jong Kim
Korea Institute of Science and Technology Information, Korea

Co-Managing Editors

Seonhee Lee
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
San 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

Hamid R. Jamali
Charles Sturt University, Australia

Boryung Ju
Louisiana State University, USA

Thomas Mandl
Universiät 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

Wolfgang G. Stock
Heinrich-Heine-Universität Düsseldorf, Germany

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

Tsutomu Shihota
St. Andrews University, Japan

2023 Copyright © Korea Institute of Science and Technology Information

JISTaP
Journal of Information Science Theory and Practice

Table of Contents

Vol. 11 No. 1 March 30, 2023
Journal of Information Science Theory and Practice • http://www.jistap.org

Articles

01 Information Verification Practices and Perception of Social Media Users on Fact-Checking Services
Rabby Q. Lavilles, January F. Naga, Mia Amor C. Tinam-isan, Julieto E. Perez, Eddie Bouy B. Palad

14 A Comparative Analysis of Research Trends in the Information and Communication Technology Field of South and North Korea Using Data Mining
Jiwan Kim, Hyunkyoo Choi, Jeonghoon Mo

31 Formulating a Conceptual Model of Digital Service Transformation Based on a Systematic Literature Review
Sawung Murdha Anggara, Agus Hariyanto, Novianto Budi Kurniawan, Arry Akhmad Arman, Suhardi

49 Social Media Celebrities in Bangladesh
Md. Sayeed Al-Zaman

61 GOMME: A Generic Ontology Modelling Methodology for Epics
Udaya Varadarajan, Mayukh Bagchi, Amit Tiwari, M. P. Satija

79 Development of Ontology for Thai Country Songs
Thunyaluk Jaitiang, Malee Kabmala, Wirapong Chansanam

89 **Call for Paper**

90 **Information for Authors**

2023 Copyright © Korea Institute of Science and Technology Information

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

RESEARCH PAPER
J Inf Sci Theory Pract 11(1): 1-13, 2023

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2023.11.1.1

Information Verification Practices and Perception of Social Media Users on Fact-Checking Services

Rabby Q. Lavilles*
College of Computer Studies, Mindanao State University-Iligan Institute of Technology, Andres Bonifacio Avenue, Iligan City, Philippines
E-mail: rabby.lavilles@g.msuiit.edu.ph

Mia Amor C. Tinam-isan
College of Computer Studies, Mindanao State University-Iligan Institute of Technology, Andres Bonifacio Avenue, Iligan City, Philippines
E-mail: miaamor.catindig@g.msuiit.edu.ph

Eddie Bouy B. Palad
College of Computer Studies, Mindanao State University-Iligan Institute of Technology, Andres Bonifacio Avenue, Iligan City, Philippines
E-mail: eddiebouy.palad@g.msuiit.edu.ph

January F. Naga
College of Computer Studies, Mindanao State University-Iligan Institute of Technology, Andres Bonifacio Avenue, Iligan City, Philippines
E-mail: january.febro@g.msuiit.edu.ph

Julieto E. Perez
College of Computer Studies, Mindanao State University-Iligan Institute of Technology, Andres Bonifacio Avenue, Iligan City, Philippines
E-mail: julieto.perez@g.msuiit.edu.ph

ABSTRACT

This study determines how social media users (SMUs) verify the information they come across on the Internet. It determines SMUs' perception of online fact-checking services in terms of their ease of use, usefulness, and trust. By conducting a focus group discussion and key informant interviews, themes were derived in determining fact-checking practices while a survey was further conducted to determine such perceived ease of use, usefulness, and trust in fact-checking services. The thematic analysis revealed major information verification practices, such as cross-checking and verifying with other sources, inspecting comments and reactions, and confirming from personal and social networks. The results showed that SMUs considered fact-checking services easy to use. However, a concern was raised about their usefulness stemming from the delayed action in addressing the information issues that need to be verified. As to perceived trust, it was found that SMUs have reservations about fact-checking services. Finally, it is believed that fact-checking services are expected to be credible and need to be promoted to mitigate any form of fake news, particularly on social media platforms.

Keywords: fact-checking, social media users, fake news, user perception, misinformation, disinformation

Received: September 8, 2022 **Revised:** November 20, 2022
Accepted: December 15, 2022 **Published:** March 30, 2023

***Corresponding Author:** Rabby Q. Lavilles
https://orcid.org/0000-0001-9798-543X
E-mail: rabby.lavilles@g.msuiit.edu.ph

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far 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.

© Rabby Q. Lavilles, January F. Naga, Mia Amor C. Tinam-isan, Julieto E. Perez, Eddie Bouy B. Palad, 2023

1. INTRODUCTION

Social media has become a medium for information sources on important issues such as health, parenting, political issues, and other important information that affects individuals and society. It influences one's behavior (Tham et al., 2020), opinions, and attitudes, particularly among the youth (Ida et al., 2020; Karamat & Farooq, 2016). In particular, young adults favor social media as a source of news (Nazari et al., 2022). While it has become a source of useful information, it is also considered a source of what is commonly referred to as "fake news." The rise of users and the continuing sharing of information on the platform is also proportional to the number of incorrect or manipulated content sources on the platform (Thurman, 2018). Sources of this content come from independent sources, individuals, or partisan-supported accounts, which raise the issue of biases in mainstream media. This may contribute to the tendency of social media users (SMUs) to share news from independent, non-mainstream sources rather than the mainstream press (Nekmat, 2020).

Misinformation and disinformation are two closely-related concepts popularly linked to "fake news" in social media. The former pertains to inaccurate information distributed unintentionally or without manipulative intent, while the latter is focused on "information that is false and deliberately created to harm a person, social group, organization or country" (Moscadelli et al., 2020). Other related terms include alternative facts, misinformation, media manipulation, and propaganda (Jahng, 2021). The spread of this type of information uses social networks or messaging services. With the limited regulation within this type of online environment, monitoring and control rely only upon its users. New emerging and unfamiliar SMUs tend to fall victim to this fake news. For example, a recent study on social media shows that older adults tend to share fake news articles to a greater extent compared to younger age groups (Guess et al., 2019).

One of the actions taken to address the spread of "fake news" in social media is to utilize fact-checking services. Fact-checking service is "the practice of systematically publishing assessments of the validity of claims made by public officials and institutions with an explicit attempt to identify whether a claim is factual" (Walter et al., 2020). For example, BBC Reality Check has been looking at some of the claims made during the debate about the 2019 UK election. These programs are of interest to the public domain because they provide analyses and assessments of information spreading on social media platforms. It has also been adopted by mainstream media as part of reporting, such as the *New York Times*, and has been recognized through journalism awards (Graves, 2013). However, recent fact-checking services include content beyond political issues, such as health-related occurrences during the pandemic (Ceron et al., 2021), to name a few.

Against this backdrop, this study intends to determine two aspects that affect SMU and fake news. First, it identifies how SMUs verify the information they encounter on social media. Secondly, it determines the perceived ease of use, perceived usefulness, and perceived trust in fact-checking services. The first objective intends to understand the practices of information verification, while the other aims to understand fact-checking services from the SMU perspective. Knowing information verification practices and the perception of SMU will support the purpose of having fact-checking services as a response to fake news propagation. Having data about the perception of users can also provide insights into the effectiveness of fact-checking services. We posit that the public must be able to appreciate and have a positive attitude toward fact-checking services to be able to fully rely on them.

2. LITERATURE REVIEW

While studies on fact-checking services have been found to influence the perception of individuals, their effectiveness in addressing the proliferation of "fake news" has mixed results. Nonetheless, indicators of effectiveness have been identified, including format and tone (Young et al., 2018), warnings and tags (Clayton et al., 2020), and the use of messages such as graphical elements (Walter et al., 2020). These fact-checking graphical elements resulted in an unexpected result. For example, integrating visuals into fact-checking messages tended to be less effective than those that did not integrate them (Walter et al., 2020). These studies provide evidence that the presentation of information lacked the element for SMUs to assert the truthfulness of the information they have seen in social media feeds.

Integrating indicators in social media shapes individual views. Rhodes (2022) shows how social media conditions are less critical of information as an effect of filter bubbles and echo chambers. While the integration of warnings or flags in social media decreases belief in the accuracy of headline news, the need to fight fake news proliferation persists (Clayton et al., 2020). Various recommendations for further studies were also mentioned, including an exploration of fact-checking services and social media. For

 https://doi.org/10.1633/JISTaP.2023.11.1.1

instance, Chung and Kim (2021) suggested that measuring how participants perceive fact-checkers would provide a better understanding of fact-checking information.

Extant literature about fact-checking has shown inconsistent results regarding its benefits or effectiveness to users. An experiment about the credibility of perception regarding fact-checking labels on news memes showed that fact-checking labels do not seem beneficial to the credibility of the perception (Oeldorf-Hirsch et al., 2020). On the other hand, to reduce belief in misinformation and false stories on Facebook, strategies on tagging were deployed. It was found that "Rated False" tags are more effective than "Disputed Tags" (Clayton et al., 2020).

External fact-checking organizations were also considered effective in addressing misinformation. Amazeen et al. (2018) concluded that "those who saw a correction were significantly more accurate in their assessment of the controversial statement than those who did not see a correction." Nyhan and Reifler (2015) argued that those participants who were shown fact-checks were able to correctly identify factual information. Similarly, in an experiment on warning labels for false news, it was also concluded that they were effective in reducing the tendency for users to share the said news (Mena, 2020).

A flurry of recent research has centered on recognizing and minimizing the danger of fake news stories spreading to social media sites. The work of Babaei et al. (2021) focuses on how users perceive truth in viral news stories. The researchers conducted online user surveys asking people to rapidly assess the likelihood of news stories being true or false, and to quantify to what extent users can recognize (perceive) a news story's accurate truth level obtained from fact-checking sites like Snopes and Politifact. In 2018, an article by Cunha et al. (2018) quantitatively analyzed how the term "fake news" is being shaped in the news media. The article provides a closer look at how the word "fake news" is influenced by newspapers and magazines around the world – a relevant social trend linked to disinformation and deception and encouraged by the rise of the Internet and online social media. The researchers studied this expression's interpretation and conceptualization through quantitative analysis of a broad corpus of news reported in 20 countries between 2010 and 2018, and supplemented the research with data collected from online search queries that help determine how public interest in the term "fake news" has evolved in different places over time and the ideas around it. These findings extend the definition of the use of the term "fake news" to help understand and describe this relevant social trend correlated with disinformation and deception more accurately.

While services for fact-checking and verification in social media have increased to counter fake news, little research has investigated how SMUs perceive the programs as a means to address the proliferation of misinformation (Brandtzaeg et al., 2018). The findings suggest that while young journalists are largely unfamiliar with or ambivalent about such services, they consider them potentially useful in the journalistic investigative process. A contrast between the opinions of journalists with those of SMUs indicates equally ambivalent SMUs. Several highlighted the value of such services, while others conveyed intense mistrust. The journalists, however, showed a more nuanced perspective, both finding such services to be potentially useful and hesitant to blindly trust one service. The researchers suggested development techniques to make online fact-checking services more efficient and accurate.

Despite the effort of fact-checking organizations, public trustworthiness with fact-checking sites, including Snopes, FactCheck.org, and StopFake, was still negative (Brandtzaeg & Følstad, 2017). Moreover, a discussion about a post on fact-checking services showed that readers have mixed reactions and are considered skeptical of fact-checks. However, interviews with journalists showed a more positive orientation toward learning information about misinformation (Brandtzaeg et al., 2018). The challenge is that human fact-checkers are said to have difficulty keeping up with the amount of misinformation and the pace at which it spreads. This challenge gives automated fact-checking systems an incentive. On the other hand, fact-checking technology is falling behind because there is no current program that does automated fact-checking. Today's professional fact-checkers conduct their research tirelessly as an art, following good data practices and investigative journalism (Manolescu, 2017).

3. CONCEPTUAL FRAMEWORK

The concepts adopted in this study were anchored on the technology acceptance model of Davis (1989) and perceived trust by Roca et al. (2009). In addition, fact-checking practices of SMUs were analyzed as supporting data about their perception of fact-checking services.

Fig. 1 shows the conceptual framework of the study, which illustrates the concepts used to determine the general perception of SMUs on fact-checking services. The first approach deals with determining how SMUs verify the information they encounter on social media plat-

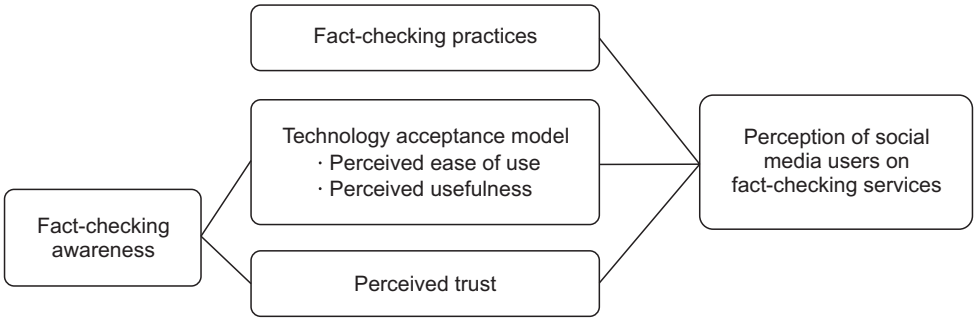


Fig. 1. Conceptual framework.

forms. The practices were derived from emergent themes from the interviews and focus group discussions (FGDs) of SMUs. The second approach adopted the concept of Davis (1989) on the perceived ease of use and perceived usefulness of fact-checking services' social media posts and their corresponding websites. Moreover, perceived trust is a concept that determines how trustworthy the fact-checking services are from the perspective of the SMU. We adopted the concept of Roca et al. (2009) that defined trust as a concept "based on the rational appraisal of an individual's ability and integrity, and on feelings of concern and benevolence."

As perceived by SMU, ease of use is a technology acceptance construct that pertains to the simplicity of navigating, exploring, and getting information from websites or social media posts. It is about how easily the system facilitated arriving at a goal or task. According to the technology acceptance model of Davis (1989), perceived ease of use leads to the perceived usefulness of the artifact. Perceived usefulness is defined by Davis (1989) as the "degree to which a person believes that using a particular system would enhance his or her job performance." Given these constructs, it can be interpreted that the more agreeable the SMUs are in terms of the perceived ease of use and perceived usefulness, the greater chance that they will use the fact-checking services. This idea will support the adoption of SMU in fact-checking services.

4. METHODOLOGY

4.1. Context of the Study

In the Philippines, social media is an influential platform with an average daily usage of 4 hours and 12 minutes (Gonzales, 2019). The country is also considered one of the top SMUs worldwide, comprising more than 73 million users in early 2020 (Kemp, 2020), or around 68% of the country's total population. These users produce and share content in their timelines which include news and current affairs information circulating on the platform. Due to this massive usage of social media, unfortunately, the country faces an active spread of misinformation, disinformation, and related phenomenon in social media as a production network to promote political-related information (Ong & Cabañes, 2018). In a way, social media has become a source or a platform for news and even for political engagement (David et al., 2019), which other people would think is the reason why the platform has been viewed as influencing political matters (Zhuravskaya et al., 2019). Another important thing to note is that during the COVID-19 outbreak, it was observed that vaccine hesitancy was highly influenced by fake news based on previous vaccine cases in the country (Yu et al., 2021).

Among other media organizations in the Philippines, Rappler and VeraFiles were tasked by Facebook as third-party fact-checkers (Magsambol, 2018). Rappler is an online news organization that has been accessible online since 2012. According to its website, it "is composed of veteran journalists trained in broadcast, print, and web disciplines working with young, idealistic digital natives eager to report and find solutions to problems." VeraFiles, on the other hand, is a non-stock, non-profit organization founded in 2008. As stated on its website, it is "published by veteran Filipino journalists taking a deeper look into current Philippine issues." Compared to the majority of the news agencies in the country, these two media organizations are relatively new.

4.2. Data Collection

A FGD was initially conducted to identify information verification practices as well as to solicit from the FGD respondents their level of awareness towards available fact-checking services in the Philippines. To recruit such respondents, purposive sampling was employed. An announcement was initially posted in several conspicuous areas around the university inviting active SMUs interested in participating in the research on fact-checking. Out

of those who responded, the researchers then purposely selected and grouped them into five or six groups based on varied criteria. It is important to note that not all of those who responded to the initial invitation were able to show up within the scheduled time set by the researchers. Such a phase gathered a total of 121 respondents. Table 1 shows the profile of the FGDs conducted.

Based on the list of respondents, individual interviews were then conducted either through face-to-face, online chat, or personal voice call methods. While the FGD was used to capture the shared experiences of SMUs, the individual interviews were intended to capture and confirm individual cases that might not be shared in a FGD. In a way, this method was used to avoid the feeling of embarrassment among the participants, who may be accused of being a sharer of fake information. Table 2 shows the summary of the profile of the participants.

The guiding research questions during the conduct of

Table 1. Focus group discussions (FGDs) conducted

	Number of participants	Average age	Date	Method of data collection
FGD1- University students	5	19	February 20, 2020	Face to face
FGD2- Students and young professionals	5	25	March 30, 2020	Online FGD (via Messenger)
FGD3- Young professionals	4	30	April 15, 2020	Facebook Messenger Chat Group
FGD4- Mixed participants	4	23	April 27, 2020	Online FGD (via Messenger)
FGD5- Working professionals	4	28	May 10, 2020	Online FGD
FGD6- Working professionals	4	37	Nov to Dec 2020	Facebook (via Messenger)

Table 2. General profile of the participants for individual interviews

Participant	Age	Gender	Educational attainment	Social media usage
P1	55	M	Vocational school	Seldom
P2	45	M	Bachelor's degree	Always
P3	40	F	Doctorate	Always
P4	22	F	College level	Frequent
P5	22	M	College degree	Always
P6	20	M	College level	Always
P7	20	F	College level	Always
P8	19	F	College level	Always
P9	35	F	College graduate	Always
P10	22	M	College graduate	Always
P11	24	F	College graduate	Always
P12	21	M	College graduate	Always
P13	25	F	College graduate	Always
P14	23	M	College graduate	Always
P15	19	G	College level	Always
P16	31	M	College degree	Always
P17	23	G	College degree	Always

the individual interviews as well as the FGDs were as follows:

How do SMUs verify the information posted on social media?

How was their experience in using the fact-checking platform?

The first question was used to facilitate how a SMU verifies the context of the post shown on their timeline. It also intends to capture the process and their experiences in using the platform. The second question intends to solicit qualitative data to support or contradict the user experiences regarding ease of use, perceived usefulness, and perceived trust in fact-checking services. Box 1 shows the guide questions during the interviews.

Box 1

Can you share your opinion and experience in using social media page as a source of information? (How often, when to share, tagging)

How did you determine if the information posted on social media is true or reliable? (use of fact check and other sources, use of search, determinants of fake news, sources of reliable info)

How was your experience in using the fact-checking platform? (ease of use, usefulness, biases)

How did fact-checking services (Rappler/Vera Files) help you?

Can you share a memorable experience about the fact-checking experience?

What do you think are the reasons why people share unreliable news?

What do you like most about fact-checking? (don't like, interesting or not)

Can you suggest anything on how to improve fact-checking?

Anything you want to share.

Although the intention was to record all interviews through a voice recorder, during the actual interview some participants refused the recording due to personal reasons. To address this, field notes were utilized to capture important information shared by the participants. Field notes were handwritten notes to capture the thoughts or important phrases shared by the participants.

For the survey, the study employed convenience sampling. While it was the intention to share and distribute the survey instruments in groups or among students in different schools through personal relations or connections of the researchers in the various regions, when the pandemic hit and eventually affected the way the research was initially conducted, the team opted for convenience sampling, as it was challenging to personally visit the contacts due to travel restrictions and the strict health and community quarantine protocols being implemented. The survey adopted a modified questionnaire by Davis (1989) by changing the items (i.e., CHART-MASTER) of the system into "fact-checking services." Moreover, a five-point Likert scale was likewise adopted instead of the original one to simplify the perception of SMUs. For example, for the Usefulness scale, the options include Agree, Slightly Agree, Neutral, Somewhat Disagree, and Disagree.

A total of 685 responses were gathered from the Google Forms distributed through referrals. The initial survey was conducted in Iligan City, Philippines to test the possible responses of the participants. The survey was then revised to eliminate confusing questions and add additional options on news sources shared on social media.

The questionnaire measured the perceived ease of use, perceived usefulness, and trust in the platform. The data gathering was conducted from February 2020 to November 2020. Considering the strategy of deploying Google Forms as a survey instrument, the study makes no claims that the data is representative of the broader population in terms of SMUs in Mindanao, the southern island of the Philippines. However, although the sample may not represent the broader population, the perception of the users does not necessarily negate the result of the survey.

It should be pointed out that the respondents participated voluntarily in the study. It was as well made clear to the participants that they have the right at any stage to withdraw from the study if needed. Google Forms also contained a voluntary consent tab that the participant must check before proceeding to the survey questionnaire to ensure informed consent. Suffice it to say, this research complied with the University Human Research Ethics Review Process.

4.3. Analysis of Data

For qualitative data from FGDs or individual interviews, the researchers used the thematic analysis method for analyzing the data. The researchers closely studied the data to identify common themes (topics, concepts, and sense trends) that repeatedly came up. Thematic analysis is a good approach to research for finding out something about the beliefs, thoughts, expertise, perceptions, or val-

 https://doi.org/10.1633/JISTaP.2023.11.1.1

ues of people from a collection of qualitative data, interview transcripts, social media profiles, or survey responses (Caulfield, 2019). The primary purpose of the interview and focus group were to determine the process or ways the participants could verify the data posted on social media sites.

For the survey data, a total of 685 respondents participated in the survey. The average age of respondents was 22 years old. Around 46% (318) were males while 364 were female. Four of the respondents preferred not to say about their gender.

For the perceived ease of use, usefulness, and trust, the researchers rehashed the 5-point scale into 3, where slightly agree to agree was grouped and converted as agree, and the same applies to the disagree scale. This scale simplified the number of participants who tended to slightly agree or to agree into one group. However, the average was still computed based on the answers.

5. FACT-CHECKING PRACTICES AND PERCEIVED ACCEPTANCE

Derived from the themes that emerged from the data, SMUs showed several practices that may have implications on how they ensure the veracity of information encountered on social media platforms. It is also important to note that, in general, the characteristics of users in how they describe themselves as active SMUs reflects their fact-checking practices.

5.1. Cross-Checking with Other Sources

Active SMUs who are aware of the widespread fake news on social media tend to cross-check for other references. Cross-checking with other sources is the process of verification where a SMU accesses other relevant references to confirm the factuality of the information on a social media platform. This is usually done using a search engine (e.g., Google) to check for related sources or if someone has pointed it out as fake news or misinformation. For instance, FGD2.2 said, "I now use Google Scholar and other authorized and trusted websites." By checking with other sources, SMUs can confirm if the information encountered in social media is accurate.

5.2. Verifying Source and Account Credibility

Account verification and account credibility pertain to accounts recognized by their organization or individual as their official social media page. The Department of Health's official account for health-related information, especially during COVID-19, is considered a credible and verified source. When validating a certain post, some users trace the source of the account. When certain posts catch SMUs' interest, they also tend to check the credibility of the source or account. This is done by checking the source, profile, and previous posts of the account. For example, when news about a death of a known or famous person pops up the tendency is to immediately check who shared it. A participant supports this idea: "and yes also considering the source of the posting. But, more of Twitter as source" (FGD2.3). In the perception of the user, before the mainstream media posts breaking news, someone will share it on Twitter or Facebook.

Most SMUs are also aware of the spread of misinformation on social media platforms. It was shared by one of the participants in one of the FGDs.

> I really think that amidst this pandemic, people should be more cautious in sharing and believing to posts in any form of media. We should be very keen on details including double-checking the account source. (FGD2.3)

There are also cases where a SMU considers mainstream media as a verified source. Even if particular news has been shared by an ordinary user but the source of the shared information is credible, it can be considered to be verified by the user.

> If the shared news is also coming from a mainstream media, I consider it as a credible source of information, so verifying it with other users may not be needed since I considered it verified considering the source information. (P8)

Indeed, checking the source of information is one recommended practice in addressing misinformation or fake news on social media. As Osatuyi (2013) explained, information credibility is operationalized by the ability of the user to verify or confirm based on the cues encountered on social media sites.

5.3. Inspecting Comments and Reactions

Those who lack awareness are influenced by the opinions they encounter in social media posts by other users, especially those they are following (Akdevelioglu & Kara, 2020), or by the common news feeds in the users' accounts. One process that has emerged as one of the verification practices is checking for comments and reactions to

the posted information. Sometimes by reading comments on the post, issues related to the post can also be confirmed if there is associated misinformation. For instance, Participant 4 of FGD 7 shared, "But my usual first action is to read comments. Conflicting views trigger me to verify it on other sites if it hits my interest." This answer indicates that comments under a particular post may trigger the verification process.

One of the first actions taken after reading a particular post is reading comments and reactions to the post, which matters to the user.

> For me, I will look at the comments to see if there are people who contradict the posts, and then I will search for the truth myself to determine if it's true or not and also if I can gain knowledge. (FGD3.2)

> But my usual first action is to read comments. Conflicting views trigger me to verify it on other sites if it hits my interest. Also, I look at who shared the news if I don't know the person. Sometimes my acceptance of the content is based on the person or author who shared it. (FGD 7.4)

Reactions to a post also may suggest something else.

> If I read a serious post and yet the majority of the reactions are laughing, then I will consider it not serious or may be satire-related post. Sometimes reactions are helpful because it also reflects the seriousness of the information posted. (P17)

> [When many shared the post about the information, for me, there is a greater tendency that it's true]. (FGD 2.3)

5.4. Confirming from Personal and Social Networks

Validating information through personal circles also emerges as one of the practices among active SMUs. This pertains to asking about the factuality of the information encountered in social media from families and social media networks. Under this theme, there are three major ways to validate the information.

First, a SMU asks directly to people around when the information is encountered. P12 confirms that sometimes it is easier to directly ask the person around you when you encounter interesting information shared on social media, as he said, "I can easily show it to them," referring to the information he wants to clarify, demonstrating by holding his phone. Indeed, it is easy to allow people to confirm, hoping that they have encountered similar information before. Moreover, P1 and P2 have already established their trust in family members or close relatives, as they both said, "I tend to share and believe information if it comes from a relative." This is another way of confirming information from the inner circles of an active SMU. They also tend to easily believe since these are people they tend to trust.

If people are not around, sharing to a group chat is another strategy for verifying posts among active SMUs. With the presence of private group chats, they become an avenue to share and verify information encountered on the platform. For example, P12 said "Because close connections see group chats, I tend to share the information I encountered with them and ask if the information I have seen is real or fake." The belief is that a limited audience or group can confirm the information easily when none is physically present when the content is encountered.

Lastly, the riskier move is to share the post to get feedback from other SMUs. For instance, "I just always share and share other posts, and I don't think if it's fake or truth" (FGD2.2). They are the type of user who has minimal knowledge about misinformation circulating on social media websites or those who are lackadaisical about it. Take, for example, P1: "I shared so that I will know if it is also true." This unconcerned SMU tends to propagate the post without critically evaluating the post as to whether it is factual or not. In the latter part of the conversation, this experience also taught a lesson to P1.

> During my college days, I really didn't care or think about what will happen next, I just shared some sort of information that I got, and after the day, I was shocked because it reached Cagayan de Oro and other places. I did cry because the sharer of the information told me that she was angry about it. .

5.5. Awareness of Respondents to Fact-Checking Services

Awareness pertains to the exposure of an individual to an advertisement or persuasive message about a certain product or service (Lavidge & Steiner, 1961). Fact-checking services pertain to the consciousness of the individual who either hears or encounters a post or information on a social media platform. In the survey, out of the 685 respondents, 424 (62%) answered "Yes" to the question, "Are you aware of Rappler and Vera Files as fact-checking

 https://doi.org/10.1633/JISTaP.2023.11.1.1

programs of Facebook?" However, when asked if they had experienced using the services, only 316 confirmed that they had. That means only 46% of 685 (or 351) have experience using it. Moreover, specific to Rappler and Vera Files as recognized fact-checking programs, only 22% (153) answered "Yes" when asked the question, "Have you used Rappler or Vera Files to check for facts?"

5.6. Perceived Ease of Use, Usefulness, and Trust in Fact-Checking Services

Out of 316 participants who confirmed their experience using a fact-checking service, 79% (250) agreed that it was easy to use, while 5% (16) did not agree. On a 5-point Likert scale, the mean is 4.06, which can be interpreted as "Agree." This can generally be interpreted as easy to use.

For perceived usefulness, 193 (61%) individuals considered fact-checking services as useful while only 26% disagreed. The average rating is 3.62, which is relatively lower than the average score for perceived ease of use. This result confirms that most SMUs considered fact-checking programs they encountered on social media as useful despite issues of fact-checking in the extant literature (e.g., Brandtzaeg & Følstad, 2017).

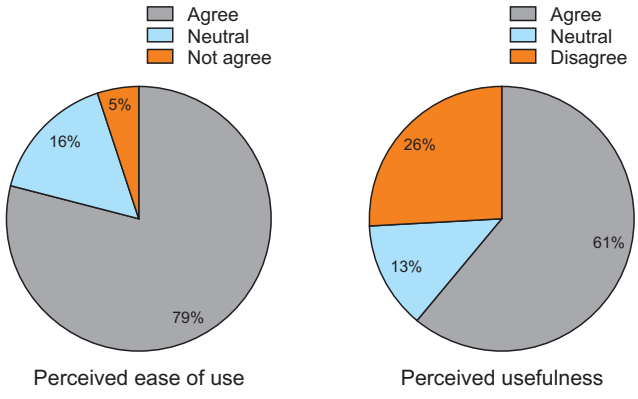


Fig. 2. Perceived ease of use and usefulness of fact-checking services.

Fig. 2 shows the comparison of perceived ease of use (A) and perceived usefulness (B) in terms of percentages. It can be seen that many disagreed on perceived usefulness compared to perceived ease of use of fact-checking services.

SMUs perceived fact-checking services as useful in their presentation of information, particularly labeling images or statements shared on social media. This presents the usefulness of fact-checking services. However, other users commented that the search support or confirmation is limited for the information they encountered in fact-checking posts. In other words, the information the users need is not available on the fact-checking platform. The need to be responsive to information that must be verified in social media cannot cope with the fact-check services' capacity to respond.

The role of trust in fact-checking is critical because it is the basis for which SMUs rely on what they have read, especially relating to political matters. To determine the trust level of SMUs to fact-check services they encountered on social media or websites, we asked them to rate their perceived trust in different mainstream media fact-check entries.

Varied responses were received for the different fact-checking services. In the pilot survey, only Rappler and VeraFiles were included. However, in the comment section, there were suggestions to include other mainstream media because they also include fact-checking services on their respective websites. Thus, we decided to include others as suggested by the respondents. For Rappler, 40% take a neutral stance, while 27% have trust and 33% choose not to trust. VeraFiles got 61% neutral and 25% on trust, and 14% distrust.

In summary, SMUs were ambivalent about the various fact-checking programs they encountered on social media, as shown in Table 3. With most users taking a neutral stance on the fact-checking service except for CNN Philippines Fact-Check, it follows that the fact checking services could not capture an important purpose of a fact-checking service. With the recognized fact-checking

Table 3. Summary of the perceived trust to fact-checking services

Fact-check/Media organization	Trust	Neutral	Distrust	No. of respondents
Fact Check Philippines	40%	52%	7%	459
Rappler Fact Check	27%	40%	33%	469
Vera Files Fact Check	25%	61%	14%	360
CNN Philippines	69%	28%	3%	424

services by prominent social media platforms, more than 50% of the respondents were neutral or answered that they were untrustful of recognized fact-checkers (Rappler and VeraFiles). This is because some participants believed these fact-checkers were one-sided or favored one political party. Similarly, the Pew Research Report on Americans polls about the usefulness and trustworthiness of fact-checking services yields the same perception. The report states that almost half of the participants "tend to favor one side" (Mitchell et al., 2019). This calls for a need to build an unbiased reputation for fact-checkers to improve usefulness and trust with social media participants.

6. DISCUSSION

The information verification practices derived as themes showed the behavior of SMUs when vital information is encountered on social media platforms. Cross-checking and source credibility are practices that can also be observed among journalists, especially for distant sources (Wintterlin, 2020), to verify the information they acquired. This confirms that such a practice can be considered a good exercise for SMUs. On the other hand, inspecting comments and reactions may indicate the need for information literacy among SMUs. The practice of referring to comments may contribute to how misinformation spreads on the platform.

Although good information verification practices are common to journalists and SMU, some practices might contribute to the proliferation of fake news on social media platforms. A socially influenced information verification practice, such as checking for comments and validating information to family, friends, or other users, strengthens the tendency of SMUs to rely on others to confirm or determine the factuality of the information. This finding supports the preference of SMUs who consider news links and recommendations from friends and family compared to those posted by journalists or news organizations. Moreover, social influence has also contributed to the belief of SMU in the information posted online.

Based on the experiences of SMUs with fact-checking websites, the general findings of this study confirmed that they are easy to use. This implies that the majority of the SMU agreed with how easy it is to confirm or determine the truthfulness of the information based on the social media posts and websites of the fact-checking programs, particularly Rappler and VeraFiles. Although the 2019 Reuters Institute's Digital News reports showed that more than half of the 38 countries studied have concerns about the ability to discern what is real and fake on the Internet (Newman et al., 2019), it can be posited that when fact-checking presents information, it can help readers to discern the truthfulness of the information. Perceived ease of use is not a significant concern for fact-checking services or programs. A more pressing issue is the identification of misinformation on social media platforms. With the competing information on the web or social media, the presentation of materials is essential to hasten the user interface. In that way, users can easily engage in the information presented.

Although the survey generally perceived fact-checking as useful, there are aspects where fact-checking services were considered useless. In this study, perceived usefulness as a construct can be considered to the degree to which the service supports users in their goals (Brandtzaeg et al., 2018). It is the degree to which SMUs can discern the factuality of information from the online content posted on social media platforms. In one of the interviews with SMUs about the use of fact-checking websites, the response was, "I do not usually use it to verify because it is easier to search using Google." Although SMUs find the fact-checking website useful, they are less significant during the verification process. This is because information on social media may not be found on fact-checking sites. Instead, a search engine was considered an alternative.

The perception of SMU influenced the adoption and usage of fact-checking services. Brandtzaeg and Følstad (2017) confirmed that positive perception is associated with the usefulness of fact-checkers, while negative perception is associated with untrustworthiness. Therefore, this study's result recommends the need to establish these perceptions. In the United States, political ideologies also differ regarding attitudes to fact-checking. For instance, mainstream liberals are more likely to report using fact-checking websites than do conservatives, who consider it less useful (Robertson et al., 2020). Clearly, it shows that fact-checking perceptions depend on the beliefs and background of an individual or group. Establishing a good reputation and credibility in the perception of SMUs can prove the goal of fact-checking services.

While addressing fake news, misinformation, and related phenomenon is the right thing to do, the adversities are challenging. The velocity and volume in the production and distribution of online information, including fake news, outpaces fact-checkers. With a handful of recognized fact-checkers, it is indeed challenging to address. A fact-checking process is a bottom-up approach that entails

 https://doi.org/10.1633/JISTaP.2023.11.1.1

the need to peruse information as objectively as possible. Fact-checking, therefore, is a vital ingredient in identifying factors for a social information system that promotes the culture of truth in the public sphere. It is alarming, though, that the trust rating of existing fact-checkers in the Philippines is low.

7. CONCLUSION AND LIMITATIONS

This study's key findings pertain to the practices of SMUs and their perception of fact-checking services in the Philippines. With the belief that fact-checking services must be credible, this study provides evidence of the perceptions of SMUs. Perception is important because the usage of fact-checking services depends on their credibility. The significant findings are summarized into major points, as explained in the succeeding paragraph.

First, the themes that emerged as the practices of information verification show two crucial perspectives on the proliferation of "fake news." Verifying through cross-checking is an example of a practice that may limit the spread of fake news. Confirmation from personal and social networks may indicate a possibility of spreading misinformation. Although there are recommended practices, awareness of fact-checking and being critical of the information posted on social media allows one to double-check the information found on social media.

Second, although around 54% of the respondents are aware of fact-checking programs (Rappler and VeraFiles), only 48% of the respondents have experienced using them. In one of the interviews, awareness of fact-checking is a reminder that not all information posted on social media is factual. Awareness itself is an indicator that SMUs are critical of information encountered on social media.

Third, fact-checking services are generally perceived as easy to use and valuable. The practice of summarizing information and labeling it with diagonal text such as "Misinformation," "Needs context," or "False" makes it easier for the user to assess the image or report they encountered on social media. However, because the information that SMUs want to be clarified is sometimes absent from the website, they tend to consider it less valuable than the ease of use.

Lastly, SMUs are ambivalent about perceived trust in fact-checking services provided by Rappler, VeraFiles, and even some mainstream media fact-check posts. This can be attributed to fact-checkers being critical of the government while the president has the majority of the support of the Philippine population based on surveys. SMUs in Mindanao tend to favor government programs because they support the president from Mindanao.

Although this study has shed light on the perspective of SMUs regarding fact-checking services, it also provides some limitations. Due to the limitations of COVID-19, recruitment for the interviews was limited. The use of convenience sampling may limit the generalization of the results. Moreover, the data collected may be reflective only at the time of data collection. However, in general, it provides insights into the perspective of SMUs on fact-checking services as one of the ways to address fake news on social media platforms.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of the Mindanao State University – Iligan Institute of Technology as well as the Consortium on Democracy and Disinformation through De La Salle University for their financial support and guidance in the conduct of research. Moreover, we are thankful to the participants, especially for those who participated during the interview, for trusting the researchers in sharing their experiences.

CONFLICTS OF INTEREST

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

REFERENCES

Akdevelioglu, D., & Kara, S. (2020). An international investigation of opinion leadership and social media. *Journal of Research in Interactive Marketing*, 14(1), 71-88. https://doi.org/10.1108/JRIM-11-2018-0155

Amazeen, M. A., Thorson, E., Muddiman, A., & Graves, L. (2018). Correcting political and consumer misperceptions: The effectiveness and effects of rating scale versus contextual correction formats. *Journalism & Mass Communication Quarterly,* 95(1), 28-48. https://doi.org/10.1177/1077699016678186

Babaei, M., Kulshrestha, J., Chakraborty, A., Redmiles, E. M., Cha, M., & Gummadi, K. P. (2021). Analyzing biases in perception of truth in news stories and their implications for fact checking. *IEEE Transactions on Computational Social Systems,* 9(3), 839-850. https://doi.org/10.1109/TCSS.2021.3096038

Brandtzaeg, P. B., & Følstad, A. (2017). Trust and distrust in online fact-checking services. *Communications of the ACM,*

60(9), 65-71. https://doi.org/10.1145/3122803

Brandtzaeg, P. B., Følstad, A., & Chaparro Domínguez, M. Á. (2018). How journalists and social media users perceive online fact-checking and verification services. *Journalism Practice,* 12(9), 1109-1129. https://doi.org/10.1080/17512786.2017.1363657

Caulfield, J. (2019). *How to do thematic analysis: Step-by-step guide & examples*. https://www.scribbr.com/methodology/thematic-analysis/#:%E2%88%BC:text=Thematic%20analysis%20is%20a%20method,meaning%20that%20come%20up%20repeatedly

Ceron, W., de-Lima-Santos, M. F., & Quiles, M. G. (2021). Fake news agenda in the era of COVID-19: Identifying trends through fact-checking content. *Online Social Networks and Media,* 21, 100116. https://doi.org/10.1016/j.osnem.2020.100116

Chung, M., & Kim, N. (2021). When I learn the news is false: How fact-checking information stems the spread of fake news via third-person perception. *Human Communication Research,* 47(1), 1-24. https://doi.org/10.1093/hcr/hqaa010

Clayton, K., Blair, S., Busam, J. A., Forstner, S., Glance, J., Green, G., Kawata, A., Kovvuri, A., Martin, J., Morgan, E., Sandhu, M., Sang, R., Scholz-Bright, R., Welch, A. T., Wolff, A. G., Zhou, A., & Nyhan, B. (2020). Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. *Political Behavior,* 42(4), 1073-1095. https://doi.org/10.1007/s11109-019-09533-0

Cunha, E., Magno, G., Caetano, J., Teixeira, D., & Almeida, V. (2018). Fake news as we feel it: Perception and conceptualization of the term "fake news" in the media. In S. Staab, O. Koltsova, & D. Ignatov (Eds.), *SocInfo 2018: Social informatics* (pp. 151-166). Springer.

David, C. C., San Pascual, M. R. S., & Torres, M. E. S. (2019). Reliance on Facebook for news and its influence on political engagement. *PLoS One,* 14(3), e0212263. https://doi.org/10.1371/journal.pone.0212263

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

Gonzales, G. (2019). *Filipinos spend most time online, on social media worldwide – Report.* https://www.rappler.com/technology/222407-philippines-online-use-2019-hootsuite-we-are-social-report/

Graves, L. (2013). *Deciding what's true: Fact-checking journalism and the new ecology of news.* (Doctoral dissertation). Columbia University, New York, NY, USA.

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances,* 5(1), eaau4586. https://doi.org/10.1126/sciadv.aau4586

Ida, R., Saud, M., & Mashud, M. (2020). An empirical analysis of social media usage, political learning and participation among youth: A comparative study of Indonesia and Pakistan. *Quality & Quantity,* 54(4), 1285-1297. https://doi.org/10.1007/s11135-020-00985-9

Jahng, M. R. (2021). Is fake news the new social media crisis? Examining the public evaluation of crisis management for corporate organizations targeted in fake news. *International Journal of Strategic Communication,* 15(1), 18-36. https://doi.org/10.1080/1553118X.2020.1848842

Karamat, A., & Farooq, A. (2016). Emerging role of social media in political activism: Perceptions and practices. *South Asian Studies,* 31(1), 381-396.

Kemp, S. (2020). *Digital 2020: The Philippines.* https://datareportal.com/reports/digital-2020-philippines

Lavidge, R. J., & Steiner, G. A. (1961). A model for predictive measurements of advertising effectiveness. *Journal of Marketing*, 25(6), 59-62. https://doi.org/10.1177/002224296102500611

Magsambol, B. (2018). *Facebook partners with Rappler, Vera Files for fact-checking program.* https://www.rappler.com/technology/social-media/200060-facebook-partnership-fact-checking-program

Manolescu, I. (2017). *ContentCheck: Content management techniques and tools for fact-checking.* https://hal.inria.fr/hal-01596563/document

Mena, P. (2020). Cleaning up social media: The effect of warning labels on likelihood of sharing false news on Facebook. *Policy & Internet,* 12(2), 165-183. https://doi.org/10.1002/poi3.214

Mitchel, A., Gottfried, J., Stocking, G., Walker, M., & Fedeli, S. (2019). *Many Americans say made-up news is a critical problem that needs to be fixed*. Pew Research Center.

Moscadelli, A., Albora, G., Biamonte, M. A., Giorgetti, D., Innocenzio, M., Paoli, S., Lorini, C., Bonanni, P., & Bonaccorsi, G. (2020). Fake news and Covid-19 in Italy: Results of a quantitative observational study. *International Journal of Environmental Research and Public Health,* 17(16), 5850. https://doi.org/10.3390/ijerph17165850

Nazari, Z., Oruji, M., & Jamali, H. R. (2022). News consumption and behavior of young adults and the issue of fake news. *Journal of Information Science Theory and Practice,* 10(2), 1-16. https://doi.org/10.1633/JISTaP.2022.10.2.1

Nekmat, E. (2020). Nudge effect of fact-check alerts: Source influence and media skepticism on sharing of news misinformation in social media. *Social Media + Society,* 6(1). https://doi.org/10.1177/2056305119897322

Newman, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, R. K.

(2019). *Reuters institute: Digital news report 2019*. Reuters Institute for the Study of Journalism.

Nyhan, B., & Reifler, J. (2015). Does correcting myths about the flu vaccine work? An experimental evaluation of the effects of corrective information. *Vaccine,* 33(3), 459-464. https://doi.org/10.1016/j.vaccine.2014.11.017

Oeldorf-Hirsch, A., Schmierbach, M., Appelman, A., & Boyle, M. P. (2020). The ineffectiveness of fact-checking labels on news memes and articles. *Mass Communication and Society,* 23(5), 682-704. https://doi.org/10.1080/15205436.2020.1733613

Ong, J. C., & Cabañes, J. V. A. (2018). Architects of networked disinformation: Behind the scenes of troll accounts and fake news production in the Philippines. *Communication Department Faculty Publication Series,* 74. https://doi.org/10.7275/2cq4-5396

Osatuyi, B. (2013). Information sharing on social media sites. *Computers in Human Behavior*, 29(6), 2622-2631. https://doi.org/10.1016/j.chb.2013.07.001

Rhodes, S. C. (2022). Filter bubbles, echo chambers, and fake news: How social media conditions individuals to be less critical of political misinformation. *Political Communication,* 39(1), 1-22. https://doi.org/10.1080/10584609.2021.1910887

Robertson, C. T., Mourão, R. R., & Thorson, E. (2020). Who uses fact-checking sites? The impact of demographics, political antecedents, and media use on fact-checking site awareness, attitudes, and behavior. *International Journal of Press/Politics,* 25(2), 217-237. https://doi.org/10.1177/1940161219898055

Roca, J. C., García, J. J., & de la Vega, J. J. (2009). The importance of perceived trust, security and privacy in online trading systems. *Information Management & Computer Security,* 17(2), 96-113. https://doi.org/10.1108/09685220910963983

Tham, A., Mair, J., & Croy, G. (2020). Social media influence on tourists' destination choice: Importance of context. *Tourism Recreation Research,* 45(2), 161-175. https://doi.org/10.1080/02508281.2019.1700655

Thurman, N. (2018). Social media, surveillance, and news work. *Digital Journalism,* 6(1), 76-97. https://doi.org/10.1080/21670811.2017.1345318

Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2020). Fact-checking: A meta-analysis of what works and for whom. *Political Communication,* 37(3), 350-375. https://doi.org/10.1080/10584609.2019.1668894

Wintterlin, F. (2020). Trust in distant sources: An analytical model capturing antecedents of risk and trustworthiness as perceived by journalists. *Journalism,* 21(1), 130-145. https://doi.org/10.1177/1464884917716000

Young, D. G., Jamieson, K. H., Poulsen, S., & Goldring, A. (2018). Fact-checking effectiveness as a function of format and tone: Evaluating FactCheck.org and FlackCheck.org. *Journalism & Mass Communication Quarterly,* 95(1), 49-75. https://doi.org/10.1177/1077699017710453

Yu, V. G., Lasco, G., & David, C. C. (2021). Fear, mistrust, and vaccine hesitancy: Narratives of the dengue vaccine controversy in the Philippines. *Vaccine,* 39(35), 4964-4972. https://doi.org/10.1016/j.vaccine.2021.07.051

Zhuravskaya, E., Petrova, M., & Enikolopov, R. (2019). Political effects of the Internet and social media. *Annual Review of Economics.* https://doi.org/10.1146/annurev-economics-081919-050239

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

RESEARCH PAPER
J Inf Sci Theory Pract 11(1): 14-30, 2023

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2023.11.1.2

A Comparative Analysis of Research Trends in the Information and Communication Technology Field of South and North Korea Using Data Mining

Jiwan Kim
News Technology Team, SBS A&T, Seoul, Korea
E-mail: kimsabino@sbs.co.kr

Hyunkyoo Choi
Korea Institute of Science and Technology Information, Seoul, Korea
E-mail: hkchoi@kisti.re.kr

Jeonghoon Mo*
Department of Industrial Engineering, Yonsei University, Seoul, Korea
E-mail: J.mo@yonsei.ac.kr

ABSTRACT

The purpose of this study is to compare research trends in the information and communication technology (ICT) field between North and South Korea and analyze the differences by using data mining. Frequency analysis, clustering, and network analysis were performed using keywords from seven South Korean and two North Korean ICT academic journals published for five years (2015-2019). In the case of South Korea (S. Korea), the frequency of research on image processing and wireless communication was high at 16.7% and 16.3%, respectively. North Korea (N. Korea) had a high frequency of research, in the order of 18.2% for image processing, 16.9% for computer/Internet applications/security, and 16.4% for industrial technology. N. Korea's natural language processing (NLP) sector was 11.9%, far higher than S. Korea's 0.7 percent. Student education is a unique subject that is not clustered in S. Korea. In order to promote exchanges between the two Koreas in the ICT field, the following specific policies are proposed. Joint research will be easily possible in the image processing sector, with the highest research rate in both Koreas. Technical cooperation of medical images is required. If S. Korea's high-quality image source is provided free of charge to N. Korea, research materials can be enriched. In the field of NLP, it calls for proposing exchanges such as holding a Korean language information conference, developing a Korean computer operating system. The field of student education encourages support for remote education contents and management know-how, as well as joint research on student remote evaluation.

Keywords: data mining, South Korea, North Korea, information and communication technology, research trends

Received: October 17, 2022 **Revised:** November 29, 2022
Accepted: December 15, 2022 **Published:** March 30, 2023

***Corresponding Author:** Jeonghoon Mo
https://orcid.org/0000-0001-5151-6486
E-mail: J.mo@yonsei.ac.kr

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far 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.

© Jiwan Kim, Hyunkyoo Choi, Jeonghoon Mo, 2023

1. INTRODUCTION

A new concept of the Fourth Industrial Revolution was presented at the annual meeting of the World Economic Forum in 2016 to solve global economic problems. The Fourth Industrial Revolution is a next-generation industrial revolution in which advanced information and communication technologies (ICT) are integrated into the economy and society as a whole, resulting in innovative changes. The core of the Fourth Industrial Revolution is defined as ICT (Telecommunications Technology Association, 2016). ICT increases the productivity of the industry and accelerates changes in the industrial structure. Many countries around the world regard ICT as an important factor in national economic development. Since the 2000s, the Republic of Korea (S. Korea) has designated ICT as a growth engine for national economic development, and has emerged as one of the top 10 global trade powers by expanding ICT infrastructure and investing heavily. The Democratic People's Republic of Korea (N. Korea) is emphasizing the importance of science and technology in building a socialist economic power, with the goal of achieving economic development beyond international sanctions. Referring to the case of S. Korea, N. Korea is pursuing a quantum jump strategy to promote economic development by focusing on ICT. Due to the common point that the two Koreas focus on the ICT field for economic development, ICT can be evaluated as a field with great potential for academic exchange, technology exchange, and economic cooperation. ICT academic exchanges between the two Koreas were relatively active compared to other fields. In 2002, scholars from S. Korea, N. Korea, and China met to jointly publish the *International standard information technology dictionary* (Kim et al., 2015). The joint publication of the dictionary was a good milestone for inter-Korean exchanges. But the exchanges are not sustainable due to political problems. Inter-Korean summits were held three times in 2018, and inter-Korean exchanges were expected to expand further. However, in 2020 N. Korea destroyed the inter-Korean joint liaison office located in the Kaesong Industrial Complex, and peaceful exchanges remain further away. However, in order to prepare for academic exchanges and technological cooperation with N. Korea that can resume at any time, an objective understanding of N. Korea's ICT trends and academic research topics is continuously required. The purpose of this study is to compare and analyze the trends of ICT research between the two Koreas using academic journals, and to provide basic data and cooperation methods to identify areas that can be continuously developed in joint research and exchanges. Comparative analysis of academic research trends between countries informs us of commonalities, differences, and theoretical maturity. It allows us to analyze whether the same topic presents different issues for different conditions, politically and culturally (Choi & Kim, 2011).

2. BACKGROUND AND LITERATURE REVIEW

Research on ICT in N. Korea which is conducted in S. Korea generally analyzes policy and technology trends, by using articles on propaganda from newspapers (*Rodong Sinmun*), new year's addresses and lessons from N. Korean leaders, policies emphasized at the Congress of the Workers' Party of N. Korea, and academic journals as basic data. Lee (2002) analyzed the trend of N. Korea's research by industry using data from the academic journal *Technology Innovation*, published in the 1990s. Since Kim Il-sung emphasized "field support research," N. Korea has focused on research on solving technical problems at production sites. Scientific and technological journals in N. Korea have also become closely related to solving problems at production sites. The number of "technology innovation" research paper posts decreased in 1993. It was evaluated that this was due to the collapse of the socialist countries and the stagnation of foreign trade, which reduced research and development activities at the production site. Despite the decrease in the number of research paper posts, the tendency to post research papers in light industry, agriculture, and fisheries rather increases. This is interpreted as the result of declaring "light industry firstism" and "agricultural firstism" through severe food shortages and consumer disasters during the arduous march. The amount of research papers submitted by companies decreased to less than half, and instead the amount of research papers submitted by research institutes and universities increased significantly. This is evaluated as a result of decreased technological innovation activities of factories and increased field support from universities and research institutes. It was revealed that the increase in the number of research paper posts by universities and research institutes did not lead to an improvement in the quality level of research, and only produced theoretical papers due to a lack of research conditions. Since 1999, when the arduous march ended, it has been confirmed that technological innovation activities have been gradually increasing, and efforts have been made to disseminate academic journals and scientific information.

Choi et al. (2021) extracted science and technology terms from newspaper (*Rodong Sinmun*) articles in the last 10 years (2011-2020) and identified major science and technology interests in N. Korea through term frequency and network analysis. When the subject of the article was divided into 16 fields, the number of articles in the construction field was the largest, and the terms in the mechanical and informatics fields were relatively numerous. N. Korea admitted that it did not meet the goal of the Five-Year Plan for National Economic Development at the 8th Congress of the Workers' Party of N. Korea, held in January 2021. Although it was evaluated that science and technology did not contribute to economic development, the party clarified its attitude to continue its path of emphasis on science and technology. The implementation tasks of the Five-Year Plan strategy are support for scientific research, such as strengthening of science and technology education in elementary and secondary schools, fostering base universities, establishing an academic unification system, expanding graduate programs, and establishing a science and technology distribution network. A plan was also established to prevent waste of research funds as much as possible by criticizing scientists' passivity and establishing a strict deliberation system for R&D. The frequency of keywords related to ICT in Kim Jong-un's reign (2012-2020) was high in the order of "program," "automation," "new century industrial revolution," and "integrated production system." N. Korea is paying attention to the digital economy, and there is a latest trend that emphasizes Artificial Intelligence (AI) and big data, which are the basis of the Fourth Industrial Revolution.

Lim (2019) studied North Korea's response to the Fourth Industrial Revolution and its achievements. The study was conducted through the perception of scholars, media, and a research journal *Economic Research* for seven years (2012-2018). In Lim's paper, N. Korea is attempting a single leap from the third industry to the fourth industry stage based on ICT, and N. Korea is promoting the Fourth Industrial Revolution in a unique way with the precision of computer numerical control (CNC) machines. In addition, N. Korea is promoting differentiation from the previous regime by developing technologies such as AI software, virtual reality (VR), and industrial information. Research institutes under the National Academy of Sciences were newly reorganized to focus on high-tech research and profit-making projects, and a strategy for commercialization. By implementing remote education and improving the education system, scientific and technological talents were fostered. Based on these efforts, some factories have been unmanned and automated. In addition, the number of research paper posts in SCI journals in the field of physics and ICT is gradually increasing. However, it is believed that there will be many restrictions on economic development due to international sanctions caused by nuclear and intercontinental ballistic missile (ICBM) factors, restrictions on technology exchanges in closed societies, and difficulties in importing advanced parts and equipment.

Baxter et al. (2021) conducted a network analysis using a N. Korean scientific journal article to study the structure of the nuclear research community from the beginning of the nuclear program to 2018. The network has shown an increase in the internal density of research cooperation, which shows that N. Korea has improved knowledge transfer over time and solidified its nuclear research community. The study confirmed that other areas of nuclear research increased interconnectivity over time.

Research data related to N. Korea are difficult to access due to the closeness of the N. Korean regime. In particular, in the face of military confrontation between the two Koreas, N. Korea is extremely reluctant to disclose information in the ICT field. It is not easy to obtain specific and objective statistics and up-to-date information. To overcome these limitations, research using N. Korea's open academic journals is a very important means of identifying the field, objectively and transparently. Research using N. Korea's academic journals is being conducted in various fields, and it overcomes the limitations of subjective interpretation and qualitative review by researchers.

3. METHODOLOGY

Scholars who study science and technology in quantitative ways have developed tools for bibliographic methods using text (Leydesdorff, 1987, 1989). Quantitative research has mainly focused on various relations such as co-authorships and co-occurring keywords (Hellsten et al., 2020). In this study, co-occurring keywords are used among quantitative research methods. Keywords were extracted from the title and body of ICT academic journals, and research trends between the two Koreas were compared and analyzed using keyword network analysis. Keyword network analysis is a network-based content analysis method that extracts meaningful words from large text sets, and selects keywords to grasp the context based on semantic connection relationships (Lee, 2018). The ICT fields covered in this study follow the method defined and classified by the OECD. The OECD defines the ICT fields

 https://doi.org/10.1633/JISTaP.2023.11.1.2

as "A combination of manufacturing and services industries that capture, transmit and display data and information electronically" (OECD, 2002). This definition is based on International Standard Industrial Classification (ISIC) Rev. 3, and has the following principles:

> #1. (For manufacturing industries) Must be intended to fulfil the function of information processing and communication including transmission and display; #2. (For manufacturing industries) Must use electronic processing to detect, measure, and/or record physical phenomena or control a physical process; #3. (For services industries) Must be intended to enable the function of information processing and communication by electronic means. (OECD, 2002)

The analysis targets and scope of this study were determined based on the above criteria, and are as shown in Table 1. In the case of S. Korea, among Korea Citation Index (KCI) registered academic journals published in Korean, all ICT academic journals with sections such as "engineering" and divisions such as "electronic/information and communication engineering" were selected as subjects of the study. In the case of N. Korea, *Information Science*, published by Science and Encyclopedia Publishing House (SEPH) and *Information Science and Technology*, published by Central Information Agency for Science and Technology (CIAST) were selected as subjects of the study. As of 2022, there are only two ICT academic journals published in N. Korea. N. Korea does not have any ICT academic journals published in English. Most of the English academic papers were published via co-authors while staying abroad. In the field of computer science, 10 papers were published in 2015 and four papers in 2016 (Choi & Noh, 2017). Compared to S. Korea, academic papers in English are extremely insufficient in quantity. Therefore, the comparison of English academic papers is meaningless.

3.1. Data Collection and Refinement

To build N. Korean paper data, the NK Tech webpage (www.nktech.net), operated by the Korea Institute of Science and Technology Information (KISTI), was crawled with the Python program to build a list of papers to be analyzed. Next, by visiting the Information Center on North Korea under the Ministry of Unification, located in the National Library of Korea, and checking the papers in each journal, keyword data was constructed by extracting

Table 1. Periodic academic journals of ICT between South and North Korea (2015-2019)

Academic journals	Papers	Journal publisher
S. Korea		
Journal of Korea Multimedia Society	757	Korea Multimedia Society (KMMS)
Journal of Broadcast Engineering	396	The Korean Institute of Broadcast and Media Engineers (KIBME)
Journal of the Acoustical Society of Korea	312	The Acoustical Society of Korea (ASK)
Journal of The Korea Institute of Electronic Communication Sciences	798	The Korea Institute of Electronic Communication Sciences (KIECS)
Journal of Korea Institute of Information, Electronics, and Communication Technology	395	The Korea Institute of Information, Electronics, and Communication Technology (KIIECT)
Journal of the Korea Institute of Information and Communication Engineering	1,360	The Korea Institute of Information and Communication Engineering (KIICE)
Journal of Korean Institute of Communications and Information Sciences	1,337	Korean Institute of Communications and Information Sciences (KICS)
N. Korea		
Information Science	1,158	Science and Encyclopedia Publishing House (SEPH)
Information Science and Technology	864	Central Information Agency for Science and Technology (CIAST)

ICT, information and communication technology.

technical nouns from the title and body of the paper. Since the keyword network analysis method aims to extract keywords and analyze relationships between keywords, objective criteria setting and consistency for keyword extraction are very important (Yang, 2019). In the case of N. Korean published papers, there are often no abstracts and selected keywords due to the short content of one to two pages per piece. Therefore, from the introduction, body, and conclusion of the paper, subject-related terminologies that appear more than three times were extracted. In the case of S. Korea, the journals to be analyzed were searched on the webpage (www.kci.go.kr) operated by KCI, and the paper list was downloaded in the form of an Excel file. Terminology was extracted from the title, abstract, and keywords of the papers. The refinement of data is divided into correcting, controlling, and removing (Lee, 2014). First, in the correction stage, all spaces in the compound nouns are attached, and special symbols are unified. Second, different spellings and similar words were unified into one. Among several words with the same meaning, they were unified into one word that appeared statistically more often. Third, among the words with a high frequency of appearance, words representing general concepts such as "research," "method," and "design" were removed. Fourth, the terminologies used in S. Korea and N. Korea are different in terms of notation and form, so a S. Korean-N. Korean-English comparison table was created, and a data list was constructed by matching words with English words.

3.2. Data Analysis

The refined keywords were coded in Excel, and the number of papers and the frequency of keywords by year were analyzed using the Knowledge Matrix Plus ver. 0.80 program (Korea Institute of Science and Technology Information, 2016). It is difficult to analyze the network because the data are too vast. Therefore, the formula 1 was used which was devised by Donohue (1973) and modified by Sun (1992). This formula removes insignificant keywords and selects high-frequency keywords that are meaningful for analysis.

$$n = \frac{-1 + \sqrt{1 + 4D}}{2}$$

D is the total number of keywords, and the value n is the number of high frequency keywords. By checking the number of appearances of the top n-th keyword, a keyword with more than that number was selected as the final keyword to be analyzed. For keyword network analysis, the two-mode network database of "papers and keywords" was converted into a one-mode network database of "keywords and keywords," and this data was analyzed by Gephi 0.9.2 (Bastian et al., 2009) and visualized by R program.

Degree centrality is an indicator of how much one node is connected to another. The combination between keywords becomes a research topic. Keywords with high degree centrality are often used simultaneously with other keywords. This means that there are many research topics directly related to those keywords (Kho, 2014). This paper used Eigenvector centrality, which modified degree centrality. Eigenvector centrality is a concept that extends degree centrality by considering not only the number of connected nodes but also how important the connected nodes are (Bonacich, 2007). The modularity algorithm of the Gephi 0.9.2 program was used (Blondel et al., 2008). Using this algorithm, keywords with high relevance were clustered to classify research topics and to understand the meaning of keywords. Modularity refers to the degree to which a particular node is close to one community and far from another, and is used to distinguish communities from each other in a network. Modularity resolution is derived by adjusting the value from 0 to 1. The number of communities increases as the value converges to 0, and the number of communities decreases as the value increases to 1. The researcher adjusts the resolution value to determine whether it has become an appropriate community. In general, a value of 0.3 or more is considered good clustering (Oh, 2017). The classification of clustering topics was based on S. Korea's ICT R&D technology classification system. The classification system is defined in Article 14 of the Information Communication and Broadcasting R&D Management Regulations of Ministry of Science and ICT Notice No. 2021-23. The ICT R&D technology classification system is divided into four levels (Sections-Divisions-Groups-Classes). In this study, the main topics of clustered communities were directly identified in the papers and the topics were set based on Groups. However, when it was difficult to see the overall picture of the ICT research trend because some subject was divided in detail based on Groups, the subject was defined as Divisions. In order to compare the research trends and interests of the two Koreas, the relationship and characteristics between keywords were expressed as structured images. The network visualization method used the Kamada–Kawai algorithm. The Kamada–Kawai algorithm is characterized by placing a node with high centrality in the center of the image. The size of the node represents degree centrality, and the thickness of the link represents the frequency of

 https://doi.org/10.1633/JISTaP.2023.11.1.2

co-occurring between the two words.

4. RESULTS

4.1. Publishing Papers and Clustering

As shown in Table 2, over the past five years, S. Korea academic papers in the field of ICT have been published 2.6 times more in quantity than in N. Korea. The amount of N. Korea academic papers published annually from 2015 to 2019 remains constant at around 400, but the amount of S. Korea academic papers tends to gradually decrease over time. This phenomenon reflects the tendency of S. Korean researchers to publish papers written in English rather than Korean. In the case of S. Korea, 11 communities were detected as a result of clustering with the modularity resolution value of 0.8. Based on the detected community results, clusters were divided into 11 major research fields of S. Korea's ICT: image processing, wireless communication, sound signal processing, computer/Internet application/security, Internet of Things (IoT), convergence media, autonomous driving/location based service (LBS)/intelligent transport system (ITS), industrial technology, networks, neural network/big data, and energy. In the case of N. Korea, eight communities were detected as a result of clustering with the modularity resolution value of 0.6. Major research fields in N. Korea's ICT were divided into a total of eight fields: image processing, computer/Internet application/security, industrial technology, natural language processing (NLP), neural network/decision-making, wireless communication, student education, and sound signal processing.

As shown in Fig. 1 and Fig. 2, image processing is the subject that occupies the highest percentage of the ICT academic papers of the two Koreas. There are 893 papers (16.7%) in S. Korea and 369 papers (18.2%) in N. Korea. S. Korean academic papers have been studied in the following order: wireless communication, 874 papers (16.3%), sound signal processing, 483 papers (9.0%), computer/Internet application/security, 474 papers (8.9%), IoT, 370 papers (6.9%), convergence media, 344 papers (6.4%), autonomous driving/LBS/ITS, 241 papers (4.5%), industrial technology, 238 papers (4.4%), networking, 214 papers (4.0%), neural network/big data, 192 papers (3.6%), energy, 116 papers (2.2%), and others, 916 papers (17.1%). N. Korean academic papers have been studied in the following order: computer/Internet application/security, 342 papers (16.9%) industrial technology, 331 papers (16.4%), NLP, 240 papers (11.9%), neural network/decision-making, 176 papers (8.7%), wireless communication, 130 papers (6.4%), student education, 113 papers (5.6%), sound signal processing, 64 papers (3.2%) and others, 257 papers (12.7%).

4.2. Characteristics of ICT Research in South Korea

4.2.1. Image Processing

Of the total 5,355 papers in S. Korea, 893 papers (16.7%) were in the field of image processing, accounting for the largest frequency. According to the top keyword of S. Korea's research for five years, the "image" keyword topped the list with 281 times (Table 3). In 2017, the "deep

Table 2. Number of papers and analysis keywords by year in South Korea and North Korea

Year	S. Korea						N. Korea					
	2015	2016	2017	2018	2019	Total	2015	2016	2017	2018	2019	Total
Papers	1,189	1,107	1,083	1,033	943	5,355	412	399	406	408	397	2,022
Total number of keywords (D)	4,968	4,913	4,679	4,668	4,248	17,442[a)]	3,040	2,869	2,750	2,833	3,130	11,594[a)]
Number of high frequency keywords (n)	69.9	69.5	67.9	67.8	64.6	131.5	54.6	53.0	51.9	52.7	55.4	107.1
Number of appearances of the 'n'th keyword	8	7	7	6	6	22	4	4	4	4	4	12
The final keywords number to be analyzed	86	87	74	91	86	134[a)]	87	86	70	80	91	113[a)]

[a)]Deduplication.

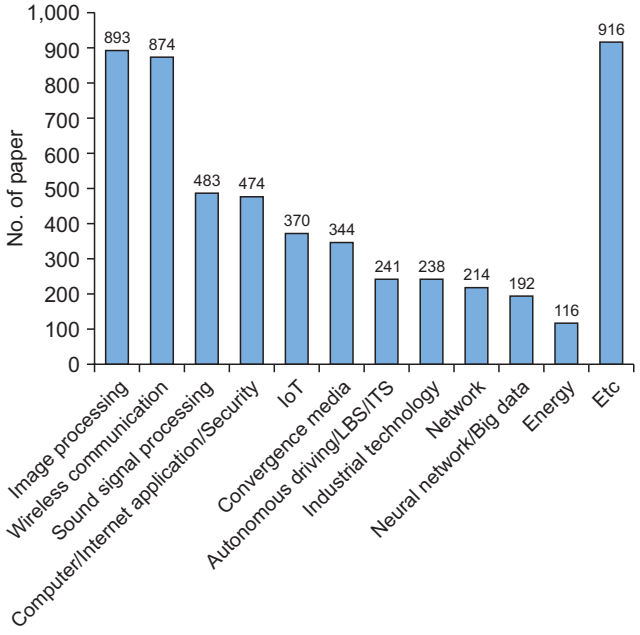


Fig. 1. Number of papers by subject of South Korea's ICT research. ICT, information and communication technology; IoT, Internet of Things; LBS, location based service; ITS, intelligent transport system.

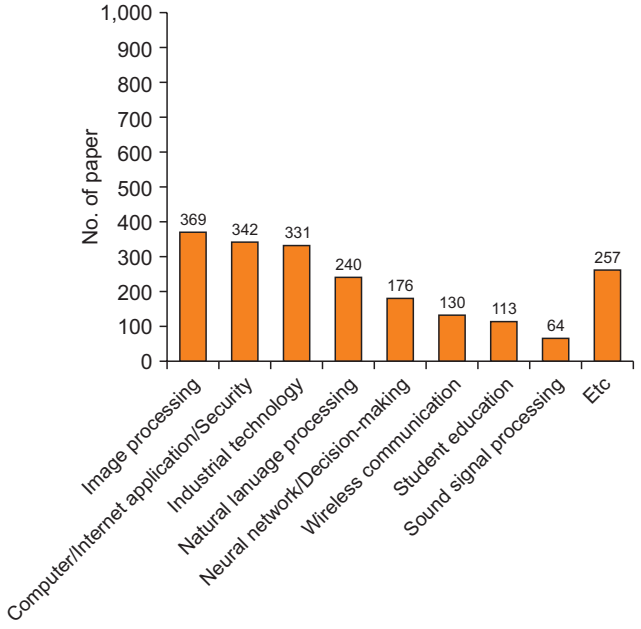


Fig. 2. Number of papers by subject of North Korea's ICT research. ICT, information and communication technology.

learning" keyword ranked high with 18 times, and the frequency increases at a rate close to double every year. It topped the list with 59 times in 2019 (Fig. 3). The "convolutional neural network" (CNN) keyword has also ranked high since 2017, with 47 times in 2018 and 42 times in 2019. Since 2018, the "deep neural network" (DNN) keyword has doubled in frequency in a year. Four keywords, "image," "deep learning," "CNN," and "DNN," have been clustered under the subject of image processing, and the frequency of these keywords has increased rapidly since 2017, indicating that deep learning image processing research is becoming very active. "VR" and "high efficiency video coding" (HEVC) were also clustered under the subject of image processing. Five keywords out of the top 20 keywords were identified as keywords related to image processing (Table 3). "HEVC" is a keyword used intensively in the image processing field, as Eigenvector centrality is the lowest at 0.910 (Table 3).

4.2.2. Wireless Communication

S. Korea's wireless communication industry was selected as a main industry and strategically fostered to overcome the economic crisis in the 1990s and upgrade the industrial structure. For this policy reason, the world's best communication infrastructure has been established so far. Due to the revitalization of the communications industry, the research frequency was 874 (16.3%), accounting for the second highest proportion of research (Fig. 1). Although "multiple-input multiple-output" (MINO) and "orthogonal frequency-division multiplexing" (OFDM) have ups and downs in ranking by year, they have been ranked high in S. Korea's top keywords for five years (Table 3). The "MIMO" and "OFDM" keywords have relatively low Eigenvector centrality values compared to other keywords. This is because "MIMO" and "OFDM" are not widely used apart from in wireless communication due to technological characteristics. "WLAN," "LTE," and "5G" have ups and downs in keyword rankings by year (Fig. 3). This is the result of reflecting the rapid changes and developments in S. Korea's wireless communication.

4.2.3. Sound Signal Processing

The papers clustered on the subject of sound signal processing consist of underwater sound and sound navigation and ranging (SONAR), 142 papers (30%), voice recognition, 56 papers (12%), media sound, 56 papers (12%), architectural sound, 20 papers (4%), ultrasound, 20 papers (4%), and others, 189 papers (39%). "SONAR" has been seen for three consecutive years with 14 to 15 frequency since 2017, and "underwater acoustic communication" is 13 times in 2019 (Fig. 3). However, the top 20 keywords for five years do not include the keyword for sound signal processing (Table 3). This result seems to have been

Table 3. Top 20 keywords for ICT research in South and North Korea for 5 years (2015-2019)

S. Korea				N. Korea			
Rank	Keywords	Freq	Eigenvector centrality	Rank	Keywords	Freq	Eigenvector centrality
1	Image	281	1.000	1	Image	153	0.998
2	IoT	238	1.000	2	Database	71	1.000
3	Deep Learning	122	1.000	3	Pixel	63	0.954
4	CNN	117	1.000	4	Evaluation Index	63	0.883
5	Big Data	95	0.994	5	AHP	58	0.953
6	Noise	91	1.000	6	Student	53	0.988
7	Smart Phone	90	1.000	7	Camera	51	0.916
8	Sensor	86	1.000	8	Neural Network	48	0.991
9	MIMO	86	0.964	9	Genetic Algorithm	48	0.949
10	UAV	84	1.000	10	Corpus	42	0.945
11	Monitoring System	82	1.000	11	Decision Making	42	0.693
12	Vehicle	81	1.000	12	Literature	41	0.912
13	WSN	79	0.993	13	Korean	40	0.946
14	Machine Learning	77	1.000	14	Delphi	38	0.877
15	VR	71	1.000	15	Voice	35	0.955
16	Energy	70	1.000	16	Evaluation	34	0.968
17	User	69	1.000	17	DEA	34	0.837
18	OFDM	68	0.980	18	Histogram	34	0.794
19	HEVC	68	0.910	19	Process	33	0.889
20	Authentication	66	0.994	20	Similarity	31	0.960

ICT, information and communication technology; IoT, Internet of Things; CNN, convolutional neural network; MINO, multiple-input multiple-output; UAV, unmanned aerial vehicle; WSN, wireless sensor network; VR, virtual reality; OFDM, orthogonal frequency-division multiplexing; HEVC, high efficiency video coding; AHP, analytic hierarchy process; DEA, data envelopment analysis.

derived because there is no keyword bias. The topics of research on sound signal processing were not concentrated to one side, and various topics were widely studied.

4.2.4. Computer/Internet Application/Security

474 papers (8.9%) were clustered on computer/Internet application/security subjects. This subject makes for the fourth highest number of research papers. It consists of computer and Internet applications, 230 papers (48%), security and encryption, 219 papers (46%), and Blockchain, 25 papers (5%). Computer and Internet application papers consist of various topics such as computer hardware and software, data communication, and Internet services. Therefore, security and encryption comprise the largest number of papers on a single topic in computer/Internet application/security cluster. Secret key distribution using quantum computing and Fintech simple certification using Blockchain are being covered as the latest topics. "Authentication" is a keyword related to security and encryption, and is ranked in the top 20 with 66 times in Table 3. "Cloud service," "software," and "interface" are also keywords clustered under this subject.

4.2.5. IoT

IoT, an element technology of the Fourth Industrial Revolution, is being used in various ways regardless of

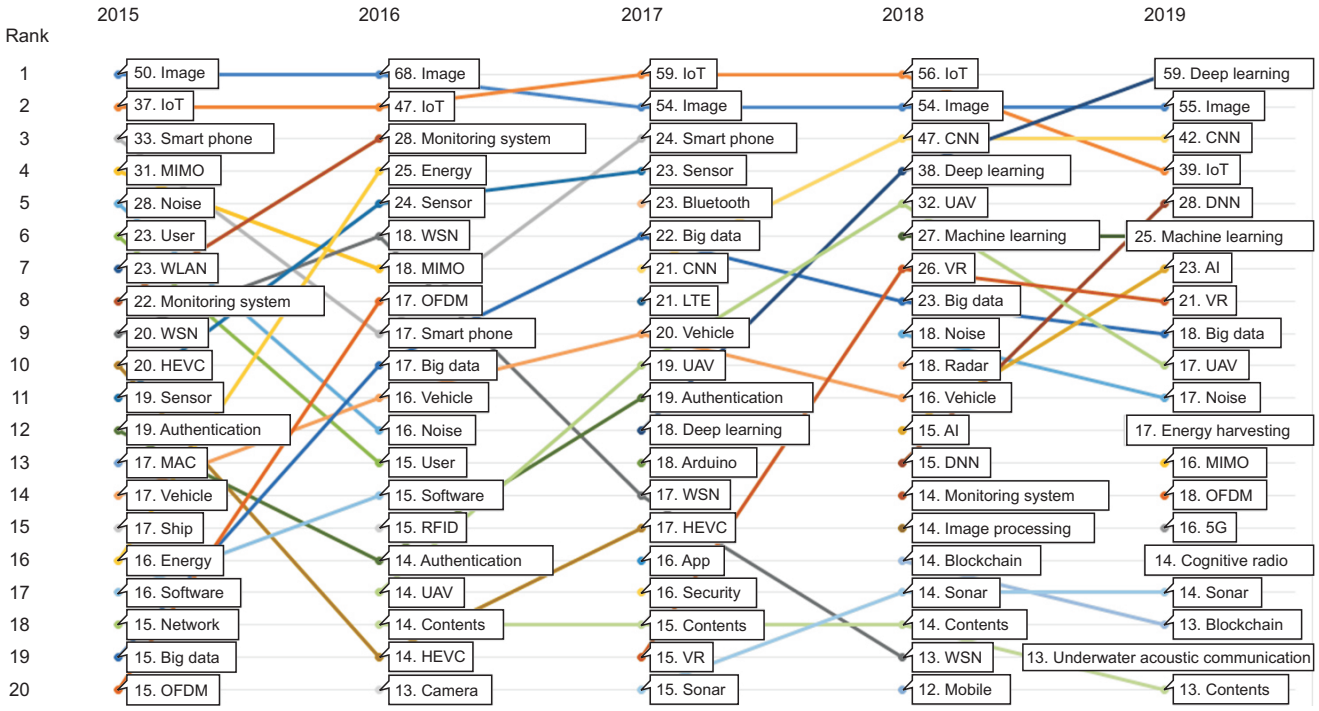


Fig. 3. South Korea ICT research top 20 keywords bump chart by year. ICT, information and communication technology; IoT, Internet of Things; MINO, multiple-input multiple-output; WSN, wireless sensor network; HEVC, high efficiency video coding; OFDM, orthogonal frequency-division multiplexing; UAV, unmanned aerial vehicle; CNN, convolutional neural network; VR, virtual reality; AI, Artificial Intelligence; DNN, deep neural network; MAC, media access control; RFID, radio-frequency Identification.

specific industrial groups. In Table 3, the frequency of the "IoT" keyword is 238 times, which is located at the top level, but the number of clustered papers is relatively small at 370 (6.9%). On the other hand, the Eigenvector centrality of the "IoT" keyword is 1.000. It can be seen that the keyword "IoT" is not limited to a single clustered topic, but is very widely used in other fields such as wireless communication and industrial technology (Fig. 4). The "IoT" keyword was clustered together with the "security" and "network" keywords, and it is confirmed that security of the IoT is a major topic in S. Korea.

4.2.6. Others

Convergence media consists of topics such as development of game and educational content using VR, mobile environments using smart phones, increasing user convenience, and ultra high definition (UHD) next-generation broadcasting systems. Convergence media research is characterized by combining technology and cultural content. Keywords such as "user" and "contents" were clustered into convergence media. The "VR" keyword clustered under the subject of image processing can also be seen in convergence media papers.

In the subject of autonomous driving/LBS/ITS, it was confirmed that there were many papers on V2X, monitoring systems using drones, and location recognition and estimation. In Fig. 3, the "unmanned aerial vehicle" (UAV) keyword has been on the upward curve since 2016, and the "vehicle" keyword has also been steadily ranked except for 2019. "UAV" is ranked as the 10th keyword and "vehicle" is the 12th keyword with the highest frequency (Table 3). Industrial technology has clustered topics such as monitoring systems of smart factories using IoT sensors, smart agriculture using machine learning, integrated railway control, railway wireless networks, and weather forecasting systems. Industrial technology is being applied in many ways regardless of the subject. Wired networks are a field where technical perfection is very high and the activity of research is gradually reduced due to the rapid development of wireless communication technology. The network field was composed of topics such as software defined networking (SDN) and network traffic control and analysis. Neural network/big data is clustered separately from the field of computer/Internet application/security, and is mainly focused on algorithm techniques, performance analysis, and improvement. It has been confirmed

 https://doi.org/10.1633/JISTaP.2023.11.1.2

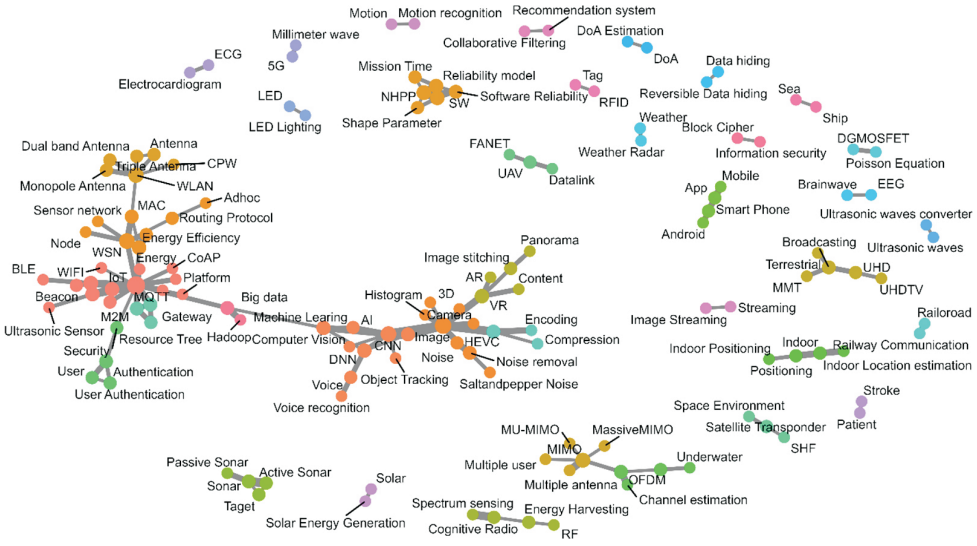


Fig. 4. Visualization of South Korea ICT research keywords network. ICT, information and communication technology; IoT, Internet of Things; CNN, convolutional neural network; MINO, multiple-input multiple-output; UAV, unmanned aerial vehicle; WSN, wireless sensor network; VR, virtual reality; AI, Artificial Intelligence; DNN, deep neural network; OFDM, orthogonal frequency-division multiplexing; HEVC, high efficiency video coding; DoA, direction of arrival; SW, software; MAC, media access control; NHPP, non-homogeneous poisson process; CoAP, constrained application protocol; MQTT, MQ telemetry transport; BLE, bluetooth low energy; MMT, MPEG media transport.

that keywords of the neural network, such as "DNN," "CNN," "deep learning," and "machine learning," are used in all fields of ICT such as image processing, industrial technology, and convergence media, rather than a single field. Energy is a recent emerging field due to carbon neutrality. Research topics such as energy storage system (ESS), solar power generation, smart grids, and wireless power have been clustered.

The ICT research trend in S. Korea has the following characteristics. First, there are many convergence studies in ICT research. The Eigenvector centrality of 14 keywords among the top 20 keywords in S. Korea is 1.000, which has very high connectivity and importance (Table 3). As shown in Fig. 4, the frequency of co-appearance of the top 20 keywords is also high, and it can be confirmed that they are used in various fields. In particular, it can be seen that the frequency of co-appearance in image processing, IoT, and wireless communication fields is high. This shows that ICT research in S. Korea is active in interdisciplinary convergence research. Second, it can be seen that ICT research in S. Korea is very sensitive to trends. As shown in Fig. 3, it can be confirmed that keywords follow the latest trend every year. New keywords are shown every year. "IoT" and "image" keywords remain at the top, but the trend line of many keywords is upward.

4.3. Characteristics of ICT Research in North Korea

4.3.1. Image Processing

Of the total 2,022 papers in N. Korea, 369 papers (18.2%) were in the field of image processing, accounting for the largest frequency. Among the keywords clustered by image processing, "image" was the most common with 153 times, followed by "pixel," 63 times, "camera," 51 times, and "histogram," 34 times (Table 3). Four out of the top 20 were composed of keywords related to image processing. Among the image processing papers, the number of research papers related to object detection is 169 papers (46%). There have been a number of studies related to object detection using CNN and object surface detection. It can be seen that some research efforts are following the latest trends. There were 54 papers related to text or document image processing written by humans. There are 21 papers (6%) related to image encryption, such as secret images and watermarking. There are 33 papers (9%) in the subject of medical imaging, such as computed tomography (CT) or X-ray image processing, research on separating blood cells from images, vein recognition, and lung tissue separation. Ryugyong Dental Hospital in Munsu Street, Pyongyang, developed a bone density measurement and implant design program using X-ray images of patients. Ryugyeong Dental Hospital was selected as the first medical institution to be an exemplary unit for infor-

matization at the National Exhibition of IT Successes 2016 (Rodong Sinmun, 2018).

4.3.2. Computer/Internet Application/Security

There were 342 N. Korean papers (17%) clustered on computer/Internet application/security. This cluster comprises the second highest number of research papers. It consists of security and encryption, 144 papers (42%), web services, 84 papers (25%), software, 59 papers (17%) and databases, 37 papers (11%). In N. Korea, which has military-first policies, the largest number of papers being related to security and encryption shows that there is a high research interest in hacking and system attacks. Among the web service studies, it was confirmed that the ad hoc on-demand distance vector (AODV) protocol, which was actively studied in S. Korea in the early 2000s, was studied in N. Korea in 2015. This proves that N. Korea lags behind S. Korea by more than a decade in web service technology gap. It can be understood that the web infrastructure that operates the web service is still backward. Since the 2000s, N. Korea has been intensively fostering the software industry rather than investing in IT infrastructure or hardware, using it as a means to overcome economic difficulties by overcoming poor manufacturing and acquiring foreign currency. Due to the high technology level of development personnel and relatively low labor costs, software development organizations are dispatched to China, and they are generating profits by developing software modules through overseas contracts (Kim & Lee, 2015). The "database" keyword was ranked second with 71 times, and the Eigenvector centrality showed the largest value at 1.000. Among the computer/Internet application/security keywords, only "database" was ranked in the top 20 (Table 3).

4.3.3. Industrial Technology

The number of papers on industrial technology in N. Korea is the third largest with 331 papers (16.4%). That is not much different from the number of papers in the first and second fields. The reason for the large number of papers related to industrial technology is N. Korea's policy and factory university system. In April 2018, N. Korea announced the "All-out Concentration Policy for Economic Building" at the plenary meeting of the Central Committee of the Workers' Party of N. Korea. With the aim of the socialist economy and the knowledge economy, the strategy planned for unmanned and automated factory production processes and modernized management activities of the

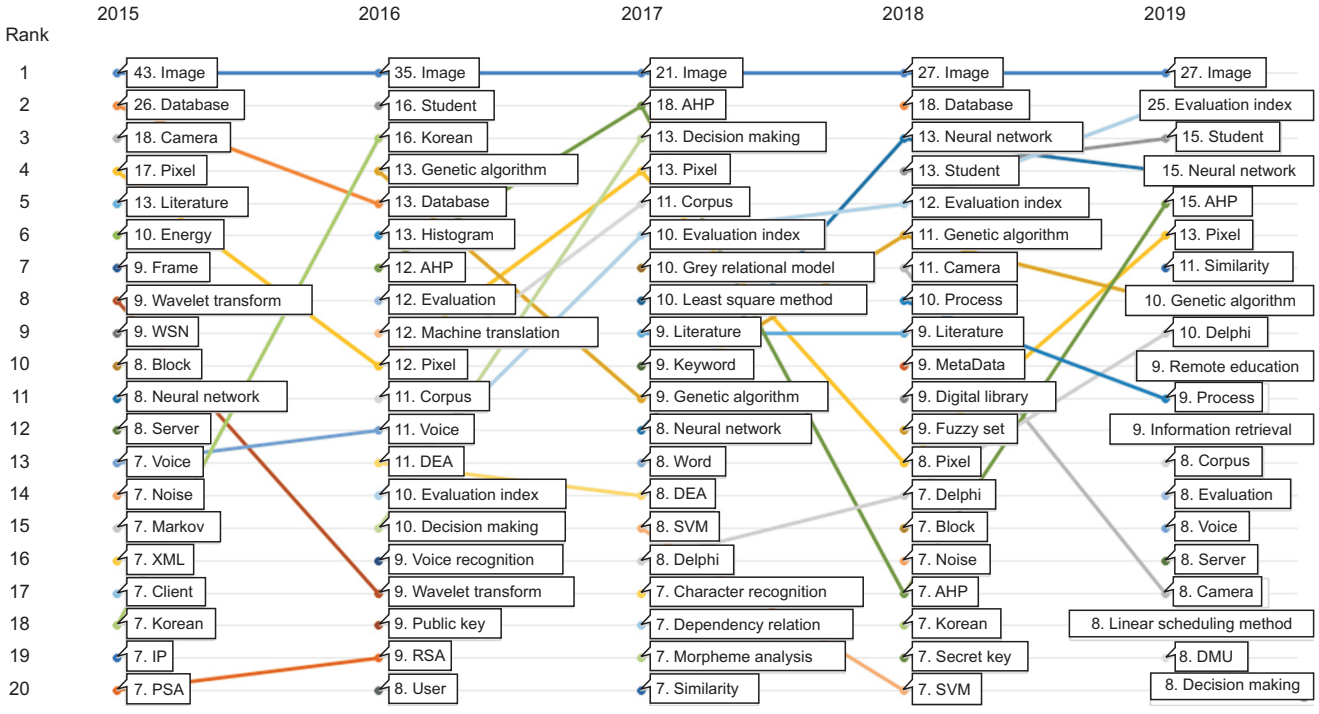


Fig. 5. North Korea ICT research top 20 keywords bump chart by year. ICT, information and communication technology; WSN, wireless sensor network; AHP, analytic hierarchy process; DEA, data envelopment analysis; SVM, support vector machine; DMU, decision making unit.

 https://doi.org/10.1633/JISTaP.2023.11.1.2

enterprise (Park, 2018). A factory university, a higher education institution established in factories and enterprises, is a regular educational institution that allow workers in the relevant enterprise or nearby areas to work and study at the same time (Cho, 2004). In addition to the purpose of training technical personnel and expanding higher education, factory universities also serve to solve technical problems arising from factories and improve productivity with creative ideas as a window for industry-academic cooperative research (Yoo, 2021). Among the keywords clustered in industrial technology are "evaluation index," "analytic hierarchy process" (AHP), "decision-making," "Delphi," "data envelopment analysis" (DEA), and "process" (Table 3). The proportion of research was in the order of technology management, 80 papers (24%), improvements in production processes at factories, 64 papers (19%), technology evaluations, 51 papers (15%), and process management systems, 29 papers (9%). As shown in Fig. 5, the keywords "AHP" and "decision-making" have ups and downs in ranking according to frequency by year. However, the "evaluation index" and "Delphi" keywords have steadily risen from 2016 to 2019. The keyword "process" has emerged rapidly since 2018, which is believed to be related to the policy announced at the plenary meeting of the Central Committee of the Workers' Party of N. Korea.

4.3.4. Natural Language Processing

The frequency of research on NLP in N. Korea was significantly higher than that of S. Korea (Fig. 6). The number of NLP papers in N. Korea is 240 (11.9%), the fourth most frequently studied topic. On the other hand, S. Korea has 35 papers (0.6%) on NLP. In both the absolute number of papers and the ratio of NLP papers to all ICT papers, N. Korea was found to be much higher. Research on machine translation was the most common with 63 papers (26%). Among machine translation, there are 27 Chinese studies and 20 English studies. In addition, QnA, 34 papers (14%) and information retrieval, 28 papers (12%) were followed. The keywords clustered by NLP are "corpus," "literature," and "Korean," which contain three of the top 20 keywords (Table 3). The AI Technology Research Institute at Kim Il-sung University in N. Korea has developed a machine translation system that can translate English, Chinese, Japanese, Russian, German, French, and Spanish (2021). It is said that the translation accuracy is almost at the level of an expert, and the translation speed is three times faster than before (Kang, 2021).

4.3.5. Student Education

Student education clustered into N. Korea's ICT research field is not classified in S. Korea (Fig. 6). There are

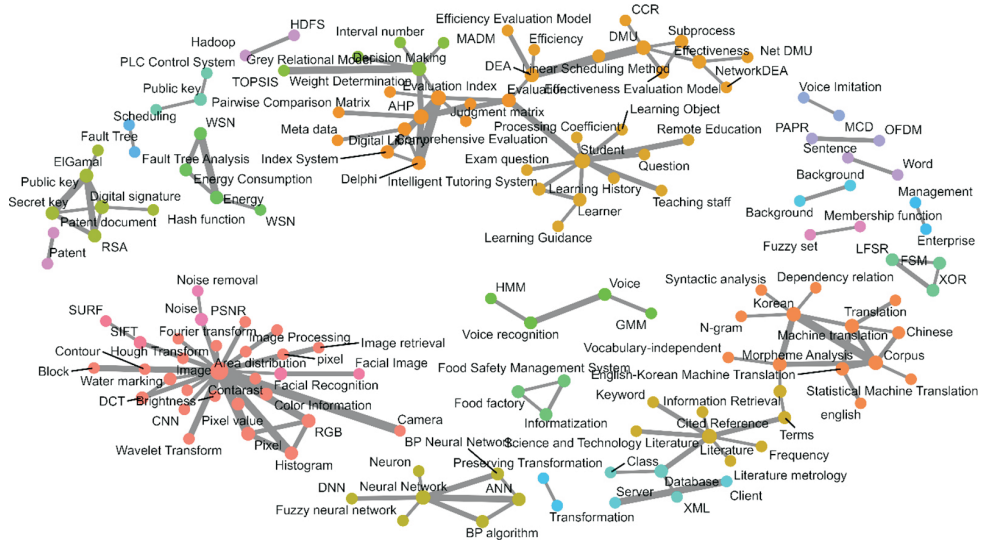


Fig. 6. Visualization of North Korea ICT research keywords network. ICT, information and communication technology; GMM Gaussian mixture model; HMM, hidden markov model; HDFS, hadoop distributed file system; TOPSIS, technique for order of preference by similarity to ideal solution; MADM, multiple attribute decision making; PAPR, peak-to-average power ratio; MCD, Mel distortion distance; LFSR, linear feedback shift registers; FSM, finite state machine; SIFT, scale invariant feature transform; SURF, speeded up robust features; DCT, discrete cosine transform; PSNR, peak signal-to-noise ratio; CNN, convolutional neural network; DNN, deep neural network; ANN, artificial neural network; OFDM, orthogonal frequency-division multiplexing; XOR, exclusive or; DMU, decision making unit; DEA, data envelopment analysis; AHP, analytic hierarchy process; WSN, wireless sensor network; PLC, programmable logic controller.

113 papers on student education. Papers on the subject of student evaluation were the most common with 49 papers (43%), followed by the subject of remote education with 26 papers (23%). The main research contents were learning evaluation of the remote education system and the evaluation of teaching ability. The Kim Jong-un regime set the goal of "Human Resources Development of National Science and Technology," revised the Education Act in 2015 and revised the constitution in 2019, and promoted modernization and informatization of education. As national computer networks are equipped and computers are distributed to schools, various attempts are being made, such as developing educational programs using remote education systems and establishing remote schools. In 2016, the university entrance examination was conducted remotely by computer (Kim & Kim, 2020). N. Korea is using ICT infrastructure to maximize educational effects and make national investments to improve the educational gap and the environment.

4.3.6. Others

The number of papers clustered with neural network/decision-making is ranked fifth with 176 papers (8.7%). The "neural network" keyword is ranked 7th with 48 frequency, and Eigenvector centrality is 0.991 (Table 3). The relatively high Eigenvector centrality of the keyword "neural network" shows that it is used as an application case in various fields rather than in research on the neural network alone. In particular, the "neural network" keyword was most prominent in the field of image processing. "Neural network" is also widely used in research on the industrial technology and NLP. Wireless communication is the second most studied topic in S. Korea with 874 papers (16.3%), but in N. Korea, it is ranked sixth with 130 (6.4%), so its research performance is relatively low. The field of wireless communication consists of wireless sensor network (WSN), 31 papers (24%), OFDM, 28 papers (22%), and code-division multiple access (CDMA), 10 papers (8%). Areas that require large-scale infrastructure investment such as wireless communication are very low in terms of research frequency and ratio compared to S. Korea. Network infrastructure requires huge capital inputs and overseas imports of equipment. However, N. Korea is under sanctions for its nuclear testing and ICBM launches. Therefore, large-scale infrastructure investment is difficult and the number of related research efforts is inevitably small. The number of papers clustered by sound signal processing is the smallest at 64 papers (3.2%). S. Korea's research on sound signal processing has a variety of subjects such as underwater sound and SONAR, voice recognition, and media sound. On the other hand, in N. Korea the number of voice recognition research is 47 papers (73%), accounting for a very high percentage of one subject. As shown in Table 3, the "voice" keyword clustered by sound signal processing was ranked as the only top 20 entry with 35 frequency. In addition to the clustered field, it was confirmed whether there were keywords related to the Fourth Industrial Revolution. At the 7th Congress of the Workers' Party of N. Korea held in May 2016, N. Korea instructed to focus on high-tech technologies, mainly based on information technology, nanotechnology, and biotechnology, and aimed to build a knowledge-based economic powerhouse. At this time, N. Korea used the word "New Century Industrial Revolution," a concept similar to the Fourth Industrial Revolution (Lim, 2019). N. Korea media are promoting that smart homes using IoT and VR were introduced at the "National Exhibition of IT Successes 2018" (Kang, 2018). In the analysis of ICT research papers, two VR papers and one IoT paper could be found. However, keywords related to the Fourth Industrial Revolution such as big data, smart home, smart city, and UAV were not found.

The ICT research trend in N. Korea has the following characteristics. First, there is a higher tendency for independent research than interdisciplinary convergence research. It can be seen that the Eigenvector centrality is relatively low compared to the top 20 keywords in S. Korea. Of the top 20 keywords in S. Korea, 14 have Eigenvector centrality of 1.000, but only one in N. Korea has Eigenvector centrality of 1.000 (Table 3). As such, the connectivity and importance of keywords are low. Second, there was a higher tendency to be influenced by Workers' Party of N. Korea policies than by research topics in the global state-of-the-art technology trends. Researchers are conducting policy-related studies in the Workers' Party of N. Korea, such as industrial technology advancement, remote education, and NLP. There were many studies that directly helped the economy of people's livelihoods.

5. SUGGESTIONS FOR INTER-KOREAN COOPERATION

The N. Korea leader's new year's address has mentioned modernization of the light industry and agricultural and fishery industries using ICT technology since 2014. In addition, ICT business profits are paid to developers to encourage commercialization of ICT technology. Through these contents, it is confirmed that N. Korea has a high

 https://doi.org/10.1633/JISTaP.2023.11.1.2

government interest and demand for economic development using ICT (Seo, 2018). The research situation and the stage of technology development are different for each detailed ICT field. Therefore, it is necessary to check the demand in each ICT field in N. Korea. In addition, specific exchange strategies should be established for each field, rather than unconditionally transplanting the developed S. Korea's ICT technology into the poor environment of N. Korea.

Image processing has a very high research rate in both Koreas and research is actively being conducted. It is expected that joint research will be quickly possible from an equal standpoint in the future. In particular, if medical imaging technology is cooperated in consideration of the poor medical environment in N. Korea, it is expected that the N. Korea authorities will try to cooperate first and favorably. In addition, it is proposed to provide UHD sports videos that S. Korea has the right to broadcast to N. Korea for research. If this happens, research materials in the field of N. Korea's image processing, which focuses on text and document images, can be enriched. Since a sports video has no political content and is not affected by sanctions against N. Korea, research exchanges are likely to proceed very stably.

Computer/Internet application/security research is also expected to be very active as a joint research subject between the two Koreas. This proposes to create a special economic zone for ICT and to build a cloud center that can gather a common database from areas where there is no significant difference in opinion such as language, ancient history, medicine, forests, and fisheries. The database aims at the unique culture of the Korean peninsula. Cloud centers in special economic zones cooperate by investing in S. Korea hardware and capital and dispatching N. Korea management personnel. With the help of international organizations such as UNESCO, there is a high probability of cooperation. In the field of security and encryption, N. Korea has a high national interest and technology. Since security and encryption technologies can be easily used in the military, exchanges and cooperation are very difficult due to sanctions against N. Korea. It may be possible to hold cyber security contests in third countries to evaluate security and encryption skills and to start exchanges by organizing a network of human resources cooperation.

The frequency of industrial technology research in N. Korea and the demand for manufacturing development are high. If capital is supported and production facilities are equipped with new models, it is expected to be of great help in improving manufacturing efficiency and revitalizing the industry. If S. Korea's IoT technology is combined with N. Korea's production process, it will be able to secure remarkable growth engines. However, manufacturing support that may be used as munitions should be excluded. Humanitarian assistance in manufacturing should be selected with experts. Industrial technology is a field that requires voluntary investment and support in the private industry. It is more efficient to conduct industrial technology exchanges through technology transfer of research institutes operated by private companies. If there is a high risk from investment in N. Korea, business operators will not invest in or support it. Prior to private exchanges, the two Korean governments should use the case of the Kaesong Industrial Complex as a teacher and eliminate investment risks in advance.

N. Korea's linguistic theory of Juche (self-reliance) defines language as a product of society and considers it important to protect ethnicity through language (Jeon & Kim, 2019). Language research is active in N. Korea. Because language policy focuses on social meaning and function and considers language an important tool for social development, it is presumed that many studies on NLP using Korean in the field of ICT were influenced by language policies. N. Korea focuses on its own computer operating system and language processing technology without relying on foreign software (Ko et al., 2007). The development of open-source-based software using its own language information is based on the activation of NLP research. Taking advantage of N. Korean language research, we would like to propose exchanges such as promoting Korean language information conferences, developing PC and mobile operating systems based on Korean, and importing Chinese machine translation technology. NLP research is also a field that can be relatively less affected by political situations or sanctions against N. Korea. If we continue to exchange information academic conferences between the two Koreas, which were held at the Korean Language Information Society in the past, it will be a good example for exchange and cooperation in other fields.

N. Korea is very interested in national investment to improve the educational environment and reduce the educational gap by using ICT infrastructure to build a strong country. In April 2020, N. Korea enacted the Remote Education Act and strengthened the legal foundation for Human Resources Development of National Science and Technology (Yang, 2020). Any N. Korean citizen was granted the right to become a student. It was stipulated as the duty of all institutions to ensure remote education by preparing educational facilities and learning places. In or-

der to upgrade educational conditions across N. Korea, a network system was established and remote education was started, and research on education evaluation methods is also steadily being conducted. If the N. Korean education system is opened up, it is predicted that it will be opened to the extent that there is no threat to the socialist system. In addition, since the obligation to remote education is imposed by law, it is judged that there will be sufficient demand for the establishment of an advanced education system. The cooperation plan related to education is as follows: free support of educational contents in S. Korea, teaching know-how to operate Korea National Open University (KNOU) in S. Korea, a joint study on student remote evaluation, a joint development of educational contents using AR and VR, the establishment of a unification education broadcasting station, and so on.

It is expected that private exchanges between the two Koreas in the field of wireless communication will not be easy. There is no possibility that N. Korea will try to cooperate first because the opening of radio wave resources and disclosure of the current status are burdensome for maintaining the N. Korea regime. However, N. Korea's research in wireless communication is poor and its technological completeness is not high, so N. Korea recognizes it as a field that requires investment first. Therefore, it is necessary to first sign an agreement between the governments and promote cooperative projects through incentives such as technology transfer and free support. Specifically, these include installation of 4G infrastructure equipment, maintenance training of wireless network systems, and personnel exchanges.

6. CONCLUSION

This study compared and analyzed research trends in the ICT fields of the two Koreas using academic papers published for five years from 2015 to 2019. In the case of S. Korea, the frequency of research on image processing and wireless communication was high, while N. Korea was in the order of image processing, computer/Internet application/security, and industrial technology. In N. Korea, NLP was found to have a much higher frequency of research than in S. Korea. Student education in N. Korea is a unique field that has not been clustered in S. Korea. Based on these research results, specific exchange strategies for each ICT field should be established. In the fields of image processing and computer/Internet application/security, the research activities of the two Koreas are high, so joint inter-Korean research is expected to be possible quickly and actively. Research in the NLP field makes the most of the advantages of N. Korea's language policy, so that S. Korea can receive and utilize know-how. It is necessary to conduct private business exchanges by confirming the demand for industrial technology. It has been confirmed that there is sufficient demand for the establishment of a remote education system, and educational content exchange is necessary. The wireless communication field needs free infrastructure support after an intergovernmental agreement. The public-private joint plan will bring about the development of inter-Korean ICT exchange and cooperation.

The limitations of this study and the characteristics of N. Korea's ICT research publications are as follows.

First, in N. Korea's ICT research, topics and keywords that can be used for military weapons were not well expressed. This means that the N. Korean authorities are tightening control over ICT. Due to the disclosure of the latest ICT information that can be used as a military weapon, it is difficult to accurately estimate the level of ICT in N. Korea only with publication research keywords. Second, N. Korean papers are composed of one to three pages, and there is a lack of evidence for research methodology and experiments. Therefore, it is difficult to judge the quality level of the studies. Third, it is believed that the policy of the N. Korean authorities and the intention of the N. Korean Kim Jong-un may have strongly affected research content and trends. It is necessary to check the impact of N. Korea's ICT policy on research activities through follow-up studies. Fourth, specific cooperation methods were proposed for each field as described above, but the United Nations currently comprehensively prohibits science and technology cooperation with N. Korea. Therefore, the process needs to find ways to bypass sanctions against N. Korea. A strategy should be established from the exemption field (Choi et al., 2020). This study is meaningful in that it compared the trends of ICT research between the two Koreas through quantitative analysis without relying on the subjective interpretation of the researcher. It is expected to be used as basic data for the establishment of the government's exchange policy and private cooperation.

CONFLICTS OF INTEREST

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

 https://doi.org/10.1633/JISTaP.2023.11.1.2

REFERENCES

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *Proceedings of the International AAAI Conference on Web and Social Media,* 3(1), 361-362. https://doi.org/10.1609/icwsm.v3i1.13937

Baxter, P., Hastings, J. V., Kim, P., & Yim, M. S. (2022). Mapping the development of North Korea's domestic nuclear research networks. *Review of Policy Research,* 39(2), 219-246. https://doi.org/10.1111/ropr.12462

Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment,* P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks,* 29(4), 555-564. https://doi.org/10.1016/j.socnet.2007.04.002

Cho, J. A. (2004). Formation of the manpower development system in North Korean factories. *North Korean Studies Review,* 8(2), 185-212. https://kiss.kstudy.com/thesis/thesis-view.asp?key=2395944

Choi, H. K., Byun, H. M., Noh, K. R., Jung, E. S., & Lee, S. H. (2020). *Research on inter-Korean cooperation measures such as unique technologies and strong technologies that lead the recent change in North Korea and human resources related thereto*. Korea Information Society Development Institute.

Choi, H. K., Byun, H. M., Noh, K. R., Park, J. S., & Hyun, M. H. (2021). *A study on the exploration of North Korea's latest policy issues and analysis of North Korean S&T terminology.* Ministry of Science and ICT.

Choi, H. K., & Noh, K. R. (2017). *A study on the analysis of the international academic paper (SCOPUS) by North Korean scientists: 2007-2016.* Korea Institute of Science and Technology Information.

Choi, Y. C., & Kim, K. K. (2011). A comparative analysis of the research trends in the study of public administration in Korea and the UK. *Korean Comparative Government Review,* 15(1), 1-26. https://doi.org/10.18397/kcgr.2011.15.1.1

Donohue, J. C. (1973). *Understanding scientific literatures: A bibliometric approach*. MIT Press.

Hellsten, I., Opthof, T., & Leydesdorff, L. (2020). N-mode network approach for socio-semantic analysis of scientific publications. *Poetics,* 78, 101427. https://doi.org/10.1016/j.poetic.2019.101427

Jeon, Y., & Kim, J. (2019). A study on improvement measures for crime prevention education of North Korean defectors by security police. *Journal of North Korean Studies,* 44(1), 153-189. https://scholar.kyobobook.co.kr/article/detail/4010027683144

Kang, J. G. (2018). *North Korea introduces big data technology, POS, and IoT.* http://www.nkeconomy.com/news/articleView.html?idxno=599

Kang, J. G. (2021). *North Korea's AI technology research institute develops a machine translation service system called Ryongma.* http://www.nkeconomy.com/news/articleView.html?idxno=4684

Kho, J. C. (2014). *A study on research trend in management of technology using keywords network analysis*. (Doctoral dissertation). Seoul, Sungkyunkwan University.

Kim, B., & Kim, J. (2020). A study on the educational conditions and environmental changes to improve the quality of North Korean education in the Kim Jong-un era. *North Korean Studies Review,* 24(1), 175-205. https://10.0.89.88/kanks.2020.24.1.006

Kim, C. W., Kim, B. S., Seo, S. Y., & Lee, W. S. (2015). *Report of the inter-Korean ICT exchange and cooperation promotion: A study on the standardization of inter-Korean ICT.* Korea Information Society Development Institute.

Kim, J. S., & Lee, C. G. (2015). Development history and implications of North Korea's IT industry. *Science & Technology Policy,* 25(8), 24-33. https://www.stepi.re.kr/skin/doc.html?fn=80aed56950619ec2b158b454216fb6ba&rs=/preview/html/202302

Ko, K, Lee, H, & Jang, S. (2007). North Korea's IT dilemma and dualist strategy: Cases of Internet policy and software industry policy. *Informatization Policy,* 14(4), 139-158. http://journal.kci.go.kr/nia/archive/articleView?artiId=ART001217366

Korea Institute of Science and Technology Information. (2016). *KnowledgeMatrix Plus (KM+) ver. 0.80.* http://mirian.kisti.re.kr/km/km_en.jsp

Lee, C. G. (2002). Technological trends in North Korea from the "technological innovation" paper. *Journal of Scientific & Technological Knowledge Infrastructure,* 11, 38-43. http://www.koreascience.or.kr/article/JAKO200245940321452.do

Lee, S. S. (2014). A content analysis of journal articles using the language network analysis methods. *Journal of the Korean Society for Information Management,* 31(4), 49-68. https://doi.org/10.3743/KOSIM.2014.31.4.049

Lee, S. S. (2018). *Network analysis methods applications and limitations.* Chungram.

Leydesdorff, L. (1987). Various methods for the mapping of science. *Scientometrics*, 11(5-6), 295-324. https://doi.org/10.1007/BF02279351

Leydesdroff, L. (1989). Words and co-words as indicators of

intellectual organization. *Research Policy,* 18(4), 209-223. https://doi.org/10.1016/0048-7333(89)90016-4

Lim, E. C. (2019). North Korean 4th industrial revolution: Strategy, implementation, achievements. *East Asian Studies,* 38(2), 1-35.

OECD. (2002). *Measuring the information economy 2002.* OECD.

Oh, S. W. (2017). *The effect of fan relationship network on business performance of the professional baseball team: Focusing on Facebook fan page.* (Doctoral dissertation). Seoul, Kyung Hee University.

Park, D. M. (2018). The rapid development of science education as a new strategy for the party. *Economic Research,* 2018(4), 33-34.

Rodong Sinmun. (2018, January 30). *Technology-oriented ethos and treatment outcomes - In the business of the staff at Ryugyong Dental Hospital*. Rodong Sinmun.

Seo, S. Y. (2018). Trends and implications of ICT policy in North Korea. *ICT & Media Policy,* 30(18), 1-15. https://kiss.kstudy.com/thesis/thesis-view.asp?key=3915713

Sun, Q. (1992). The division of words of high frequency and low frequency. *Information Science,* 13(2), 28-32.

Telecommunications Technology Association. (2016). *A dictionary of information and communication terms.* http://word.tta.or.kr/dictionary/dictionaryView.do?word_seq=100949-17

Yang, E. H. (2020). *North Korea adopted the "remote education act" institutionalization of "online education."* https://www.news1.kr/articles/?3905670

Yang, J. H. (2019). Analysis of research trends on 'The Journal of Korea Open Association for Early Childhood Education' using the language network analysis methods. *Journal of Korea Open Association for Early Childhood Education,* 24(1), 221-246. https://doi.org/10.20437/KOAECE24-1-09

Yoo, Y. K. (2021). *North Korea, modernizing production with "example factory."* https://www.pressian.com/pages/articles/2021051309585314611#0DKU

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

RESEARCH PAPER
J Inf Sci Theory Pract 11(1): 31-48, 2023

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2023.11.1.3

Formulating a Conceptual Model of Digital Service Transformation Based on a Systematic Literature Review

Sawung Murdha Anggara
Sekolah Teknik Elektro dan Informatika, Institut Teknologi Bandung, Bandung, Indonesia
E-mail: 33221030@mahasiswa.itb.ac.id

Agus Hariyanto
Sekolah Teknik Elektro dan Informatika, Institut Teknologi Bandung, Bandung, Indonesia
Solotech University, Surakarta, Indonesia
E-mail: mr.aghato@gmail.com

Novianto Budi Kurniawan
Badan Pusat Statistik, Jakarta, Indonesia
E-mail: nbudik@bps.go.id

Arry Akhmad Arman
Sekolah Teknik Elektro dan Informatika, Institut Teknologi Bandung, Bandung, Indonesia
E-mail: arry.arman@itb.ac.id

Suhardi*
Sekolah Teknik Elektro dan Informatika, Institut Teknologi Bandung, Bandung, Indonesia
E-mail: suhardi@itb.ac.id

ABSTRACT

Digital service transformation study is a part of research in the field of digital transformation, which is devoted to exploring the transformations that occur in digital service products, which have been intensely explored in recent years to address digital disruption. Several concepts and definitions of digital service transformation have emerged as a result of an approach from the point of view of digital transformation and digital services concepts. This paper is organized to provide a foundational understanding of digital service transformation terminology. This paper uses the systematic literature review method to compile 52 qualified articles from previous studies. We conduct an analysis and synthesis of articles to answer research questions. The results of this study are a descriptive summary of research in the digital service transformation field, determining digital service transformation terminology and components, and also a proposed digital service transformation model to explain the position of transformation in digital service products in the overall transformation process. We construct this model using the findings of previously determined components synthesis.

Keywords: digital transformation, digital service, service transformation, digital service transformation, conceptual model, systematic literature review

Received: April 22, 2022 **Revised:** November 18, 2022
Accepted: December 24, 2022 **Published:** March 30, 2023

***Corresponding Author:** Suhardi
https://orcid.org/0000-0002-0486-3562
E-mail: suhardi@itb.ac.id

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far 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.

© Sawung Murdha Anggara, Agus Hariyanto, Novianto Budi Kurniawan, Arry Akhmad Arman, Suhardi, 2023

1. INTRODUCTION

1.1. Research Background

Digital technology, digital innovation, and digitization have fundamentally changed every social, cultural, and economic aspect of human life. This phenomenon has changed business aspects such as processes, products, services, and relationships (Osmundsen et al., 2018). Digitalization is experiencing dynamic development along with the presence of Industry 4.0 and rapid environmental changes such as the emergence of the COVID-19 pandemic (Kutnjak, 2021). Industry 4.0 is supported by digital technology and innovations known as SMACIT (Vial, 2019) such as social media, mobile apps, data analytics, cloud computing, and the Internet of things (IoT), to add potential value to the enterprise. Meanwhile, the emergence of pandemic impacts reducing social interaction, social distancing, lockdowns, and movement restrictions force enterprises to accelerate new value creation by optimizing digital technology and digitization.

The new value-creation process utilizing digital technology, innovation, and digitization is referred to as digital transformation (Aguiar et al., 2019). Digital transformation is a new effort in utilizing digital artifacts, symbols, and systems to enable business development such as improving customer experience, streamlining business processes, or creating new business models (Fitzgerald et al., 2014). Digital transformation covers almost all aspects and dimensions of the organization (Gebayew et al., 2018; Zaoui & Souissi, 2021). Service and its innovation as an organization's product is the one aspect which is transformed (Mihardjo et al., 2020a). Another source states that for a successful implementation of digital transformation, service is the core of all (Matzner et al., 2018). Service innovation, including digital services, is an important focus to consider in the transformation process. Digital services that experience a transformation process through digital technology/innovation become part of digital transformation (Matzner et al., 2018). Some refer to it as digital service transformation (Klötzner, 2016; Matzner et al., 2018).

Until this research was conducted, not so many studies have specifically discussed digital service transformation terms. Some publications mentioning digital service transformation equate this concept with digital transformation (Mihardjo et al., 2020a, 2020b) or service transformation (Omar et al., 2017; Weerakkody et al., 2016) or digital servitization (Klötzner, 2016). The digital service transformation concept still varies according to the understanding of each researcher. This difference is caused, as an example, in manufacturing enterprises; relevant research on digital service transformation is still insufficient, particularly in terms of empirical testing and transformation mechanisms (Shen et al., 2021). Another study stated there is a gap in digital service ecosystem transformation theoretical knowledge. The conceptual boundaries are still ambiguous, and research in this area lacks impact (Tana et al., 2019), which becomes a research potential in the context of digital service transformation (Augenstein, 2017).

1.2. Scope and Purpose

This research aims to formulate a conceptual model of digital service transformation to provide a more comprehensive and systematic understanding in answering problem statements. It will also discuss identified components for the digital service transformation model. The contribution of this research is to enrich research related to digital transformation as well as to add insight regarding research on digital service transformation for further research.

This publication is structured as follows. Section 2 will discuss research related to digital service transformation, including digital transformation, service transformation, and digital service. Section 3 discusses methodology and data collection. Section 4 discusses the results and discussions related to digital service transformation. Section 5 provides conclusions from this research and shows the limitations and plans for future research.

2. RELATED WORKS

2.1. Digital Transformation

The presence of digital technology triggers digital disruption and forces digital transformation in organizations to adapt to these conditions, thus creating new organizational value (Vial, 2019). The results change organizational structure, processes, and business models. Digital transformation brings benefits such as increasing customer satisfaction, improving customer experience, increasing productivity, reducing costs, and gaining competitive advantages (Gebayew et al., 2018).

According to Verina and Titko (2019), three main components need to be aligned for the successful implementation of digital transformation, such as technology, human resources, and business/management processes. Digital transformation is initiated by driver components including digital technology, executive support, and increasing competition, as well as customer demands. The results include creating business value, increasing effi-

 https://doi.org/10.1633/JISTaP.2023.11.1.3

ciency and reducing costs, providing competitive advantages, and other benefits. Fig. 1 shows an image of a digital transformation process model in an organization.

The digital transformation process occurs primarily in three areas: customer experience, operational processes, and business models (Westerman et al., 2011). Each area contains nine building elements in total. According to Vial (2019), the impact of digital transformation on value creation and organizational structure can be positive or negative. These components were also identified as critical success factors for successful digital transformation implementation (Morakanyane et al., 2020).

Digital transformation affects business process, including products, offerings, and models (Bosch, 2018). Businesses are always aware of their customers' needs. They have shifted from traditional methods (through surveys and interviews) to digital ones (personalization platforms and big data) due to the emergence of digital technology. The purchased item was converted into a service product, followed by physical offerings evolving into digital media. Manufacturing companies have moved from producing only products to producing products including services. Meanwhile, the service industry transforms physical services into digital services. This process of transforming products and services can be considered as digital service transformation.

2.2. Service Transformation

Service transformation is an idea changing traditional products and services by adopting manufacturing concepts to create new value, such as profits (Matzner et al., 2018). Service transformation has the same meaning as servitization, service infusion, or service transition (Dombrowski & Fochler, 2018). Servitization is a paradigm shift in which organizations initially only selling products switch into selling products with integrated services on these products (Klötzner, 2016). Service transformation is an attempt to generate new markets and reshape existing ones, which they call service innovation with a focus on creating value for the organization (Kandampully et al., 2021).

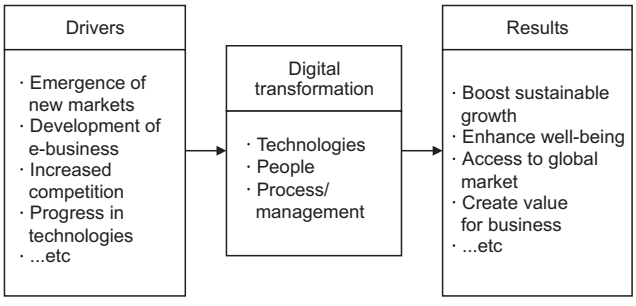


Fig. 1. Conceptual model for digital transformation (Verina & Titko, 2019).

Service transformation has a close relationship with digital transformation. Since 2011 and 2013, service transformation and digital transformation have been the subject of continuous and pertinent research (Dombrowski & Fochler, 2018). It is essential to recognize that digital transformation technology is a key driver of service transformation. Service transformation driven by digital technology and all types of smart machines are just called digital transformation (Matzner et al., 2018). Despite coming from different fields, service transformation and digital transformation can converge as the concept of digital servitization, the transition to intelligent service-product software systems that generate value via monitoring, control, optimization, and independent management (Shen et al., 2021).

2.3. Digital Service

Generally, services are intangible products or solutions produced by suppliers by utilizing existing resources to meet customer demands (Alter, 2008). From the standpoint of value creation, services are experienced and co-created by customers (Ben Letaifa et al., 2016). Digital service is service that utilizes digital technology as support. It is also known as advanced service technologies that rely on digital components embedded in physical products (Linde et al., 2023).

Digital services are objects of change in digital transformation. Organizations must use a combination of products, services, and software to derive value from digital services. These digital services must be personalized, dynamic, digital, and have high-quality solutions that emphasize perceptibility and connectivity, as a combination of physical products and services with digital value added (Shen et al., 2021). Digital services can be code-based software modules acquired and communicated via digital transactions, which are then sent to users via Internet protocols and supported by technological infrastructure (Ghazawneh, 2019).

3. METHODOLOGY

The discussion about digital service transformation literature review follows the systematic literature review guidelines in Kitchenham et al. (2009) and Henriette et al.

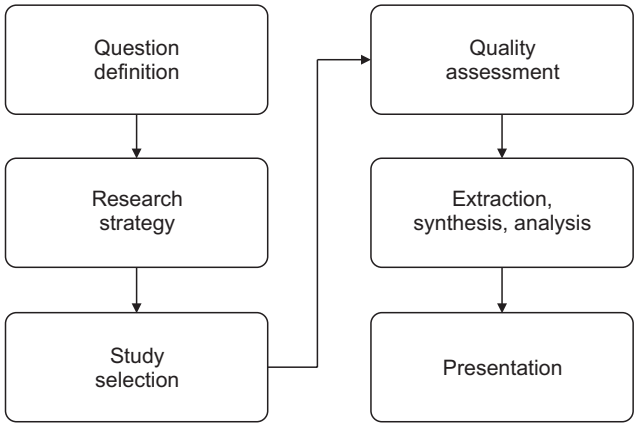


Fig. 2. Systematic literature review method (Henriette et al., 2015; Kitchenham et al., 2009).

(2015) as mentioned in Fig. 2.

3.1. Research Questions

We formulate the following research questions:

1. What research related to digital service transformation has been done?
2. What are the definitions contained in the research related?
3. What are the components of digital service transformation?
4. How can we develop a digital service transformation model based on the specified components?

3.2. Research Strategy

The main keyword used in this research is "Digital Service Transformation," using quotation marks (" ") to indicate that the search objective wants to get publications that mention the term *digital service transformation* both in the content, title, and metadata. To expand on the research, keywords are added to its building elements such as digital transformation, service transformation, and digital service. Keywords for searches include:

1. "Digital Service Transformation";
2. "Digital Service" AND ("Digital Transformation" OR "Digitalization" OR "Digitalisation");
3. "Digital Service" AND ("Service Transformation" OR "Servitization" OR "Servitisation").

The search of articles was conducted regardless of time limit by using some sources including ACM, IEEE, and Scopus (Henriette et al., 2015). The article search was conducted on September 14, 2021.

3.3. Study Selection

This section defines included/excluded criteria for found publications. The criteria for filtering include: written in English, scientific publication, focus on discussing digital service transformation, digital transformation, or service transformation on digital service. The scientific publication criteria accept papers, journals, conferences, proceedings, theses, dissertations, and research reports, and exclude white papers, reports, wire feeds, web pages, and blogs. Other criteria for rejected articles are informal literature surveys, duplicated articles from the same studies, and articles discussing the procedures used for evidence-based software engineering or systematic literature review. Articles in the form of short papers, abstracts, and incomplete articles will also be rejected (Gebayew et al., 2018).

3.4. Quality Assessment

To evaluate the rigor and credibility of the chosen articles, it is necessary to define criteria for quality assessment. Evaluation requires a thorough examination of existing articles. The criteria for quality include an adequate description of the research context, a clear statement of research objectives, an explanation of the research question, a concise statement of the findings, and others (Henriette et al., 2015). One focus of quality assessment is the suitability of the research design to research objectives. The authors select articles based on criteria that emphasize transformation in the provision of digital services. Articles containing these elements but lacking correlation will be rejected.

4. RESULT AND DISCUSSION

4.1. Research Related to Digital Service Transformation

Results from searching databases obtained 427 articles in total. After processing study selection and quality assessments, 52 qualified articles were obtained. Conducting descriptive analysis is one of the data presentation processes. Grouping includes publication year, bibliographic form, research method, and implementation area. Grouping is based on year in Table 1, bibliographic form in Table 2, research method in Table 3, and implementation area in Table 4.

Based on Table 1, the term *digital service transformation* was still not widely known until 2016, despite hav-

 https://doi.org/10.1633/JISTaP.2023.11.1.3

Table 1. Articles based on year of publication

Year	Count	References
2006	1	(Zysman, 2006)
2011	1	(Westerman et al., 2011)
2014	1	(Fichman et al., 2014)
2015	1	(Yeager, 2015)
2016	4	(Berghaus & Back, 2016; Klötzner, 2016; Truong, 2016; Weerakkody et al., 2016)
2017	7	(Augenstein, 2017; Fleig & Augenstein, 2017; Frach et al., 2017; Ismail et al., 2017; Mceachern & Cholewa, 2017; Omar et al., 2017; Ramadani et al., 2017)
2018	9	(Bracken & Greenway, 2018; Ghazawneh, 2018; Granados, 2018; Henfridsson et al., 2018; Matzner et al., 2018; Osmundsen et al., 2018; Swan, 2018; West et al., 2018; Zaoui & Souissi, 2018)
2019	9	(Ghazawneh, 2019; Jordan, 2019; McParland & Connolly, 2019; Sklyar et al., 2019a, 2019b; Tana et al., 2019; Verina & Titko, 2019; Vial, 2019; Zaki, 2019)
2020	10	(Bannister & Connolly, 2020; Bian, 2020; Germak et al., 2020; Gomes et al., 2020; Liu & Deng, 2020; Mihardjo et al., 2020a, 2020b; Park et al., 2020; Rajab, 2020; Sila & Martini, 2020)
2021	8	(Barravecchia et al., 2021; Kandampully et al., 2021; Manser Payne et al., 2021; Newsom & Jones, 2021; Rackwitz et al., 2021; Scupola & Mergel, 2021; Senyo et al., 2021; Shen et al., 2021)
2023	1	(Frennert, 2023)

Table 2. Articles based on bibliographic forms

Type	Count	References
Journal	23	(Augenstein, 2017; Bannister & Connolly, 2020; Barravecchia et al., 2021; Chatfield & Reddick, 2019; Fichman et al., 2014; Frennert, 2020; Gomes et al., 2020; Henfridsson et al., 2018; Kandampully et al., 2021; Manser Payne et al., 2021; Matzner et al., 2018; McParland & Connolly, 2019; Omar et al., 2017; Rackwitz et al., 2021; Scupola & Mergel, 2021; Senyo et al., 2021; Shen et al., 2021; Sila & Martini, 2020; Sklyar et al., 2019b; Vial, 2019; Weerakkody et al., 2016; Zaki, 2019; Zaoui & Souissi, 2018)
Conference	14	(Berghaus & Back, 2016; Germak et al., 2020; Ghazawneh, 2018, 2019; Granados, 2018; Mceachern & Cholewa, 2017; McParland & Connolly, 2019; Mihardjo et al., 2020a, 2020b; Osmundsen et al., 2018; Ramadani et al., 2017; Verina & Titko, 2019; West et al., 2018; Yeager, 2015)
Working paper	2	(Tana et al., 2019; Zysman, 2006)
Thesis	6	(Bian, 2020; Jordan, 2019; Klötzner, 2016; Rajab, 2020; Swan, 2018; Truong, 2016)
Book articles	7	(Augenstein, 2016; Bracken & Greenway, 2018; Frach et al., 2017; Ismail et al., 2018; Jones, 2021; Liu & Deng, 2020; Westerman et al., 2011)

ing been coined in 2006. The number of publications has increased over the past few years and is becoming more commonly known as of 2021. According to Table 2, publications that use the digital service transformation term the most are journals, followed by conferences and book articles. Grouping by research method according to Palvia et al. (2006) in Table 3 shows that proposing conceptual frameworks or models is most frequently discussed. Other sources discuss how to assess the implementation of digital service transformation using surveys and interviews. Others examine digital service transformation using literature analysis. Others monitor implementation using case studies and other qualitative methods. Table 4 shows digital service transformation has been implemented in numerous service areas, with the private sector taking the lead. There are also those studies that cannot be classified in any sector. Specific discussions in the private sector include finance, industry, manufacturing, health, and trans-

Table 3. Articles based on research method

Method	References
Framework and conceptual model	(Augenstein, 2016, 2017; Bannister & Connolly, 2020; Bian, 2020; Bracken & Greenway, 2018; Chatfield & Reddick, 2019; Granados, 2018; Jones, 2021; Jordan, 2019; Kandampully et al., 2021; Manser Payne et al., 2021; Matzner et al., 2018; Mceachern & Cholewa, 2017; Mihardjo et al., 2020a, 2020b; Rackwitz et al., 2021; Rajab, 2020; Ramadani et al., 2017; Verina & Titko, 2019; Vial, 2019; West et al., 2018; Yeager, 2015; Zaki, 2019)
Literature analysis	(Barravecchia et al., 2021; Chatfield & Reddick, 2019; Ghazawneh, 2018; Gomes et al., 2020; McParland & Connolly, 2019; Mihardjo et al., 2020a; Sklyar et al., 2019b; Zaoui & Souissi, 2018)
Case study	(Frach et al., 2017; Ghazawneh, 2019; Omar et al., 2017; Scupola & Mergel, 2021; Sklyar et al., 2019a; Truong, 2016; Weerakkody et al., 2016)
Survey	(Frennert, 2020; Germak et al., 2020; Klötzner, 2016; Manser Payne et al., 2021; Mihardjo et al., 2020b; Shen et al., 2021)
Mathematics analysis	(Liu & Deng, 2020; Mihardjo et al., 2020a; Zysman, 2006)
Qualitative	(Jordan, 2019; Osmundsen et al., 2018; Senyo et al., 2021; Sila & Martini, 2020; Swan, 2018; Tana et al., 2019)
Interview	(Frennert, 2020; Ismail et al., 2018; Klötzner, 2016; Scupola & Mergel, 2021; Sklyar et al., 2019a; Westerman et al., 2011)

Table 4. Articles based on implementation area

Sector	Area	Count	References
Private	Financial	4	(Bian, 2020; Granados, 2018; Tana et al., 2019; Truong, 2016)
	Industry	7	(Ghazawneh, 2018; Mihardjo et al., 2020a, 2020b; Sklyar et al., 2019a; West et al., 2018; Yeager, 2015; Zaki, 2019)
	Health	2	(Frennert, 2020; Mceachern & Cholewa, 2017)
	Manufacture	2	(Klötzner, 2016; Shen et al., 2021)
	Transportation	1	(Rajab, 2020)
Public	Government	13	(Bracken & Greenway, 2018; Chatfield & Reddick, 2019; Frach et al., 2017; Germak et al., 2020; Jones, 2021; Jordan, 2019; Omar et al., 2017; Rackwitz et al., 2021; Ramadani et al., 2017; Scupola & Mergel, 2021; Senyo et al., 2021; Swan, 2018; Weerakkody et al., 2016)
Unclassified	Organization in general	23	(Augenstein, 2016, 2017; Bannister & Connolly, 2020; Barravecchia et al., 2021; Bian, 2020; Fichman et al., 2014; Ghazawneh, 2019; Gomes et al., 2020; Henfridsson et al., 2018; Ismail et al., 2018; Kandampully et al., 2021; Liu & Deng, 2020; Manser Payne et al., 2021; Matzner et al., 2018; McParland & Connolly, 2019; Osmundsen et al., 2018; Sila & Martini, 2020; Sklyar et al., 2019b; Verina & Titko, 2019; Vial, 2019; Westerman et al., 2011; Zaoui & Souissi, 2018; Zysman, 2006)

portation in which their products are services. Industries and manufacturers not only participate in providing products but also services supporting products.

Research on digital service transformation also includes value co-creation of the digital service transformation process (Manser Payne et al., 2021; Scupola & Mergel, 2021), the role of servitization, digitalization, and innovation performance in manufacturing companies (Shen et al., 2021), research on the service-product system as the combination of services and products as part of servitization (Barravecchia et al., 2021), digital platformization as a transformation strategy (Senyo et al., 2021), optical communication technologies for 5G (Liu & Deng, 2020), servitization in the context of IoT (West et al., 2018),

 https://doi.org/10.1633/JISTaP.2023.11.1.3

Table 5. Relevant studies about digital service transformation

Ref.	Research focus	DST definitions	Elements/Components
(Klötzner, 2016)	Service transformation and servitization on manufacturing industries	...is the process of converting manufactured products into commercial services that are provided via digital channels, such as the Internet. The general goal is to provide the customer with a subscribable digital solution that fulfills the same needs as the physical product.	Servitization, digital servitization, digital transformation, digital service, integrated digital service, add-on digital service, purely digital service, digital transformation maturity model
(Zaki, 2019)	Four trajectories of digital transformation to reshape digital service	…process to address the challenge: how to create new propositions for customers that will deliver a return on investment, how to change from a rules-based culture to support innovation and take risks, and how to solve the challenge while existing business is still running	Big data, customer value, customer service, Service encounters, business model, digital service, digital economy, machine learning, AI, digital technologies, digital strategy, customer experience, data-driven business models
(Zysman, 2006)	Algorithmic for service transformation	The services that ride on the product platform become the differentiated asset that creates value for the firm	Digital technology, financial engineering, products and services, algorithmic service transformation
(Matzner et al., 2018)	Role of digital transformation in service management	Service transformation that is being enabled by digital and mostly smart machines	Customer contact, simultaneity, demand fluctuation, customization, service orientation, network ubiquity, blockchain, distributed ledger technology
(Ghazawneh, 2019)	Five main categories of digital services to develop created by the service provider by digital transformation	Although digital technologies are significant for digital service transformation, the processes, knowledge, and experience of developing these digital services are equally important which is facilitated by the adoption of digital platforms and ecosystems	Digital service; service transformation; platforms; service ecosystems; services experience; service process; service capabilities; service environment; service delivery; service provider: business, interaction, technological; service experience: website service experience, app service exp., soc-med exp., Internet bot exp.; service process: standardized, semi-standardized, non-standardized; service capabilities: technology capability, platform capability, hardware-device capability; service environment: user experience, service integration, service customization; service delivery: service delivery cost, service delivery review, service delivery infrastructure
(Rajab, 2020)	To explore how values and emotional engagement impact well-being in Self-Service-Technology	…the increased efficiency of service technology and its cost-effectiveness, service providers are rushing and competing toward digital transformation	Digital well-being, service quality, service design, emotional engagement, personalization, aviation, distance learning, self-service technology, service quality model
(Ghazawneh, 2018)	Role of digital transformation in the development of digital services	Digital transformation of services involves the digitalization of services from analog to digital and the change of the actual process generated by digitization	Digital services, service transformation, digital platforms, service ecosystems, service development

Table 5. Continued

Ref.	Research focus	DST definitions	Elements/Components
(Liu & Deng, 2020)	Overview of the emerging optical communication technologies	Digital service transformation which will minimize operating costs, deliver efficiencies, and drive revenue growth	5G oriented optical network, optical communication technology, optical transport network
(Barravecchia et al., 2021)	To explore research topics and future trends for product-service systems research	Digitalization supports companies in the development of new services, business models, and innovative products	Product-service systems, Text mining, latent Dirichlet allocation, topic landscape, servitization, technical services, extended service blueprint, digital servitization
(Chatfield & Reddick, 2019)	Process of building framework for IoT-enabled smart government	Developing and exploiting IoT-enabled dynamic capabilities to drive, further, and accelerate the process of digital transformation of public administration and governance	Digitalization, servitization, service ecosystem, digitization, centralization, integration
(Kandampully et al., 2021)	To explore paths of service transformation and identify situations and contexts of it	Digitalization is not only an antecedent and an outcome of service transformation. It is also one of its main partners in the effort to maintain viability, achieve growth, and realize superior service performance over the long term in a quickly changing and challenging business environment.	Service transformation, reformation, renovation, reincarnation, digitalization
(Scupola & Mergel, 2021)	To explore value co-creation in digital service transformation	Digital transformation is a holistic effort to revise core processes and services of government beyond the traditional digitization efforts. It evolves along a continuum of transition from analog to digital to a full-stack review of policies, current processes, and user needs and results in a complete revision of the existing and the creation of new digital services.	Digital transformation, public value, co-production, public service delivery, public administration, digital strategy, digital policy, digital service, service policies, digital transformation of public administration, economic value, administrative value, citizen value, democratic/societal value
(Augenstein, 2017)	Design intelligent business model transformation tool	…since the age of digitalization, services have become more and more important in companies' strategies. Additional pressure is created through the fact that consumers are more than ever able to compare these services on the markets.	Business model, business model canvas, design science, development tools, enterprise resource planning, business model mining
(Shen et al., 2021)	To explore the role of servitization and digitalization in manufacturing	The transition to intelligent product-service software systems that generate and capture value through monitoring, control, optimization, and self-management. According to this view, enterprises must use a combination of products, services, and software to derive value from digital services.	Servitization, digitalization, dynamic capabilities, enterprise innovation performance, digital servitization transformation, enterprise innovation, PSM-DID
(Sklyar et al., 2019b)	How to organize digital servitization from a service ecosystem perspective	The utilization of digital tools for the transformational processes whereby a company shifts from a product-centric to a service-centric business model and logic	Digitalization, servitization, service ecosystem, digitization, embeddedness, centralization, integration, digital servitization, digital service

 https://doi.org/10.1633/JISTaP.2023.11.1.3

Table 5. Continued

Ref.	Research focus	DST definitions	Elements/Components
(Senyo et al., 2021)	How does digital platformisation facilitate public sector transformation	Public sector transformation focused on ICTs to improve internal and external processes as well as structures to promote efficiency, transparency, and accountability	Digital platforms, digital platformisation, technology affordance, transformational affordance framework, developing economies, public sector transformation, paperless port, new public management
(Omar et al., 2017)	Enabling digital service transformation in UK public sector	Digitally-enabled service transformations in a government attempt to improve service delivery in public organizations	Digitally enabled transformation, public sector case, institutional theory, structuration theory
(Weerakkody et al., 2016)	Digitally-enabled service transformation in the public sector	As a policy instrument, digital initiatives continue to flourish in expectancy of enhancing the public service delivery system, lowering operational cost, and bettering government controls on information and data as well as transparency and efficiency	Public sector, service transformation, institutional theory, digital government, organizational change, digital-led service transformation, isomorphism process
(Truong, 2016)	To explore innovative edges of the financial technology	The disrupted financial markets have rapidly been transformed by the financial technology companies and transformed digital services	Blockchain, digital, finance, Fintech, investment, online, payment service, peer-to-peer, third-party integration
(Frach et al., 2017)	How to assess and compare digitalization in the public sector	A holistic approach to digital transformation emphasizes digital service transformation but includes workforce and internal process transformation measures	Digitalization, e-government, digital government, public service digitalization index, digital service delivery, digital enterprise, digital strategy, shared service
(Vial, 2019)	Understanding the framework and components of digital transformation	Transforming the organization to become a platform provider enabling customers to act as complements of the firm's digital services	Strategic responses, value propositions, value networks, digital channels, agility and ambidexterity, security/privacy, operational efficiency, organizational performance

framework for IoT (Chatfield & Reddick, 2019), designing an intelligent business model transformation tool (Augenstein, 2017), and others.

4.2. Defining Digital Service Transformation

The majority of collected articles do not explain the definition of the digital service transformation. Only an article with explicit definitions from Klötzner (2016) stated that digital service transformation is the process of transforming manufactured products into commercial services delivered through digital channels such as the Internet. This has the same meaning as servitization, as mentioned in Shen et al. (2021) and Sklyar et al. (2019b).

According to some researchers, digital service transformation is synonymous with digital transformation or service transformation. A study by Mihardjo et al. (2020a) implicitly describes digital service transformation as an organizational transformation based on digital service. Studies have also defined digital service innovation (Mihardjo et al., 2020b), digitally transformed public service (Germak et al., 2020), transformation on service ecosystem (Tana et al., 2019), transformation on enterprise services (Augenstein, 2017), or transformation in public services (Senyo et al., 2021). The same definition is also conveyed by Frennert (2023), Liu and Deng (2020), and Scupola and Mergel (2021). Others defined this term as service transformation enabled by digital technology and smart machines or digitally enabled service transformation (Omar et al., 2017; Weerakkody et al., 2016). Other studies consider it as a digitally implemented service process (Zysman, 2006), service transformation in the enterprise (Kandampully et al., 2021), service transformation in the industry (Ghazawneh, 2019; Zaki, 2019), or in government (Ramadani et al., 2017). Although these studies have considered similar terminology referring to service transformation, the definition does not mention the use of dig-

ital elements (Dombrowski & Fochler, 2018; Kandampully et al., 2021; Klötzner, 2016; Matzner et al., 2018); however, digital services require technology, platforms, and digital ecosystems which act as digital elements. Therefore, despite being classified as a service transformation, digital service transformation has the same meaning as digital transformation.

The variation of digital service transformation terminology based on each researcher's understanding is displayed in Table 5. This table outlines the research focus, proposed concept of digital service transformation, and terminology-related components. Based on definition analysis from Table 5, the authors attempt to formulate digital service transformation definition as: *a process of creating new/ improvement of digital service products as part of a business transformation and organizational strategy that aims to generate new added value through the use of digital technology.* The authors limit the scope of this definition to digital services because digital transformation, in general, can encompass many aspects of an organization. This definition will impact component determination and the proposed model.

4.3. Collecting Components of Digital Service Transformation

Based on its definition, the components of digital service transformation can be a combination of digital transformation and digital service components. According to Verina and Titko (2019), the conceptual model of digital transformation consists of three component categories: enablers, processes, and results. Similarly, Westerman et al. (2011) stated that digital transformation is enabled by digital capabilities, which initiate the process' key components (customer experience, operational processes, and

Table 6. Digital transformation components

Components	Sub-components	References
Drivers	Digital capabilities (digitization of tasks and processes, unified data and processes, analytics capability, business and IT integration, solution delivery)	(Verina & Titko, 2019; Vial, 2019; Westerman et al., 2011)
	Customer behavior and expectations	(Verina & Titko, 2019; Vial, 2019)
	Competitive landscape (new market, new business opportunities, increased competition)	(Verina & Titko, 2019; Vial, 2019)
	Executive support	(Verina & Titko, 2019; Vial, 2019)
	Law and regulation	(Verina & Titko, 2019; Vial, 2019)
Processes	Business model (digitally modified business, new digital business, digital globalization)	(Verina & Titko, 2019; Westerman et al., 2011)
	Operational process (process digitization, worker enablement, performance management)	(Verina & Titko, 2019; Westerman et al., 2011)
	The customer experience (customer understanding, top-line growth, customer touchpoints)	(Verina & Titko, 2019; Westerman et al., 2011)
	Use of digital technologies (social, mobile, analytics, cloud, IoT, platform, ecosystem)	(Verina & Titko, 2019; Vial, 2019)
Results	Value creation (value proposition, value networks, digital channels, agility, ambidexterity)	(Verina & Titko, 2019; Vial, 2019)
	Structural changes (organizational structure, organizational culture, leadership, employee roles, skills)	(Verina & Titko, 2019; Vial, 2019)
	Security and privacy	(Verina & Titko, 2019; Vial, 2019)
	Operational efficiency	(Verina & Titko, 2019; Vial, 2019)
	Organizational performance	(Verina & Titko, 2019; Vial, 2019)
	Industry and society improvements	(Verina & Titko, 2019; Vial, 2019)

 https://doi.org/10.1633/JISTaP.2023.11.1.3

Table 7. Digital service components

Components	Sub-components	References
Service function	Service as acts for the benefit of others	(Matzner et al., 2018)
	Service as outcomes	(Matzner et al., 2018)
	Service as exchange or value co-creation	(Matzner et al., 2018)
Service form	Integrated with product	(Klötzner, 2016)
	Add-on of product	(Klötzner, 2016)
	Purely service	(Klötzner, 2016)
Service development	Planning	(Scupola & Mergel, 2021)
	Design	(Scupola & Mergel, 2021)
	Management	(Scupola & Mergel, 2021)
	Delivery	(Scupola & Mergel, 2021)
	Assessment	(Scupola & Mergel, 2021)
Service value	Public value (economic, administrative, democratic, citizen)	(Jordan, 2019; Sklyar et al., 2019a; Zaki, 2019)
	Organizational value	(Jordan, 2019; Sklyar et al., 2019a; Zaki, 2019)
	Customer value	(Jordan, 2019; Sklyar et al., 2019a; Zaki, 2019)
	Shareholder value (market capitalization)	(Jordan, 2019; Sklyar et al., 2019a; Zaki, 2019)
Service objectives	Business	(Ghazawneh, 2018; Westerman et al., 2011)
	Interaction	(Ghazawneh, 2018; Westerman et al., 2011)
	Technological	(Ghazawneh, 2018; Westerman et al., 2011)
	Process	(Westerman et al., 2011)
Service experience	Website service experience	(Ghazawneh, 2018, 2019)
	App service experience	(Ghazawneh, 2018, 2019)
	Social media service experience	(Ghazawneh, 2018, 2019)
	Internet bot service experience	(Ghazawneh, 2018, 2019)
Service process	Standardized process	(Ghazawneh, 2018, 2019)
	Semi-standardized process	(Ghazawneh, 2018, 2019)
	Non-standardized process	(Ghazawneh, 2018, 2019)
Service capabilities	Technology capabilities	(Ghazawneh, 2018, 2019)
	Platform capabilities	(Ghazawneh, 2018, 2019)
	Hardware-device capabilities	(Ghazawneh, 2018, 2019)
Service environment	User experience	(Ghazawneh, 2018, 2019)
	Service integration	(Ghazawneh, 2018, 2019)
	Service customization	(Ghazawneh, 2018, 2019)
Service delivery	Cost delivery	(Ghazawneh, 2018, 2019)
	Review delivery	(Ghazawneh, 2018, 2019)
	Infrastructure delivery	(Ghazawneh, 2018, 2019)

business models). Similar to what is stated by Vial (2019), digital transformation is enabled by the use of digital technologies. Table 6 describes digital transformation components.

Digital service also has several component categories based on the collected articles. The components consist of: function (Matzner et al., 2018), form (Klötzner, 2016), development (Scupola & Mergel, 2021), value (Jordan, 2019; Sklyar et al., 2019b; Zaki, 2019), and objectives and developing processes (Ghazawneh, 2018). The components of digital services are outlined in Table 7.

4.4. Proposing Conceptual Model of Digital Service Transformation

The digital service transformation model is proposed to explain the position of digital service transformation and what or how its role is in the overall digital transformation. This model is proposed in Fig. 3. There are three primary building blocks in the model: driver/enabler, process/event, and result/impact, adopting from Vial (2019) and Verina and Titko (2019). These three primary building blocks become a foundation for this proposed model, as we consider the transformation of digital service products to be part of digital transformation as a whole. Drivers are a collection of components as enablers and determine the direction of transformation. A process is a group of components that are being changed/processed. Results are a group of impact components expected from a transformation.

Digital service transformation is triggered or enabled by a block which is named a 'driver.' The digital capabilities component is highlighted because service objectives and functions (both being digital service components) are considered as transformation driver/enablers. Due to this service capability, a digital service can transform a process block. There, digital service component groups are categorized according to each digital transformation domain that participates in/experiences transformation (a business model, customer experience, use of digital technologies, and operational processes). The group of digital service components consists of a service process, service environment, service development, service capabilities, service delivery, and service experience. Here we can see which digital service components have changed following the digital transformation domain in which they are located. The transformation process then produces the result/impact explained in the result block. In this block, the impact can be value increasing of the digital services provided (in the service value component) or a change of digital services form (in the service form component). These two components are part of value creation domain, which is one of digital transformation's components.

This model is built based on the stages presented in Baslyman et al. (2020), which include component extraction, model design, and model validation. In this study, we did not validate our proposed model. This model adopts from Vial (2019) and Verina and Titko (2019) as the basis of its design. The component extraction process has been done in the previous section. From component extraction, component metrics are compiled to show reference indicators and strengthen concepts from the proposed model. Next, we organize the digital service component collec-

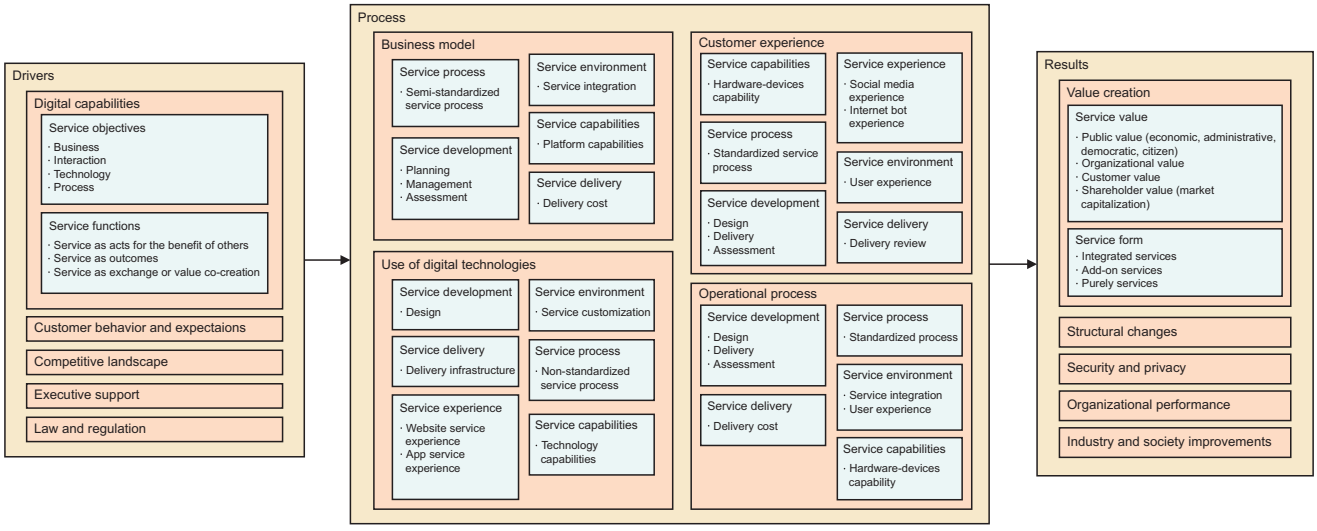


Fig. 3. Digital service transformation conceptual model.

 https://doi.org/10.1633/JISTaP.2023.11.1.3

tions into digital transformation blocks.

As previously mentioned, the three main building blocks of this proposed model are drivers, processes, and results. Each block contains several digital transformation domains, including five in the driver block, four in the process, and five in the results. There are digital capabilities in the driver block that include digital service components (service objectives and functions). Service objectives contain digital service goals, namely business, interaction, technology goals, and process. Business objectives are related to financial bottom line objectives, customer loyalty, brand formation, marketing, management, or strategies. Interaction is concerned with experienced users and process interaction design. Technology is related to technology and its components. Processes are related to improving routine services. Service functions relate to functions that can be provided by digital services, which include acts for the benefit of others as outcomes, and an exchange or value co-creation.

Furthermore, in the process block there are four domains where digital transformation occurs. The following is a description of digital service components:

1) The business model has five service components:
   a. Service capabilities relate to platform capabilities. The determination of the platform is carried out as part of the transformation process. Various types of digital platforms have their advantages, features, and business goals. Here, the service provider must ensure the appropriate platform for transformation needs.
   b. Service development. The digital service transformation process in this component includes planning, management, and assessment. Planning includes reorganizing the strategy and digital strategy work team, establishing community/governance groups, regulations, and teamwork. Management includes collaborating with externals or managing employee contracts. Transformation assessment is necessary through business forums, regulations, executive boards, and other comparable assessment forms.
   c. Service environment relates to how a service can be integrated or adapted to other services. From a business perspective, this transformation process is essential to increase the user base and provide alternative options to existing business models.
   d. Service processes relate to the management of semi-standard services in order to transform processes adhering to similar standards in order to create a new business opportunity.
   e. Service delivery in digital service transformation includes determining changes in service costs to end-users.

2) Customer experience has six service components:
   a. Service capabilities relate to device and hardware capabilities, in this case referring to device fragmentation across platforms where digital services are transformed. The greater the number of connected devices, the greater the customer-business interaction.
   b. Service experience. The transformation is accomplished by enhancing the user experience when interacting with social media and Internet bots. The social media experience identifies the integration of digital services with various social media tools, technologies, and networks. The variety of integrated social media channels and the level of interaction between these channels will impact the use of digital services. Experience with Internet bots is crucial due to the importance of feedback management and follow-up.
   c. The service environment in customer experience includes transformation of the digital service users' experience. Mapping the user experience is the foundation for creating a digital strategy to analyze customer growth rates. The transformation process is executed on services with specific user experience groups based on the results of growth rate analysis and organizational leadership decisions.
   d. Service process. The transformation in this case pertains to the standardized service process as the only service capable of user interaction.
   e. Service delivery. Transformation here pertains to the review of delivery to determine the end user's ability to interact with the organization or other users. Organizations must comprehend how the digital service transformation is implemented in order to create the simplest interaction features for end-users to review or be reviewed.
   f. Service development includes design, assessment, and delivery. The design includes transformation in digital service review and user tests as part of digital service design. The assessment includes the transformation of service assessments through forums and survey websites. The delivery includes transformation in the creation and delivery of digi-

tal services to additional service users.

3) The use of digital technology has five service components:
   a. Service capabilities involve technology transformation to design and develop digital services. Organizations must be able to evaluate the robustness, performance, adaptability, and efficacy of the technology employed in the development of digital services.
   b. Service environment includes transformation of service personalization to meet user requirements. In this case, service providers must collaborate with the corresponding service industry and end-users in order to adapt the customization process according to their needs.
   c. Service experience involves transformation in the experience of utilizing website or application services. The website service experience shows that organizations need to transform based on the technology used by users according to their experience in using websites. Likewise, the same is done for app service users on mobile devices.
   d. Service processes for non-standard services. This service is not standardized yet, so organizations must leverage their digital technology to adapt to the needs and preferences of users.
   e. Service delivery entails identifying the technology infrastructure that enables the transformation of digital services for end-users. Digital service providers must collaborate across different infrastructure levels, including platform, ecosystem, and digital marketplace levels to ensure the transformation goes according to plan.

4) The operational process has five service components:
   a. Service development. The transformed components include design, delivery, and assessment. The design includes transformation in reviews, user tests, and prototype tests in digital service design. The assessment process includes the transformation of service assessment through improvement forums and website surveys for employees. Service delivery entails the transformation of the creation and delivery process to end-users into a routine business operation.
   b. Service processes. The transformation relates to standardized services as routine operations of the organization.
   c. Service delivery includes transformation for cost-efficiency of services to end-users. The provisioning costs and associated costs outlined in the digital strategy must be followed by routine operational processes.
   d. The service environment includes service integration and user experience. Transformation occurs during the integration of services to enhance operational digitization. The user experience transformation process includes employee participation in customizing digital services.
   e. Service capabilities are associated with the transformation of digital device capabilities, which result in more integrated services, strengthening employee interaction and communication.

Value creation is part of the service component included in the results block. Value creation includes service value and service form. Service value is a value expected to be achieved. These values can take the form of public values (economic, administrative, democratic, or community values) that government organizations are expected to achieve. Other values include organizational value, which is anticipated for the organization's growth and development; customer value, which is anticipated from customer satisfaction and the fulfillment of expectations; and shareholder value, which is an anticipated profit for private organizations. Service form is an expected result from transformation itself, which includes integrated services with goods products, add-on services as complementary functions for products, and pure service, which is unbound and independent from products.

5. CONCLUSION

This paper discusses the terminology of digital service transformation based on the results of a systematic literature study. This study was able to identify previous research related to digital service transformation. This study also defines the concept and terminology of digital service transformation and its components based on the results of literature analysis. Furthermore, the results also include a proposed digital service transformation model that explains the position of this terminology and its role in overall digital transformation.

Digital service transformation is the process of creating or improving digital service products as part of a transformation of business and organization strategy that aims

 https://doi.org/10.1633/JISTaP.2023.11.1.3

to create new added value through the use of digital technology. The digital service transformation components collected in this study are organized into reference metrics and might become a basis to develop the proposed model. Nonetheless, the model we propose has not yet been validated.

This study is expected to give a comprehensive picture of how digital services are transforming, which can serve as a reference for further related studies. This paper will become a foundation for further research, including validating the proposed model and creating a reference model and life cycle model of digital service transformation.

CONFLICTS OF INTEREST

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

REFERENCES

Aguiar, T., Bogea Gomes, S., da Cunha, P. R., & da Silva, M. M. (2019, October 28-31). Digital transformation capability maturity model framework. *2019 IEEE 23rd International Enterprise Distributed Object Computing Conference (EDOC)* (pp. 51-57). IEEE.

Alter, S. (2008). Service system fundamentals: Work system, value chain, and life cycle. *IBM Systems Journal,* 47(1), 71-85. https://doi.org/10.1147/sj.471.0071

Augenstein, D. (2017). Designing an intelligent business model transformation tool. In M. Eibl, & M. Gaedke (Eds.), *Informatik 2017* (pp. 2357-2363). Gesellschaft für Informatik.

Bannister, F., & Connolly, R. (2020). The future ain't what it used to be: Forecasting the impact of ICT on the public sphere. *Government Information Quarterly,* 37(1), 101410. https://doi.org/10.1016/j.giq.2019.101410

Barravecchia, F., Franceschini, F., Mastrogiacomo, L., & Zaki, M. (2021). Research on product-service systems: Topic landscape and future trends. *Journal of Manufacturing Technology Management,* 32(9), 208-238. https://doi.org/10.1108/JMTM-04-2020-0164

Baslyman, M., AlGhamdi, A., & AlMuhaysh, S. (2020, August 31). Toward achieving the core goals of digital business transformation: A preliminary study. *2020 IEEE Tenth International Model-Driven Requirements Engineering (MoDRE)* (pp. 68-72). IEEE.

Ben Letaifa, S., Edvardsson, B., & Tronvoll, B. (2016). The role of social platforms in transforming service ecosystems. *Journal of Business Research,* 69(5), 1933-1938. https://doi.org/10.1016/j.jbusres.2015.10.083

Berghaus, S., & Back, A. (2016, September 4-6). Stages in digital business transformation: Results of an empirical maturity study. *MCIS 2016 Proceedings* (pp. 1-17). Association for Information Systems.

Bian, Z. (2020). *Test automation process assessment*. (Master's thesis). Metropolia University, Espoo, Finland.

Bosch, J. (2018). Towards a digital business operating system. In E. Cabello, J. Cardoso, L. Maciaszek, & M. van Sinderen (Eds.), *ICSOFT 2017: Software technologies* (pp. 296-308). Springer.

Bracken, M., & Greenway, A. (2018). *How to achieve and sustain government digital transformation.* IDB. https://doi.org/10.18235/0001215

Chatfield, A. T., & Reddick, C. G. (2019). A framework for Internet of Things-enabled smart government: A case of IoT cybersecurity policies and use cases in U.S. federal government. *Government Information Quarterly,* 36(2), 346-357. https://doi.org/10.1016/j.giq.2018.09.007

Dombrowski, U., & Fochler, S. (2018, July 31-August 2). Servitization as a key driver for digital transformation of manufacturing companies' spare parts service. *2018 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)* (pp. 291-296). IEEE.

Fichman, R. G., Dos Santos, B. L., & Zheng, Z. (2014). Digital innovation as a fundamental and powerful concept in the information systems curriculum. *MIS Quarterly,* 38(2), 329-343. https://doi.org/10.25300/misq/2014/38.2.01

Fitzgerald, M., Kruschwitz, N., Bonnet, D., & Welch, M. (2014). Embracing digital technology: A new strategic imperative. *MIT Sloan Management Review,* 55(2), 1-12.

Fleig, C., & Augenstein, D. (2017). *Developing a business model transformation tool.* https://doi.org/10.5445/ir/1000076445

Frach, L., Fehrmann, T., & Pfannes, P. (2017). Measuring digital government: How to assess and compare digitalisation in public sector organisations. In S. Falk, A. Römmele, & M. Silverman (Eds.), *Digital government: Leveraging innovation to improve public sector performance and outcomes for citizens* (pp. 25-38). Springer.

Frennert, S. (2023). Moral distress and ethical decision-making of eldercare professionals involved in digital service transformation. *Disability and Rehabilitation: Assistive Technology,* 18(2), 156-165. https://doi.org/10.1080/17483107.2020.1839579

Gebayew, C., Hardini, I. R., Panjaitan, G. H. A., Kurniawan, N. B., & Suhardi. (2018, October 22-26). A systematic literature review on digital transformation. In Suhardi, A. Z. R. Langi, & A. A. Arman (Eds.), *2018 International Conference on Information Technology Systems and Innovation (ICITSI)* (pp. 260-265). IEEE.

Germak, O. V., Gugueva, O. A., & Kalacheva, N. A. (2020). Automation of the materials' provision and the data from the cartographic and geodetic fund. *IOP Conference Series: Materials Science and Engineering,* 913(4), 042028.

Ghazawneh, A. (2018, August 22-24). The dynamics of transformation in the development of digital services. In B. Andersson, B. Johansson, S. Carlsson, C. Barry, M. Lang, H. Linger, & C. Schneider (Eds.), *27th International Conference on Information Systems Development* (pp. 595-606). Lund University.

Ghazawneh, A. (2019). Digital services development and the dynamics of transformation by service providers. In B. Andersson, B. Johansson, C. Barry, M. Lang, H. Linger, & C. Schneider (Eds.), *Advances in information systems development* (pp. 97-112). Springer.

Gomes, S. B., Santoro, F. M., & Mira da Silva, M. (2020). An ontology for BPM in digital transformation and innovation. *International Journal of Information System Modeling and Design,* 11(2), 52-77. http://doi.org/10.4018/IJISMD.2020040103

Granados, O. (2018). Rethinking branch banking network. In A. Morales, C. Gershenson, D. Braha, A. Minai, & Y. Bar-Yam (Eds.), *Unifying Themes in Complex Systems IX* (pp. 220-227). Springer.

Henfridsson, O., Nandhakumar, J., Scarbrough, H., & Panourgias, N. (2018). Recombination in the open-ended value landscape of digital innovation. *Information and Organization,* 28(2), 89-100. https://doi.org/10.1016/j.infoandorg.2018.03.001

Henriette, E., Feki, M., & Boughzala, I. (2015, October 2-5). The shape of digital transformation: A systematic literature review. *MCIS 2015 Proceedings* (pp. 431-450). Association for Information Systems.

Ismail, M. H., Khater, M., & Zaki, M. (2017). *Digital business transformation and strategy: What do we know so far?* Cambridge Service Alliance.

Jordan, M. (2019). *Creating public value through service design: Facilitation of human-centred service transformation in the city-state of Hamburg.* (Master's thesis). Laurea University of Applied Sciences, Vantaa, Finland.

Kandampully, J., Bilgihan, A., Bujisic, M., Kaplan, A., Jarvis, C. B., & Shukla, Y. (2021). Service transformation: How can it be achieved? *Journal of Business Research,* 136, 219-228. https://doi.org/10.1016/j.jbusres.2021.07.033

Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology,* 51(1), 7-15. https://doi.org/10.1016/j.infsof.2008.09.009

Klötzner, H. (2016). *The digital service transformation in the manufacturing industry.* (Master's thesis). UPF Barcelona School of Management, Barcelona, Spain.

Kutnjak, A. (2021). Covid-19 accelerates digital transformation in industries: Challenges, issues, barriers and problems in transformation. *IEEE Access,* 9, 79373-79388. https://doi.org/10.1109/ACCESS.2021.3084801

Linde, L., Frishammar, J., & Parida, V. (2023). Revenue models for digital servitization: A value capture framework for designing, developing, and scaling digital services. *IEEE Transactions on Engineering Management,* 70(1), 82-97. https://doi.org/10.1109/TEM.2021.3053386

Liu, X., & Deng, N. (2020). Emerging optical communication technologies for 5G. In A. E. Willner (Ed.), *Optical fiber telecommunications VII* (pp. 751-783). Academic Press.

Manser Payne, E. H., Peltier, J., & Barger, V. A. (2021). Enhancing the value co-creation process: Artificial intelligence and mobile banking service platforms. *Journal of Research in Interactive Marketing,* 15(1), 68-85. https://doi.org/10.1108/JRIM-10-2020-0214

Matzner, M., Büttgen, M., Demirkan, H., Spohrer, J., Alter, S., Fritzsche, A., Ng, I. C. L., Jonas, J. M., Martinez, V., Möslein, K. M., & Neely, A. (2018). Digital transformation in service management. *Journal of Service Management Research,* 2(2), 3-21. https://doi.org/10.15358/2511-8676-2018-2-3

McEachern, A., & Cholewa, D. (2017). Digital health services and digital identity in Alberta. In F. Lau, J. Bartle-Clar, G. Bliss, E. Borycki, K. Courtney, & A. Kuo (Eds.), *Studies in health technology and informatics* (pp. 222-227). IOS Press.

McParland, C., & Connolly, R. (2019, September 12-14). Employee monitoring in the digital era: Managing the impact of innovation. *2019 ENTRENOVA Conference Proceedings* (pp. 548-557). Society for Advancing Innovation and Research in Economy.

Mihardjo, L. W. W., Sasmoko, Alamsjah, F., & Elidjen. (2020a). Moderating effects of green IS on the relationship between organizational agility, customer experience and digital service innovation to achieve sustainable performance. *IOP Conference Series: Earth and Environmental Science,* 426, 012118. https://doi.org/10.1088/1755-1315/426/1/012118

Mihardjo, L. W. W., Sasmoko, Alamsjah, F., & Elidjen. (2020b). Role of green information system in developing corporate reputation and co-creation-innovation to attain sustainable performance. *IOP Conference Series: Earth and Environmental Science,* 426, 012120. https://doi.org/10.1088/1755-1315/426/1/012120

Morakanyane, R., O'Reilly, P., & Mcavoy, J. (2020, January 7-10). Determining digital transformation success factors. *Proceedings of the 53rd Hawaii International Conference on*

System Sciences (pp. 4356-4365). HICSS.

Newsom, A., & Jones, J. (2021). *MyTN: Digitizing and transforming the citizen experience within Tennessee state government.* https://www.nascio.org/awards-library/awards/digitizing-and-transforming-the-citizen-experience-within-tennessee-state-government

Omar, A., Weerakkody, V., & Sivarajah, U. (2017). Digitally enabled service transformation in UK public sector: A case analysis of universal credit. *International Journal of Information Management,* 37(4), 350-356. https://doi.org/10.1016/j.ijinfomgt.2017.04.001

Osmundsen, K., Iden, J., & Bygstad, B. (2018, September 28-30). Digital transformation: Drivers, success factors, and implications. *MCIS 2018 Proceedings* (pp. 1-15). MCIS.

Palvia, P., Midha, V., & Pinjani, P. (2006). Research models in information systems. *Communications of the Association for Information Systems,* 17. https://doi.org/10.17705/1CAIS.01747

Park, Y. J., Sang, Y., Lee, H., & Jones-Jang, S. M. (2020). The ontology of digital asset after death: Policy complexities, suggestions and critique of digital platforms. *Digital Policy, Regulation and Governance,* 22(1), 1-14. https://doi.org/10.1108/DPRG-04-2019-0030

Rackwitz, M., Hustedt, T., Hammerschmid, G. (2021). Digital transformation: From hierarchy to network-based collaboration? The case of the German "Online Access Act". *DMS – Der Moderne Staat – Zeitschrift für Public Policy, Recht und Management,* 14(1-2021), 101-120. https://doi.org/10.3224/dms.v14i1.05

Rajab, V. (2020). *DesİQual: Destination in motion*. (Master's thesis). University of Lapland, Rovaniemi, Finland.

Ramadani, L., Kurnia, S., & Breidbach, C. (2017, December 12-14). Advancing ICT4D research through service-dominant logic. *ACIS 2017 Proceedings* (pp. 1-9). ACIS.

Scupola, A., & Mergel, I. (2021). Value co-creation and digital service transformation: The case of Denmark. *SSRN Electronic Journal.* https://doi.org/10.2139/ssrn.3896269

Senyo, P. K., Effah, J., & Osabutey, E. L. C. (2021). Digital platformisation as public sector transformation strategy: A case of Ghana's paperless port. *Technological Forecasting and Social Change,* 162, 120387. https://doi.org/10.1016/j.techfore.2020.120387

Shen, L., Sun, C., & Ali, M. (2021). Role of servitization, digitalization, and innovation performance in manufacturing enterprises. *Sustainability,* 13(17):9878. https://doi.org/10.3390/su13179878

Sila, I. K., & Martini, I. A. O. (2020). Transformation and revitalization of service quality in the digital era of revolutionary disruption 4.0. *Jurnal Ilmiah Manajemen Bisnis Dan Inovasi Universitas Sam Ratulang,* 7(2), 197-205. https://ejournal.unsrat.ac.id/index.php/jmbi/article/view/30257/29245

Sklyar, A., Kowalkowski, C., Sörhammar, D., & Tronvoll, B. (2019a). Resource integration through digitalisation: A service ecosystem perspective. *Journal of Marketing Management,* 35(11-12), 974-991. https://doi.org/10.1080/0267257X.2019.1600572

Sklyar, A., Kowalkowski, C., Tronvoll, B., & Sörhammar, D. (2019b). Organizing for digital servitization: A service ecosystem perspective. *Journal of Business Research,* 104, 450-460. https://doi.org/10.1016/j.jbusres.2019.02.012

Swan, K. (2018). *Leading with legitimacy in government design labs: Introducing design thinking to the Finnish immigration service*. (Master's thesis). Aalto University School of Business, Espoo, Finland.

Tana, S., Breidbach, C., & Turpin, A. (2019, June 8-14). I want a Lamborghini: An ethnography of cryptocurrency communities. *Proceedings of the 27th European Conference on Information Systems (ECIS)* (pp. 1-11). Association for Information Systems.

Truong, O. (2016). *How fintech industry is changing the world.* (Doctoral dissertation). Centria University of Applied Sciences, Kokkola, Finland.

Verina, N., & Titko, J. (2019, May 9-10). Digital transformation: Conceptual framework. In V. Skvarciany, & J. Stankevičienė (Eds.), *International Scientific Conference: Contemporary Issues in Business, Management and Economics Engineering' 2019* (pp. 719-727). Vilnius Gediminas Technical University.

Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *Journal of Strategic Information Systems,* 28(2), 118-144. https://doi.org/10.1016/j.jsis.2019.01.003

Weerakkody, V., Omar, A., El-Haddadeh, R., & Al-Busaidy, M. (2016). Digitally-enabled service transformation in the public sector: The lure of institutional pressure and strategic response towards change. *Government Information Quarterly,* 33(4), 658-668. https://doi.org/10.1016/j.giq.2016.06.006

West, S., Kujawski, D., & Bächler, L. (2018, November 22-23). Servitization of office lighting in the context of IoT. In E. Gomes, M. P. Cunha, & F. Vendrell-Herrero (Eds.), *Book of abstracts: 7th International Conference on Business Servitization* (pp. 162-169). OmniaScience.

Westerman, G., Calméjane, C., Bonnet, D., Ferraris, P., & McAfee, A. (2011). *Digital transformation: A roadmap for billion-dollar organizations.* https://www.capgemini.com/wp-content/uploads/2017/07/Digital_Transformation__A_Road-Map_for_Billion-Dollar_Organizations.pdf

Yeager, K. (2015, June 7-10). DC microgrid performance excel-

lence in electricity renewal. *2015 IEEE First International Conference on DC Microgrids (ICDCM)* (pp. 377-380). IEEE.

Zaki, M. (2019). Digital transformation: Harnessing digital technologies for the next generation of services. *Journal of Services Marketing,* 33(4), 429-435. https://doi.org/10.1108/JSM-01-2019-0034

Zaoui, F., & Souissi, N. (2018). Onto-digital: An ontology-based model for digital transformation's knowledge. *International Journal of Information Technology and Computer Science,* 10(12), 1-12. https://doi.org/10.5815/ijitcs.2018.12.01

Zaoui, F., & Souissi, N. (2021, June 5-12). A framework for a strategic digital transformation. In M. El Mohajir, M. Al Achhab, B. E. El Mohajir, B. K. Ane, & I. Jellouli (Eds.), *2020 6th IEEE Congress on Information Science and Technology (CiSt)* (pp. 502-508). IEEE.

Zysman, J. (2006). The algorithmic revolution? The fourth service transformation. *Communications of the ACM,* 49(7), 48. https://doi.org/10.1145/1139922.1139947

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

RESEARCH PAPER
J Inf Sci Theory Pract 11(1): 49-60, 2023

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2023.11.1.4

Social Media Celebrities in Bangladesh

Md. Sayeed Al-Zaman*
Department of Journalism and Media Studies, Jahangirnagar University,
Savar, Dhaka, Bangladesh
E-mail: msalzaman@juniv.edu

ABSTRACT

This study endeavors to answer four research questions: (a) Who are the celebrities? (b) How popular are they? (c) What do they post on social media platforms? (d) How do their followers react to them? Following quantitative methods, this study analyzes the top 195 Facebook celebrities of Bangladesh and their 9,441 Facebook posts' interactions. The result suggests that actors, music-related celebrities, and sports stars are the most popular celebrities. While male celebrities have a higher frequency, female celebrities have higher average followers than their male counterparts. Celebrities mostly share photos on Facebook, perhaps to recreate their public images and provide regular updates to their followers. The followers also engage the most in photos with affection and surprise reactions. Their reactions to celebrities' Facebook content are highly positive. Some strengths and limitations of this study are also discussed.

Keywords: celebrity, social media, online reactions, Facebook, Bangladesh

Received: October 10, 2022 **Revised:** November 20, 2022
Accepted: January 1, 2023 **Published:** March 30, 2023

***Corresponding Author:** Md. Sayeed Al-Zaman
https://orcid.org/0000-0003-1433-7387
E-mail: msalzaman@juniv.edu

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far 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.

© Md. Sayeed Al-Zaman, 2023

1. INTRODUCTION

This study endeavored to understand the celebrity culture in contemporary Bangladesh, centering on social media. Studies on online celebrity culture have been growing worldwide, though their focus is primarily on the Global North (Lau, 2021). Besides this, global academic scholarship has overlooked some regions and countries like Bangladesh, where celebrity culture is thriving. As with regional bias, celebrity studies also seem limited to a few specific disciplines, such as business studies and psychology (Abidin, 2018; Aljasir, 2019; Bennett, 2014; Chung & Cho, 2017; Hou, 2019; Jacobsen, 2020; Kim & Kim, 2020; Ledbetter & Redd, 2016). These two issues have created a knowledge vacuum that demands further scholarly attention and exploration. Since social media usage has dramatically increased worldwide in recent years, transforming the global celebrity culture, research on social media celebrities has become highly relevant and imperative. Previous research shows that social media celebrities can influence followers more than traditional celebrities can (Ahmadi et al., 2022). For this reason, many celebrities are also interested in using social media to reach an expanded audience (Selkie, 2022). However, social sciences and humanities scholars have paid limited attention to online celebrity culture. As a result, a functional and appropriate theoretical framework remains underdeveloped (Couldry, 2015). Taking these issues into account, this research studies online celebrity culture in Bangladesh.

Why is a study on Bangladesh necessary? In Bangladesh, 49.55 million people (29.7% of the population) use social media, and 4.6 million users (+10.1%) increased its ranks in 2021-22. The community of social media users is expanding rapidly, allowing celebrities to exert immense influence over their followers' behavior (Selkie, 2022). However, due to a lack of proper academic investigation, insights regarding online celebrities are limited, impeding the understanding of people's choices and the cultural trends of the country. Not only culture is influenced; celebrities often knowingly or unknowingly endorse products and information that may negatively affect people's health (Selkie, 2022). The parasocial relationship between celebrities and followers can also inspire followers toward positive actions. During the COVID-19 pandemic, for example, national cricket stars donated a large amount of money to the government fund to combat the crisis, which was well-received by their followers and later motivated more people to do the same (Sakib, 2020).

Based on these rationales and in the absence of relevant research, this study attempted to answer four research questions in the context of Bangladesh: Who are the celebrities, how popular are they, what do they post on social media platforms, and how do their followers react to them? Following a quantitative content analysis method and descriptive statistical analysis, this study found that the most popular celebrities on social media are actors, musicians, and sports stars. Also, celebrities mainly share photos on social media, perhaps to recreate their public image and provide regular updates to their followers. The following section discusses relevant literature to provide a background for this research.

2. BACKGROUND

2.1. Who Are the Celebrities?

In its simplest form, *celebrity* means being popular to many people, who are known as *followers* or *fans*. Theoretically, celebrity-*ism* is "understood to be a quality of individuals" (Rojek, 2015, p. 1). A celebrity is an individual who receives public attention. Are celebrities and influencers the same? While some scholars use the terms interchangeably, some argue that the two are different and that an influencer is *the* epitome of a celebrity (Abidin, 2018; Hou, 2019). However, this paper does not maintain a strict and subtle division between the concepts. Celebrities can be of three types: *ascribed celebrities*, who are popular due to their hereditary status; *achieved celebrities*, who become popular through their recognized talents; and *celetoids* (a portmanteau of *celebrity* and *tabloids*), who are short-lived celebrities, ascending to fame often through scandal, crime, or exceptional performances (Driessens, 2021; Rojek, 2015). In the age of social media, the prevalence of celetoids seems to have increased.

Who are and can be celebrities nowadays? Some studies have investigated the professions of celebrities who are popular in the public domain and their level of popularity (Schimmelpfennig, 2018). Most studies we found relevant were from business and economics disciplines. For example, Um (2013), following a quantitative content analysis method, investigated the dominant professions of celebrity endorsers in the United States (US) and Korea. They explored the following five professions of celebrities: actors, athletes, musicians, comedians, and others. In both countries, the result shows that actors (59.7% in Korea and 63.8% in the US) were the most popular celebrities for commercials. Also, female celebrities (46.8% in Korea and 61% in the US) were more popular than male celebrities (31.7% in Korea and 33.3% in the US). Um's study further

 https://doi.org/10.1633/JISTaP.2023.11.1.4

illustrated that celebrities' professions were more varied in the US than in Korea, since the US "modeled such professions as CEOs, chefs, supermodels, and TV personalities" (Um, 2013, p. 167). The study did not mention what professions the *others* category includes.

In a similar study, Schimmelpfennig (2018) explored popular celebrities' professions in German advertisements: actors, athletes, models, musicians, and broadcast stars. Of them, according to their appearance, actors (23%) were the most popular, followed by athletes (22%) and models (14%) (Schimmelpfennig, 2018, p. 7). Following a quantitative content analysis of female magazine advertisements, St. James (2010) studied celebrities' gender- and profession-based proportions. The study found as the five top following celebrity professions: musicians (42.4%), actors (40.1%), athletes (12.1%), TV personalities (3.5%), and models (2%). Of them, 94.2% were female, and 8.2% were male celebrities. Note that previous studies preferred to use the term *athletes* instead of *sports stars*.

In another study, Gergaud et al. (2012) analyzed the top 100 celebrities published by Forbes. Professions of celebrities in the study were grouped into four broad categories with their frequencies: actors/broadcast stars (n=25), musicians/singers (n=7), sports (n=7), and others (n=10). Of them, 35 were males, 14 were females, 42 were Caucasians, 3 were Hispanics, 2 were Afro-Americans, and 2 were Asians. Like Um (2013), this study also did not explain what the *others* category includes. Also, gender and ethnicity were conflated, further limiting the study. Beyond the business discipline, celebrities' professions are also a concern in health studies. Ahmad et al. (2020) investigated deceased celebrities from 1999 to 2017 in the US due to drug overdose. Their result shows the 10 top professions of celebrities: athletes, musicians, actors, entertainers, models, writers, DJs, directors, activists, and wrestling managers.

Overall, these studies cannot provide at least two relevant insights. First, they dealt with celebrities' statistics collected from traditional media instead of social media. Second, they did not analyze celebrities' level of popularity. Instead, they relied only on celebrities' frequencies. Therefore, this study asks the following research questions:

RQ1: Who are the most popular celebrities on Facebook?
RQ2: How popular are they?

2.2. Celebrities and Social Media

Media plays a pivotal role in creating celebrities because recurrent media representation of individuals shapes and reconstructs their public image (Driessens, 2021). In the past, traditional media was the sole agent of producing celebrities. However, social media now occupies an important space in this discussion (Abidin, 2018). It empowers users to self-represent themselves and their talents and to reach a large number of fans easily. Such self-branding and direct, instantaneous, and regular communication seem effective, which has changed how we perceive celebrity culture in the pre-digital era (Bennett, 2014).

It is important to remember that the pervasiveness of social media is now erasing the distinction between real-life celebrities and social media celebrities. Nowadays, while many celebrities emerge from social media, many others are real-life celebrities who may wish to maximize their fandom through social media platforms. Given the present context, more studies are required to identify the features of contemporary online celebrity culture.

Meanwhile, researchers, mainly from psychology and business studies disciplines, have produced numerous papers focusing on celebrity endorsement, the psychology of fandom, branding, and advertising intertwining social media (Abidin, 2018; Aljasir, 2019; Bennett, 2014; Chung & Cho, 2017; Hou, 2019; Jacobsen, 2020; Kim & Kim, 2020; Ledbetter & Redd, 2016). Also, previous studies dealing with celebrity culture relied on only a few specific theories. For example, online celebrity studies from the psychology discipline incorporated parasocial interaction theories (Bennett, 2014; Chung & Cho, 2017; Kim & Kim, 2020; Kim & Song, 2016; Ledbetter & Redd, 2016). Again, studies from business and marketing disciplines relied on celebrity endorsement theory (Aljasir, 2019; Arefin, 2018; Chung & Cho, 2017; Schimmelpfennig, 2018; Shabnaz et al., 2019; St. James, 2010; Um, 2013). However, pure social theory is still absent in celebrity studies (Couldry, 2015), which compels communication scholars to depend on theories from other fields (Hou, 2019; Kim & Kim, 2020; Ledbetter & Redd, 2016). Moreover, classical theories used in celebrity studies seem less competent and less applicable to celebrity-fan interactions on social media platforms (Aljasir, 2019). These conceptual limitations call for theoretical and methodological experimentations.

Another limitation existing in previous literature is that celebrity research is highly West-centric, leaving other regions and countries out of focus (Shabnaz et al., 2019). For example, two keyword searches (Search 1: "celebrity" OR "celebrities" AND "social media" OR "online" OR "Facebook" OR "Twitter" AND "Bangladesh"; Search 2:

"celebrity" OR "celebrities" AND "Bangladesh") in the Scopus database yielded no result, indicating no significant research on the celebrity culture of Bangladesh.

However, we found some underexposed studies conducted by some native scholars from the business discipline. For example, Shabnaz et al. (2019) surveyed 568 women from Dhaka, the capital of Bangladesh, to investigate how celebrity endorsements influence consumers' intention to choose personal care products. Like similar previous studies (Schimmelpfennig, 2018; Um, 2013), this Bangladesh-based study focuses on consumer behavior and product marketing. In addition, similar research explored Bangladeshi consumers' music preferences (Arefin, 2018). The growing online community, impactful social media celebrity culture, the prevalence of celetoids, and the absence of proper scholarly investigation provide rationales for the present study. Therefore, this exploratory study further asks two questions:

RQ3: What types of content do celebrities post on Facebook?
RQ4: How do the followers react to these posts?

2.3. Framing Social Media Reactions

Two dominant theories from psychology are used widely in scholarly literature to frame and understand human reactions: the circumplex model of affect, and basic emotion theory. The Circumplex Model of Affect by Russell (1980) includes two primary variables: valence and arousal. Valence is the inclination of emotion, which could be positive and negative, and arousal is the level of emotional intensity, which could be high or low. Combining these two variables, this model includes 28 different emotional reactions in four categories: high arousal positive valence, low arousal positive valence, high arousal negative valence, and low arousal negative valence (Russell, 1980).

In contrast, the basic emotion theory proposed by Ekman (1992) seems more simplified, integrating fewer broad reactions in the framework: happiness, sadness, disgust, fear, anger, and surprise. Like the circumplex model, this theory also relies on human facial expressions. Interestingly, theorists cannot agree on universal emotions to date, allowing researchers to produce different sets of *basic* emotions (Kowalska & Wróbel, 2017). In the study of social media reactions, however, the basic emotion theory seems more relevant than others. For example, in February 2016, Facebook, the world's most populated social media platform, introduced five major reactions (i.e., love, wow, haha, sad, and angry) influenced by the basic emotion theory (Al-Rawi, 2020; Dewey, 2016). Previous studies also used this theory as a theoretical lens to analyze social media users' reactions to different issues (Eberl et al., 2020; Freeman et al., 2019, 2020; Geboers et al., 2020; Giuntini et al., 2019; Jost et al., 2020; Tran et al., 2018; Varanasi et al., 2018).

3. MATERIALS AND METHODS

We prepared two datasets for two different analyses: The first answered RQ1 and RQ2, and the second answered RQ3 and RQ4. We focused on Facebook because most Bangladeshi social media users use this platform (86.73%) compared to the others, such as Twitter (7.73%) and YouTube (3.58%) (StatCounter, 2020). Moreover, of the total Facebook users, 41.4% are found to be highly addicted to it, making the platform influential and worth exploring (Raisa, 2018).

For dataset one, we collected the popularity data of Bangladeshi social media celebrities from SocialBakers (http://socialbakers.com), an AI-powered social media marketing platform. It has a rich collection of Facebook data, which is used widely in scholarly literature (Brosnan & Gavin, 2015; Mamur, 2015). From this website, we collected data on the top 200 most popular Bangladeshi Facebook celebrities based on their followers. The data included the names of the celebrities, who they are, their numbers of followers, and the monthly growth of their followers. However, information about some celebrities was missing from the website. In such cases, we collected it from their official Facebook fan page or by searching the Internet.

Although the website attributed nine professions to celebrities, we found it problematic from at least two aspects. First, some celebrities are known in more than one profession (e.g., Tahsan is a musician and actor). In these cases, we tried to determine their most popular and accepted identities to their fans, which often contradicted the attributed professions on the website. Second, some celebrities were incorrectly identified on the website, so we amended them (e.g., Sajeeb Wazed is a politician, not a businessman). Also, we found some information incorrect. For example, the list included Abdur Razzak, a non-Bangladeshi and Pakistani lawyer. Therefore, we eliminated such celebrity profiles (n=5) from our dataset. Our final dataset included 195 celebrity profiles and their details. This dataset was optimized to analyze the professions and genders of the most popular celebrities (RQ1)

 https://doi.org/10.1633/JISTaP.2023.11.1.4

and their popularity patterns (RQ2). For RQ1, we coded the professions based on previous studies and the website's attributed professions, along with a few newly attributed professions (Ahmad et al., 2020; Gergaud et al., 2012; Schimmelpfennig, 2018; St. James, 2010; Um, 2013).

For dataset two, we collected celebrities' Facebook posts and users' engagement in those posts. To understand what the celebrities post on Facebook (RQ3) and how users react to them (RQ4), we selected the five most-followed celebrity profiles: Shakib Al Hasan (sports star; 14.20 million followers), Mushfiqur Rahim (sports star; 12.01 million followers), Pori Moni (actress; 9.09 million followers), Tahsan (singer; 8.89 million followers), and Hanif Sanket (broadcast star; 8.77 million followers). We set the time range between 21 May 2011 and 20 May 2021 and extracted all posts' data from their official Facebook pages. In this data extraction process, we used CrowdTangle (http://crowdtangle.com), a public insights tool owned and operated by Facebook (CrowdTangle, 2021). Researchers now widely use it, and we have authorized access to harvest Facebook data. We collected 9,441 posts ($M_{Celebrity}$=1,882) and 475.44 million users' interactions ($M_{Celebrity}$=95.09): 50,359.26 reactions per post.

CrowdTangle has built-in filters/features for extracting Facebook data. It categorizes the content types as links, statuses, photos, videos, live videos complete, live videos scheduled, native videos, and YouTube videos. To avoid discrepancy and confusion, we merged the five video-related categories (i.e., videos, live videos complete, live videos scheduled, native videos, and YouTube videos) and named them as *videos*. We measured the interactions with comments, shares, like, love, wow, sad, haha, angry, and care attributes. Facebook's five main reaction buttons (i.e., love, haha, wow, sad, and angry) allow users to express their emotions (Dewey, 2016). In the light of previous literature, for the present study, we reformulated the meaning and valence of these reaction buttons as follows:

- *Love* (valence: positive; meaning: positivity, affection, empathy, support, kindness, liking, happiness, amused, proud, confident, satisfied, contented, and comforted).
- *Sadness* (valence: negative; meaning: empathy, grief, tragedy, support, misery, guilt, and depression).
- *Haha* (valence: positive; meaning: happy, laughter, fun, mockery, and rejection).
- *Wow* (valence: positive; meaning: muted disdain, surprise, shock, skepticism, interest, confused, and aroused).
- *Anger* (valence: negative; meaning: denial, aggressiveness, strong and overt disdain, dislike, disturbance, restlessness, and affliction) (Dewey, 2016).

We used descriptive statistical analysis to yield frequency counts (n), percentages (%), mean (M), and standard deviation (SD) of the relevant variables. It is a basic statistical analysis to extract descriptive insights from a dataset (Denis, 2018). Also, we performed inferential statistical analyses (i.e., correlation analysis) where necessary. The datasets were prepared using Microsoft Excel, and the statistical analyses were performed using IBM Statistical Package for the Social Sciences (SPSS) 25 (IBM Co., Armonk, NY, USA).

4. RESULTS AND DISCUSSION

This study aimed to explore four aspects of online celebrity culture: celebrities from those professions and genders that are the most popular, their popularity, what they post on social media, and how their fans react to them.

The result shows that 13 types of celebrities are more prevalent in Bangladesh (Table 1). Of them, actors (33.85%) have the highest frequency, followed by singers (24.10%) and musicians (14.36%). Despite their proximity, singers and musicians are categorized separately: singers are indicated as the persons who *only* sing using their vocals, and musicians write and compose music and often sing. We combined both categories: music-related celebrities constitute 38.46% of the total celebrities.

Sports stars (8.72%) also have a higher frequency than some other celebrities, such as writers (5.13%) and models (3.08%). Although sports stars are in the fourth position on the list, their followers (22.2%) are the second highest after actors (39.3%). More interestingly, sports stars have the highest mean value (M=3.93 M), followed by Islamic scholars (M=2.22 M), ranking ninth in the list with 2.05% frequency and 2.9% followers. In terms of average monthly followers' growth, actors have the highest percentage (48.5%), followed by singers (31.8%) and sports stars (8.4%). However, mean growth indicates that broadcast stars ($M_{Celebrity}$=0.18 M) have the highest average growth, followed by actors ($M_{Celebrity}$=0.15 M) and singers ($M_{Celebrity}$=0.13 M). Studies on Korean (59.7%), American (63.8%), and German (23%) celebrity endorsements of brands suggest that actors are the most popular celebrities (Schimmelpfennig, 2018; Um, 2013), which supports the findings of this study.

Another study (St. James, 2010) shows the seemingly

Table 1. Descriptions of the most popular celebrities

Celebrity	n	%	Followers				Growth			
			Followers[a)]	%	Mean[a)]	Standard deviation[a)]	Followers[a)]	%	Mean[a)]	Standard deviation[a)]
Actor	66	33.85	118.21	39.3	1.79	1.89	9.55	48.5	0.15	0.23
Singer	47	24.10	54.19	18.0	1.15	1.61	6.26	31.8	0.13	0.46
Musician	28	14.36	13.16	4.4	4.70	0.54	0.18	0.9	0.00	0.02
Sports star	17	8.72	66.76	22.2	3.93	4.27	1.65	8.4	0.10	0.10
Writer	10	5.13	14.40	4.8	1.44	2.17	0.46	2.3	0.05	0.08
Model	6	3.08	1.81	0.6	0.30	0.24	0.02	0.1	0.00	0.00
Broadcast star	4	2.05	10.72	3.6	2.68	4.10	0.71	3.6	0.18	0.20
Entrepreneur	4	2.05	0.69	0.2	0.17	0.11	-0.00	0.0	-0.00	0.00
Islamic scholar	4	2.05	8.86	2.9	2.22	1.50	0.34	1.7	0.08	0.05
DJ	3	1.54	3.61	1.2	1.20	1.13	0.12	0.6	0.04	0.04
Politician	3	1.54	4.52	1.5	1.51	1.34	0.03	0.1	0.00	0.01
YouTuber	2	1.03	2.96	1.0	1.48	0.68	0.20	1.0	0.10	0.10
Doctor	1	0.51	0.10	0.3	0.10	0	0.17	0.9	0.17	0.00

[a)]Values are in millions.

Table 2. Descriptions of genders of the most popular celebrities

Gender	n	%	Followers				Growth			
			Followers[a)]	%	Mean[a)]	Standard deviation[a)]	Followers[a)]	%	Mean[a)]	Standard deviation[a)]
Male	125	64.10	189.17	62.9	1.51	2.34	9.21	46.8	0.07	0.24
Female	70	35.90	111.70	37.1	1.60	1.87	10.48	53.2	0.15	0.31

[a)]Values are in millions.

contradictory result that musicians (42.4%) were more popular than actors (40.1%) for advertising purposes. However, unlike other studies and the present one, this research combined singers and musicians into a single category. On the other hand, Gergaud et al. (2012) combined actors and broadcast stars as the most popular celebrity category, making it difficult to determine the share of actors. These results infer that the popularity of actors is not limited to only advertisements and marketing. Instead, they are overall the most popular celebrities. Also, the popularity of actors seems not country-specific.

This study also found that male celebrities (64.10%) have higher frequencies than female celebrities (35.90%), which corresponds to their number of followers as well (Table 2). Previous studies provided limited insights regarding celebrities' gender. A relevant study suggests that male celebrities are more common than female celebrities in advertising (Um, 2013), which is similar to the gender-based frequency of the present study. However, the average number of followers of female celebrities (M_{gender}=1.60 M) is slightly higher than for male celebrities (M_{gender}=1.51 M), which also corresponds to their followers' growth. It indicates that though female celebrities are fewer in number, they have a higher number of fans on average than male celebrities. An Eta coefficient test between celebrity professions and followers shows a strong association, and

 https://doi.org/10.1633/JISTaP.2023.11.1.4

between professions and follower growth shows a weak association. Again, associations between gender and followers and gender and follower growth are weak or negligible (according to Quinnipiac University; Akoglu, 2018).

Female actors are higher in frequency (n=45) with a higher share of followers (73.3%) compared to male actors (n=21 and 26.7% followers) (Table 3). Their mean value is also higher than for male actors. Since actors are the most popular celebrities, female actors' higher frequency and number of followers again suggest that female celebrities might have higher popularity than male celebrities. We can observe a somewhat similar result for music-related celebrities. Most of the singers (n=34) and musicians (n=27) are male, although the mean value of female singers' followers (M=1.47 M) is higher than for the male singers' followers (M=1.03 M). Interestingly, it seems some celebrity professions are gender specific. For example, celebrities who are entrepreneurs, doctors, Islamic scholars, politicians, and YouTubers are all males, whereas only models are all females. It could also be an outcome of limited data. A chi-square test (χ^2=72.474) between celebrities' genders and professions shows a strong positive association (p<0.05).

Regarding Facebook content, celebrities mostly share photos (56.49%) on Facebook, followed by links (17.43%) and videos (14.23%) (Table 4). They are less likely to post statuses (11.85%). Most photos are related to their everyday activities and achievements, perhaps to build a

Table 3. Gender-based descriptions of the most popular celebrities

Celebrity	Gender	n	Followers				Growth			
			Followers[a)]	%	Mean[a)]	Standard deviation[a)]	Followers[a)]	%	Mean[a)]	Standard deviation[a)]
Actor	Female	45	86.64	73.3	1.92	2.09	7.13	74.6	0.20	0.25
	Male	21	31.57	26.7	1.50	1.34	2.42	25.4	0.11	0.19
Broadcast star	Female	2	1.68	15.6	0.84	0.78	0.37	52.7	0.19	0.25
	Male	2	9.04	84.4	4.52	6.01	0.34	47.3	0.17	0.24
DJ	Female	1	0.82	22.6	0.82	0.00	0.07	61.5	0.07	0.00
	Male	2	2.79	77.4	1.40	1.53	0.05	38.5	0.02	0.03
Doctor	Male	1	0.10	100	0.10	0.00	0.17	100	0.17	0.00
Entrepreneur	Male	4	0.69	100	0.17	0.11	-0.00	100	-0.00	0.00
Model	Female	6	1.81	100	0.30	0.26	0.02	100	0.00	0.00
Islamic scholar	Male	4	8.86	100	2.22	1.49	0.34	100	0.08	0.05
Musician	Female	1	0.23	1.7	0.23	0.00	0.02	8.40	0.02	0.00
	Male	27	12.93	98.3	0.48	0.55	0.17	91.6	0.00	0.01
Politician	Male	3	4.52	100	1.51	1.34	0.03	100	0.00	0.01
Singer	Female	13	19.07	35.2	1.47	1.47	2.84	45.4	0.22	0.56
	Male	34	35.13	64.8	1.03	1.66	3.42	54.6	0.10	0.42
Sports star	Female	1	1.07	1.6	1.07	0.00	0.03	1.80	0.03	0.00
	Male	16	65.68	98.4	4.11	4.35	1.62	98.2	0.10	0.10
Writer	Female	1	0.40	2.7	0.40	0.00	-0.00	-0.20	-0.00	0.00
	Male	9	14.01	97.3	1.56	2.27	0.46	100	0.05	0.09
YouTuber	Male	2	2.96	100	1.48	0.68	0.19	100	0.10	0.10

[a)]Values are in millions.

positive image in front of their fans, which is a common practice among celebrities (Abidin, 2018). Interestingly, the percentage gap between photos and other content further suggests how important and favored sharing photos is to celebrities. Links shared on Facebook contain mainly news items (Baresch et al., 2011), suggesting that celebrities share a significant percentage of news links that might or might not be relevant to themselves. Further studies are required to explore what topical content celebrities share on Facebook and how they are relevant to their public identity.

The posts from celebrities generate mainly like reactions (85.46%), followed by love (8.90%) (Table 5). Angry reactions (0.03%) are at the bottom of the list, meaning followers rarely prefer reacting angrily to the celebrities' posts. One reason for the like reaction button's highest value is that the five major reaction buttons were introduced only in 2016 (and the care reaction button in 2020), allowing Facebook users to express their diverse emotional reactions (Dewey, 2016). Before 2016, users had to express their emotions through only the like button.

Regarding expressiveness, the like button is ambiguous and difficult to interpret because no specific meaning can consistently be attributed to this button. For example, the like button can simultaneously express positivity and negativity regarding emotional valence and happiness, sadness, anger, surprise, disgust, and fear regarding specific emotional reactions (Giuntini et al., 2019). For such reasons, previous studies also excluded this button from interpretation (Jost et al., 2020); we did the same in this study.

Therefore, from the results, we can infer that the followers are most likely to express affection and liking and least likely to express disapproval and rage. Also, most of their reactions belong to positive valences, meaning followers are almost always positive about their celebrity icons. A higher percentage of comments suggests that followers express their emotions through reaction buttons and written text.

A contingency table shows that photos from celebrities generate the highest share of interactions (80.79%). Although links have a higher frequency than videos, they generate a lower share of interactions (5.22%) than video (10.55%) (Table 6). This suggests that photos have the highest frequency among all content and attract more followers. A point-biserial correlation also shows that photos and total reactions are positively associated (r_{pbis}=0.290; $p<$ 0.01), unlike others (Table 7). However, all other engage-

Table 4. Facebook content posted by celebrities

Types	n	%
Photo	5,333	56.49
Link	1,646	17.43
Video	1,343	14.23
Status	1,119	11.85
Total	9,441	100

Table 5. Descriptions of the followers' interactions

Interaction metrics	Interactions[a)]	%	Mean	Standard deviation
Likes	406.30	85.46	43,035.74	71,188.21
Love	42.32	8.90	4,482.08	18,120.84
Comments	10.97	2.31	1,161.80	4,642.80
Shares	6.32	1.33	669.08	4,282.89
Haha	3.60	0.76	381.40	2,288.23
Sad	2.50	0.53	264.57	4,468.34
Wow	2.02	0.42	213.48	821.79
Care	1.30	0.27	136.84	10,063
Angry	0.13	0.03	14.26	71.11
Total	475.44	100		

[a)]Values are in millions.

 https://doi.org/10.1633/JISTaP.2023.11.1.4

Table 6. Different interactions in different types of content

Types	Total interactions[a)]	%	Comments[b)]	Shares[b)]	Likes[b)]	Love[b)]	Wow[b)]	Haha[b)]	Sad[b)]	Angry[b)]	Care[b)]
Photo	384.13	80.79	76.0	37.0	82.1	78.4	80.6	64.2	77.1	70.1	72.8
Video	50.17	10.55	13.2	59.2	8.9	17.8	13.1	16.1	8.0	19.7	24.1
Link	24.80	5.22	4.2	2.0	5.6	1.7	3.4	17.6	0.6	8.0	1.6
Status	16.35	3.44	6.7	1.8	3.5	2.1	2.9	2.1	14.3	2.2	1.5

[a)]Values are in millions. [b)]Values are in percentages.

Table 7. Associations between content types and reactions

Types	Total interactions	Comments	Shares	Likes	Love	Wow	Haha	Sad	Angry	Care
Photo	0.290**	0.098**	-0.062**	0.312**	0.109**	0.126**	0.026*	0.025*	0.055**	0.045**
Video	-0.062**	-0.007	0.201**	-0.093**	0.025*	-0.008	0.009	-0.011	0.031**	0.038**
Link	-0.191**	-0.087**	-0.064**	-0.189**	-0.102**	-0.096**	0.001	-0.026*	-0.050**	-0.057**
Status	-0.154**	-0.040**	-0.048**	-0.157**	-0.075**	-0.072**	-0.050**	0.004	-0.060**	-0.044**

*p<0.05; **p<0.01.

ment measures except sharing have the highest percentages for photos; followers are more likely to share celebrities' videos (59.2%) than photos or other content. This further suggests that although followers interact more with photos, they tend to circulate their icons' videos more often.

The number of shares is also positively associated with only videos (r_{pbis}=0.201; p<0.01). After photos, links (17.6%) posted by celebrities receive more haha reactions, followed by videos (16.1%). It suggests that celebrities share links to something either hilarious or joyful, which is why fans express laughter, or something unacceptable, which is why fans express denial and mockery. Status updates (14.3%) from celebrities receive a higher percentage of sad reactions than videos (8%) and links (0.6%). It could be possible that celebrities' Facebook statuses often contain tragic elements so that their fans express empathy. However, the association between status and sad reactions seems statistically not significant. Of the significant associations, the correlation coefficient analysis shows that only photos' association with likes is moderate (r_{pbis}=0.312; p<0.01), while most of the other associations are either weak or negligible (Akoglu, 2018). We presume that such smaller correlation values are the consequences of outliers' influence, a common phenomenon when a study deals with a large amount of dispersed data (Grubbs, 1969; Penn State University, 2021).

5. CONCLUSION

As stated in the previous section, this exploratory study investigated the most popular Facebook celebrities' professions and genders, their popularity patterns, shared content types, and followers' reactions to them. The results highlight that actors, music-related celebrities, and sports stars are the most popular celebrities both in Bangladesh and around the globe. Male celebrities are more in numbers, but female celebrities are more popular than male celebrities in average followers. It seems some professions of celebrities are gender specific. Celebrities mostly share photos on Facebook, perhaps to recreate their public image and provide regular updates to their followers. The followers also engage the most in photos with affection and surprise reactions. Overall, followers' reactions to celebrities' Facebook content are highly positive.

Beyond these novel findings, this study is different from others, at least in two ways. First, to the best of our knowledge, no prior celebrity-related studies in Bangladesh have focused on celebrities' Facebook content, popularity patterns, and followers' reactions. Previous studies in other contexts, except for a few (Bennett, 2014; Ledbetter & Redd, 2016) were mostly concerned with celebrities' relations to various psychological and economic issues (Aljasir, 2019; Arefin, 2018; Chung & Cho, 2017; Schimmelp-

fennig, 2018; Shabnaz et al., 2019; Um, 2013). Second, no research dealt with celebrity culture and celebrity-fan communication in Bangladesh. In fact, Bangladesh, where social media and online celebrity culture are thriving, has been under-researched in celebrity studies.

Apart from these broad scholarly contributions, this study theoretically and methodologically contributes to information studies by analyzing Facebook's interaction data. Although many may consider such data as less significant metadata, it provides essential insights regarding human communication. In information studies, the priority of such interaction data is increasing, and the present study contributes to this growing body of literature. Furthermore, global scholarship of information studies mainly focuses on social media usage in prominent countries, primarily from the Global North, leaving the Global South behind the curtain. This research is a response to this academic favoritism and knowledge gap.

Despite offering some novel insights and scholarly contributions, this study has a few limitations that call for further academic investigation. First, we presume that a larger sample would yield more reliable and generalizable results. However, this study analyzed a decent amount of data, which is larger enough than many previous studies considered. For example, Gergaud et al. (2012) examined only 100 celebrities for their category-based frequencies, which is smaller than the present study. Second, another limitation of this study is related to the data source. Although SocialBakers promised to provide comprehensive data to their clients, it has some limitations. This AI-based platform could not include some influential celebrities on its list. For example, Sohel Taj, a renowned politician with 1.1 million Facebook followers, was not on the list and should have been in the 80th position. Therefore, future studies should collect and analyze more inclusive data for more comprehensive results. Third, we analyzed interaction data for only a few celebrities, which may impact the generalizability of these results. Hence, we recommend not overgeneralizing the findings of RQ3 and RQ4. Lastly, this study only investigated the context of Bangladesh. The country's unique communication infrastructure, online climate, and celebrity culture might differ from other countries, limiting the applicability of our findings to other contexts.

CONFLICTS OF INTEREST

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

REFERENCES

Abidin, C. (2018). *Internet celebrity: Understanding fame online.* Emerald Publishing.

Ahmad, Z., Kim, J., Udovica, A., & Lee, R. (2020). Comparison of fatal recreational drug overdoses between celebrities and non-celebrities. *STEM Fellowship Journal,* 6(1), 12-18. https://doi.org/10.17975/sfj-2020-003

Ahmadi, A., Fakhimi, S., & Ahmadi, Y. (2022). Instagram celebrities and positive user responses. The mediating role of user "like". *Journal of Contemporary Marketing Science,* 5(1), 65-80. https://doi.org/10.1108/JCMARS-01-2021-0002

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine,* 18(3), 91-93. https://doi.org/10.1016/j.tjem.2018.08.001

Aljasir, S. (2019). Are classic theories of celebrity endorsements applicable to new media used by Arabs? A qualitative investigation of Saudi social media users. *Journal of Creative Communications,* 14(1), 15-30. https://doi.org/10.1177/0973258618822608

Al-Rawi, A. (2020). Networked emotional news on social media. *Journalism Practice,* 14(9), 1125-1141. https://doi.org/10.1080/17512786.2019.1685902

Arefin, S. S. (2018). *The importance of celebrity endorsement & artist management for a music app to increase brand value in Bangladesh*. (Bachelor's thesis). BRAC University, Dhaka, Bangladesh.

Baresch, B., Knight, L., Harp, D., & Yaschur, C. (2011). Friends who choose your news: An analysis of content links on Facebook. *International Symposium on Online Journalism,* 1(2), 1-24. https://isoj.org/wp-content/uploads/2018/01/Baresch2011.pdf

Bennett, L. (2014). Fan/celebrity interactions and social media: Connectivity and engagement in Lady Gaga fandom. In L. Duits, K. Zwaan, & S. Reijnders (Eds.), *The Ashgate research companion to fan cultures* (pp. 109-120). Routledge.

Brosnan, M., & Gavin, J. (2015). Are "friends" electric?: Why those with an autism spectrum disorder (ASD) thrive in online cultures but suffer in offline cultures. In L. D. Rosen, N. A. Cheever, & L. M. Carrier (Eds.), *The Wiley handbook of psychology, technology, and society* (pp. 250-270). John Wiley & Sons.

Chung, S., & Cho, H. (2017). Fostering parasocial relationships with celebrities on social media: Implications for celebrity endorsement. *Psychology & Marketing,* 34(4), 481-495. https://doi.org/10.1002/mar.21001

Couldry, N. (2015). Why celebrity studies needs social theory (and vice versa). *Celebrity Studies,* 6(3), 385-388. https://doi.org/10.1080/19392397.2015.1062657

CrowdTangle. (2021). *Facebook.* https://www.crowdtangle.com/

Denis, D. J. (2018). *SPSS data analysis for univariate, bivariate, and multivariate statistics.* John Wiley & Sons.

Dewey, C. (2016). *A quick, no-nonsense guide to using Facebook's new reactions.* https://www.washingtonpost.com/news/the-intersect/wp/2016/02/24/a-quick-no-nonsense-guide-to-using-facebooks-new-reactions

Driessens, O. (2021). Celebrity and celetoid. In G. Ritzer (Ed.), *The Blackwell encyclopedia of sociology* (pp. 1-3). John Wiley & Sons.

Eberl, J. M., Tolochko, P., Jost, P., Heidenreich, T., & Boomgaarden, H. G. (2020). What's in a post? How sentiment and issue salience affect users' emotional reactions on Facebook. *Journal of Information Technology & Politics,* 17(1), 48-65. https://doi.org/10.1080/19331681.2019.1710318

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion,* 6(3-4), 169-200. https://doi.org/10.1080/02699939208411068

Freeman, C., Alhoori, H., & Shahzad, M. (2020). Measuring the diversity of Facebook reactions to research. *Proceedings of the ACM on Human-Computer Interaction,* 4(GROUP), 1-7. https://doi.org/10.1145/3375192

Freeman, C., Roy, M. K., Fattoruso, M., & Alhoori, H. (2019, June 2-6). Shared feelings: Understanding Facebook reactions to scholarly articles. In M. Bonn, D. Wu, S. J. Downie, & A. Martaus (Eds.), *Proceedings of the 2019 ACM/IEEE Joint Conference on Digital Libraries (JCDL)* (pp. 301-304). IEEE.

Geboers, M., Stolero, N., Scuttari, A., Van Vliet, L., & Ridley, A. (2020). Why buttons matter: Repurposing Facebook's reactions for analysis of the social visual. *International Journal of Communication,* 14, 1564-1585. https://ijoc.org/index.php/ijoc/article/view/11657

Gergaud, O., Ginsburgh, V., & Livat, F. (2012). Success of celebrities: Talent, intelligence or beauty? *Economics Bulletin,* 32(4), 3120-3127. https://ideas.repec.org/a/ebl/ecbull/eb-12-00415.html

Giuntini, F. T., Ruiz, L. P., Kirchner, L. D. F., Passarelli, D. A., Dos Reis, M. D. J. D., Campbell, A. T., & Ueyama, J. (2019). How do I feel? Identifying emotional expressions on Facebook reactions using clustering mechanism. *IEEE Access,* 7, 53909-53921. https://doi.org/10.1109/ACCESS.2019.2913136

Grubbs, F. E. (1969). Procedures for detecting outlying observations in samples. *Technometrics,* 11(1), 1-21. https://doi.org/10.2307/1266761

Hou, M. (2019). Social media celebrity and the institutionalization of YouTube. *Convergence,* 25(3), 534-553. https://doi.org/10.1177/1354856517750368

Jacobsen, L. B. (2020). Vitafiction and virality: Celebrities fictionalizing the self online. *Convergence,* 26(4), 912-926. https://doi.org/10.1177/1354856518818081

Jost, P., Maurer, M., & Hassler, J. (2020). Populism fuels love and anger: The impact of message features on users' reactions on Facebook. *International Journal of Communication,* 14, 2081-2102. https://ijoc.org/index.php/ijoc/article/view/13400

Kim J., & Song, H. (2016). Celebrity's self-disclosure on Twitter and parasocial relationships: A mediating role of social presence. *Computers in Human Behavior,* 62, 570-577. https://doi.org/10.1016/j.chb.2016.03.083

Kim, M., & Kim, J. (2020). How does a celebrity make fans happy? Interaction between celebrities and fans in the social media context. *Computers in Human Behavior,* 111, 106419. https://doi.org/10.1016/j.chb.2020.106419

Kowalska, M., & Wróbel, M. (2017). Basic emotions. In V. Zeigler-Hill, & T. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (pp. 1-6). Springer.

Lau, D. W. S. (2021). Aamir Khan and celebrity humanitarianism in Asia: Towards a cosmopolitical persona. *Celebrity Studies,* 12(2), 234-249. https://doi.org/10.1080/19392397.2021.1912193

Ledbetter, A. M. & Redd, S. M. (2016). Celebrity credibility on social media: A conditional process analysis of online self-disclosure attitude as a moderator of posting frequency and parasocial interaction. *Western Journal of Communication,* 80(5), 601-618. https://doi.org/10.1080/10570314.2016.1187286

Mamur, N. (2015). Critical pedagogical approaches to visual culture in Turkish preservice education. *Studies in Art Education,* 56(4), 355-368. http://www.jstor.org/stable/45185098

Pennsylvania State University. (2021). *Cautions about correlation and regression.* https://online.stat.psu.edu/stat800/lesson/cautions-about-correlation-and-regression

Raisa, T. S. (2018). Impact of Facebook obsession among university students in Bangladesh. *International Journal of New Technology and Research,* 4(2), 89-94. https://www.ijntr.org/download_data/IJNTR04020063.pdf

Rojek, C. (2015). Celebrity. In D. T. Cook, & J. M. Ryan (Eds.), *The Wiley Blackwell encyclopedia of consumption and consumer studies* (pp. 1-3). John Wiley & Sons.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology,* 39(6), 1161-1178. https://doi.org/10.1037/h0077714

Sakib, S. N. (2020). *Bangladesh: Celebrities, youth resolve to fight COVID-19.* https://www.aa.com.tr/en/asia-

pacific/bangladeshcelebrities-youth-resolve-to-fight-covid-19/1779034

Schimmelpfennig, C. (2018). Who is the celebrity endorser? A content analysis of celebrity endorsements. *Journal of International Consumer Marketing,* 30(4), 220-234. https://doi.org/10.1080/08961530.2018.1446679

Selkie, E. (2022). Influence at the intersection of social media and celebrity. *JAMA Network Open,* 5(1), e2143096. https://doi.org/10.1001/jamanetworkopen.2021.43096

Shabnaz, S., Shajahan, B., Maryam, H., & Shovon, A. R. (2019, September 18-20). Celebrity endorsement in advertisement: A study on consumers' perception for selected personal care products in Bangladesh. *The Proceedings of the 2nd International Conference on Business and Management (ICBM 2019)* (pp. 271-274). BRAC University.

St. James, M. (2010). Female sports celebrities targeting female teenagers: A content analysis of magazine advertising. *Journal of Business & Economics Research,* 8(1), 1-14. https://doi.org/10.19030/jber.v8i1.653

StatCounter. (2020). *Social media stats Bangladesh.* https://gs.statcounter.com/social-media-stats/all/bangladesh

Tran, T., Nguyen, D., Nguyen, A., & Golen, E. (2018, May 20-24). Sentiment analysis of marijuana content via Facebook emoji-based reactions. In D. Huang, & S. Secci (Eds.), *Proceedings of the 2018 IEEE International Conference on Communications (ICC)* (pp. 1-6). IEEE.

Um, N. H. (2013). The role of culture in creative execution in celebrity endorsement: The cross-cultural study. *Journal of Global Marketing,* 26(3), 155-172. https://doi.org/10.1080/08911762.2013.804613

Varanasi, R. A., Dicicco, E., & Gambino, A. (2018). Facebook reactions: Impact of introducing new features of SNS on social capital. In C. Stephanidis (Ed.), *HCI International 2018 – Posters' extended abstracts* (pp. 444-451). Springer.

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

RESEARCH PAPER
J Inf Sci Theory Pract 11(1): 61-78, 2023

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2023.11.1.5

GOMME: A Generic Ontology Modelling Methodology for Epics

Udaya Varadarajan
Documentation Research and Training Centre (DRTC), Indian Statistical Institute, Bangalore Centre, India
Department of Library and Information Science, University of Calcutta, Kolkata, India
E-mail: udayav@drtc.isibang.ac.in

Mayukh Bagchi
Department of Information Engineering and Computer Science (DISI), University of Trento, Trento, Italy
E-mail: mayukhbagchi@drtc.isibang.ac.in

Amit Tiwari*
Documentation Research and Training Centre (DRTC), Indian Statistical Institute, Bangalore Centre, India
Department of Library and Information Science, University of Calcutta, Kolkata, India
E-mail: amittiwari@drtc.isibang.ac.in

M. P. Satija
Department of Library and Information Science, GNDU, Amritsar, India
E-mail: satija_mp@yahoo.com

ABSTRACT

Ontological knowledge modelling of epic texts, though being an established research arena backed by concrete multilingual and multicultural works, still suffers from two key shortcomings. Firstly, all epic ontological models developed till date have been designed following ad-hoc methodologies, most often combining existing general purpose ontology development methodologies. Secondly, none of the ad-hoc methodologies consider the potential reuse of existing epic ontological models for enrichment, if available. This paper presents, as a unified solution to the above shortcomings, the design and development of GOMME - the first dedicated methodology for iterative ontological modelling of epics, potentially extensible to works in different research arenas of digital humanities in general. GOMME is grounded in transdisciplinary foundations of canonical norms for epics, knowledge modelling best practices, application satisfiability norms, and cognitive generative questions. It is also the first methodology (in epic modelling but also in general) to be flexible enough to integrate, in practice, the options of knowledge modelling via reuse or from scratch. The feasibility of GOMME is validated via a first brief implementation of ontological modelling of the Indian epic *Mahabharata* by reusing an existing ontology. The preliminary results are promising, with the GOMME-produced model being both ontologically thorough and competent performance-wise.

Keywords: ontology development methodology, GOMME methodology, modelling epics, *Mahabharata* ontology, ontology reuse methodology, digital humanities

Received: July 21, 2022 **Revised:** December 2, 2022
Accepted: January 10, 2023 **Published:** March 30, 2023

***Corresponding Author:** Amit Tiwari
https://orcid.org/0000-0003-4147-3931
E-mail: amittiwari@drtc.isibang.ac.in

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far 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.

© Udaya Varadarajan, Mayukh Bagchi, Amit Tiwari, M. P. Satija, 2023

1. INTRODUCTION

The most cited definition of ontology is that it is "an explicit specification of a shared conceptualization" (Gruber, 1995). Ontology development methodologies are practices that focus on the life cycle of ontologies, their development process, and the techniques, tools, and languages employed to create them (Gómez-Pérez et al., 2004). The research arena of Knowledge Representation (KR) has witnessed a plethora of knowledge modelling methodologies exploiting a diverse array of KR formalisms including semantic nets (Lehmann, 1992), frames (Lassila & McGuinness, 2001), ontologies (Guarino et al., 2009; Phoksawat et al., 2019) and now knowledge graphs (Bagchi & Madalli, 2019; Ehrlinger & Wöß, 2016). Such methodologies can be generic (Fernández-López et al., 1997; Noy & McGuinness, 2001) or tailor-designed for specialized domains such as for health (Das & Roy, 2016), machine translation (Mahesh et al., 1996), and smart cities (Espinoza-Arias et al., 2019). Tailor designed or ad-hoc methodologies most often combine an existing general methodology or methodologies. They are created as and when necessary, in particular, for modelling knowledge in diverse arenas of digital humanities such as folklore (Abello et al., 2012), narrative information (Damiano & Lieto, 2013; Swartjes & Theune, 2006), and epics and mythology (Hyvönen et al., 2005; Iqbal et al., 2013; Syamili & Rekha, 2018).

1.1. Motivation

As mentioned in the introduction, the ad-hoc approaches (for instance, in Syamili & Rekha, 2018) inherit the architectural characteristics and advantages of the combinations of established methodologies. But they suffer from three crucial shortcomings specific to modelling epic texts and other genres in digital humanities in general. Firstly, as observed from the literature study, there exists no methodology with a focus on modelling and representing epics (Greene, 1961) and on works in digital humanities in general. A generic ontology development methodology will not be satisfactory, because modelling epics requires classificatory and ontological elegance and should adhere to well-established literary cardinals (similarly for other literary genres). Secondly, the existing models across the domains in digital humanities do not discuss potential reuse of existing ontological formalizations. Thirdly, in accordance with dynamism in knowledge organization and representation (Giunchiglia et al., 2014; Kent, 2000), none of the approaches for modelling domains in digital humanities stress iterative knowledge development and modelling at the methodological level itself. This is important to ensure that the model is constantly perfected from the feedback mechanism within. The motivation of the present work lies in the above shortcomings.

1.2. Novelty

In response to the aforementioned challenges, this work develops and proposes GOMME. The novelty of the work lies in being the first dedicated methodology for iterative ontological modelling of literary epics. The other novel factors of the work are: Firstly, GOMME is grounded in the novel characteristic design cardinal of repurposing and flexibility, which makes the methodology (or appropriate fragments of it) completely customizable and extensible for modelling domains in (digital) humanities and social sciences beyond epics. Secondly, the methodology is not only grounded in knowledge modelling best practices and application satisfiability norms, but also adheres to canonical norms for epics (Greene, 1961) which makes it innately suitable for modelling epics while, at the same time, preserving (some of) its unique literary features (detailed in Section 3). Thirdly, the novel fact that GOMME, in practice, offers the flexibility of not only modelling an epic from scratch but also is natively grounded in the ontology reuse paradigm. It thereby accommodates reuse and potential enrichment of any formal or conceptual model of an epic, if available.

The remainder of the paper is organized as follows: Section 2 details the state-of-the-art frameworks in modelling literary works and narrative information, and highlights the research gaps. Section 3 describes, in fine detail, the foundations and steps of the GOMME methodology. Section 4 presents the proof of the methodology via a brief case study of modelling the Indian epic *Mahabharata*, showing the feasibility and advantages of the GOMME methodology. Section 5 discusses some notable implications of the work and Section 6 concludes the paper with a discussion of future prospects of the present work.

2. LITERATURE REVIEW

Ontology development methodologies are "activities that concern with the ontology development process, the ontology life cycle, and the methodologies, tools, and languages for building ontologies" (Gómez-Pérez et al., 2004). For over two decades, research has been happening in the domain of methodologies for ontology development. For the current work, the most commonly reused

 https://doi.org/10.1633/JISTaP.2023.11.1.5

and acclaimed methodologies were studied. Ontology Development 101 (Noy & McGuinness, 2001) is an initial guide for amateur ontology developers; NeON methodology (Suárez-Figueroa et al., 2015) focuses on building large-scale ontologies with collaboration and reuse; and DILIGENT (Vrandečić et al., 2005) is a methodology that lays emphasis on ontology evolution and not on initial ontology designing. TOVE (Gruninger & Fox, 1995a) highlights ontology evaluation and maintenance. ENTERPRISE (Uschold & King, 1995) discusses the informal and formal phases of ontology construction sans identification of the ontology concepts. Yet Another Methodology for Ontology (YAMO) (Dutta et al., 2015) is a methodology for ontology construction for large-scale faceted ontologies. The principles of facet analysis and analytico-synthetic classification guide the YAMO methodology. The methodology was used in the construction of ontology for food.

Following this, grounded in the general methodologies as aforementioned, a literature search was conducted to identify specific methodologies for narrative/literary ontological models, via the models developed in the similar domains. These works are similar in scope to our objective of examining the state-of-the-art for modelling epics. The existing models do not follow any standard, matured ontology development methodologies. Instead, they have ad-hoc methodologies. Varadarajan and Dutta (2021a) conducted a study of such models (Table 1 details the same). The study discovered a lack of any methodology dedicated to the development of narrative/literary ontologies.

We now focus on briefly elucidating some of the above works. Constructed on the basis that the narrative situation (Klarer, 2013) needs characters and objects which form a larger story once connected, the ontology by Mulholland et al. (2004) provides intelligent support for the exploration of digital stories to encourage heritage site visits. Hyvönen et al. (2005) develop a framework and national ontology for representing the culture and heritage of Finland on the Semantic Web. Developed as an upper ontology, it is based on the Finnish General Thesaurus. The National Library of Finland is the competent authority for its maintenance. The ontology by Nakasone and Ishizuka (2006) is constructed with the generic aspects of

Table 1. Selected ontology models for study

Ontology name	Purpose	Domain	Ontology design methodology
Ontology for story fountain (Mulholland et al., 2004)	To describe stories and related themes	Literature	Not available
Finnish national ontology (FinnONTO) (Hyvönen et al., 2005)	To express the national Finnish Thesaurus as ontologies	Cultural heritage	Not available
Ontology model for storytelling (Nakasone & Ishizuka, 2006)	To build a coherent event centric generic storytelling model	Literature	Not available
The archetype ontology (AO) (Damiano & Lieto, 2013)	To link the various resources in the archive through narrative relations	Cultural heritage	Not available
Quran ontology (Iqbal et al., 2013)	To develop ontology that provides contextual information for the correct interpretation of Quran	Religious text/Epic	Mixed existing methodologies
ODY-ONT (Khan et al., 2016)	Explicit representation of the story found in any text	Epic	Not available
Transmedia fictional worlds ontology (Branch et al., 2017)	To represent the elements of the fictional world, connecting characters, places, and events	Literature	Not available
BK Onto (Yeh, 2017)	To capture biographical information	Literature	Not available
Ontology for Greek mythology (Syamili & Rekha, 2018)	To develop an ontology for Greek mythology	Mythology/Epic	Mixed methodologies
Drammar ontology (Damiano et al., 2019)	Represent drama elements independent of media and task formally	Literature	NeOn

storytelling as its founding philosophy. The purpose of a domain independent model was to provide coherence to the events in the story. The Archetype Ontology (Damiano & Lieto, 2013) is built to explore a digital archive via narrative relations among the resources. To query and retrieve contextual information along with Quran verses, an ontology was developed by Iqbal et al. (2013). A merging of various methodologies was adopted to model this ontology. To represent narration in a literary text, the ODY-Onto (Khan et al., 2016) was constructed. The developed ontology is part of a system constructed for querying information from literary texts. Transmedia ontology (Branch et al., 2017) allows users to search for and retrieve the information contained in the transmedia world. The ontology will help in inferring connections between transmedia elements such as characters, elements of power associated with characters, items, places, and events. The ontology contains 72 classes and 239 properties. Biographical Knowledge Ontology (BK onto) was created to capture biographical information. The ontology was deployed in the Mackay Digital Collection Project Platform (http://dlm.csie.au.edu.tw) for linking "the event units to correlate to the contents of external digital library/archive systems so that more diverse digital collections can be presented in StoryTeller system" (Yeh, 2017). Ontology for Greek Mythology (Syamili & Rekha, 2018) developed an ontology for the relationship among Greek mythological characters. For ontology development, a combination of the existing methodologies was reused. The Drammar ontology (Damiano et al., 2019) was developed to represent the elements of drama independent of the media and task. Drama, as a domain, is evolving, but there is a concrete manifestation of drama such as in screenplays, theatrical performances, radio dramas, and movies. The top four classes of dramatic entities are (1) DramaEntity, the class of the dramatic entities, i.e., the entities that are peculiar to drama; (2) DataStructure, the class that organizes the elements of the ontology into common structures; (3) DescriptionTemplate, which contains the patterns for the representation of drama according to role-based templates; and (4) ExternalReference, the class that bridges the description of drama to commonsense and linguistic concepts situated in external resources.

Observations from the literature study indicate that the generic ontology development methodologies involve domain identification, term collection, relationship and attributes establishment, representation, and documentation. These methodologies are for ontology development from scratch. Rarely do such methodologies integrate the reusing of existing ontologies. Similar is the case with theoretically established methodologies.

The methodologies used for ontology construction in narrative/literary domains, on the other hand, are ad-hoc and assume previous knowledge in ontology construction. The steps in the methodologies often start with the ontology population. Further investigation found that the ontology development methodologies for narrative/literary domains lack an iterative process and that these methodologies are mostly linear. A linear methodology is unidirectional and does not include traversing in a feedback loop, therefore leaving very little space for improvement. Keeping these research gaps in mind, the following section will detail our proposed GOMME methodology.

3. THE *GOMME* METHODOLOGY

This section elucidates the GOMME methodology in detail. In Section 4, a first brief implementation of GOMME for modelling the Indian epic *Mahabharata* is detailed. The *Mahabharata*, the longest epic in the history of world literature, was originally written by Ved Vyasa in Sanskrit. It comprises around 100,000 stanzas (Lochtefeld, 2002). The *Mahabharata* narrates the story of a kingdom called Hastinapur, where two brothers' families, Pandavas and Kauravas, fought for control over the kingdom. Each of the sides was directly backed up by several other kingdoms. In the end, Vyasa illustrated the war between the two sides, and the Pandavas won with the help of Lord Krishna. Several connecting mythological stories make *Mahabharata* more comprehensive (Rajagopalachari, 1970). The stages of GOMME are discussed as follows.

3.1. Initial Scoping Exercise

The initial scoping exercise (see Fig. 1) marks the inception for modelling epic texts in the context of the GOMME methodology. The step is an exercise in consolidating a set of objective-focused design guidelines for initiating, scoping, and informing the steps of the methodological architecture in an all-encompassing manner. A non-exhaustive candidate set of such design concerns was identified and discussed as follows: Primary design concern is scopic review, i.e., to achieve a shared convergence in delineating the existing formal modelling attempts in the chosen (epic) domain, culminating in research gaps yet to be addressed. The next was addressing feasibility and novelty of research ideation, which involves brainstorming and federating the theoretical novelty and implementational feasibility of the proposed research

 https://doi.org/10.1633/JISTaP.2023.11.1.5

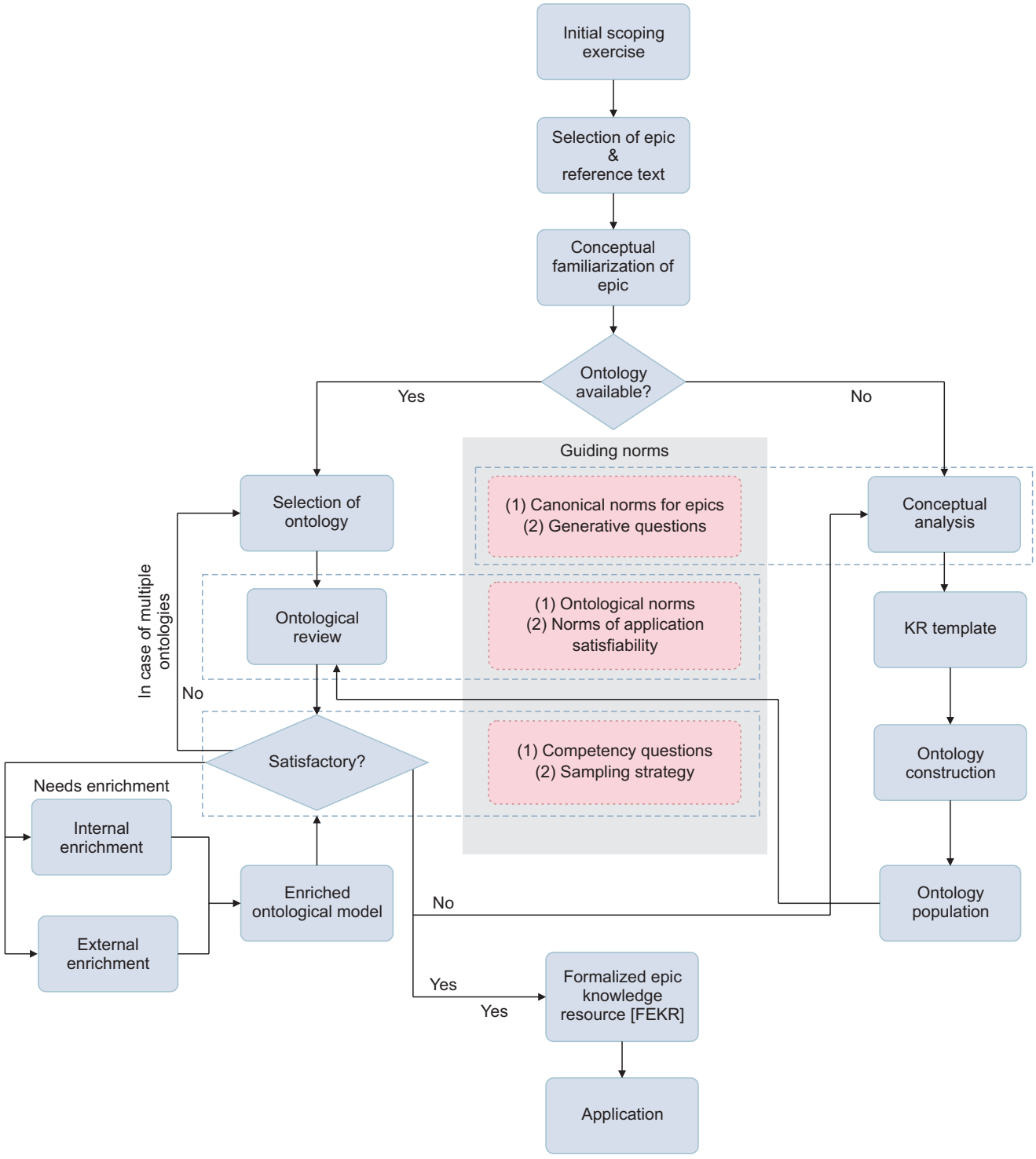


Fig. 1. The GOMME methodology. KR, knowledge representation.

ideation in sync with the research scope as understood in the scopic review. The third design concern was related to technologies and optimization, which is an attempt at a broad identification of conceptual tools (often transdisciplinary) and technological paradigms required for optimal effectuation of the architecture. The next design concern was architectural agility, meaning to ground the overall architectural flow in the notion of technological agility, thus

realizing it as not only methodologically a-technological per se but also incremental, adaptive, and scalable (for instance, via maturity models; see Tiwari & Madalli, 2021 for a recent study). The final concern was purposing and extensibility, i.e., to incorporate the scope for repurposing the architectural components (or the architecture in its entirety), thereby facilitating its extension in varied contexts situated within the scope of knowledge modelling.

The above-mentioned design concerns are indicative of generic aspects which might be considered in the initial scoping exercise and are completely flexible in allowing added tailored formulations for specific research contexts, for instance, if we extend for other categories of works in the digital humanities. The methodology recommends the involvement of the entire Project Development Team (PDT), by which we mean the team responsible for end-to-end coordination of the GOMME methodology for a particular project, for the initial scoping exercise.

3.2. Selection of Epic and Reference Text

An epic usually transgresses commonplace literary praxis and might dwell in conventions as varied as proposition, invocation, enumeration, or in medias res (Cooren, 2015). Therefore, mindful selection of an epic and a related reference text is crucial for robust development and evolution of the proposed methodological architecture. The process of deciding on an epic must involve the following conventions in varying degrees. The conventions are cultural affinity, literary expertise, and computational amenability. By cultural affinity, it is recommended to select an epic which embodies, to a certain extent, cultural coherence and rootedness with (at least some of) the PDT members. Literary expertise indicates that the selection of an epic should also factor in the availability of requisite literary expertise, without which an inceptive understanding and subsequent conceptual analyses will prove knotty; and computational amenability implies that the selected epic should have the ability to elicit, engineer, actualize, and manipulate its specific representation computationally. The convention is non-trivial; for instance, the Buddhist epic *Mahavamsa* written in the Pali language, a language for which literary expertise is scant, thus potentially has low computational amenability.

Similarly, the approval of a standard reference text(s) for the selected epic should also be based on certain non-normative norms, the core of which are scholarly standing, contributor foothold, equilibrium of comprehensiveness, and scholarly currency. Scholarly standing means that the chosen reference text is expected to have high standing amongst relevant scholarly communities. Book reviews, exegeses, and bibliographic citations can be some parameters cumulatively determining a text's scholarly authority. The contributor foothold refers to the scholarly reputation of the contributors (editors, translators, etc.) to the reference text. In general, it is best to select the work of those contributors who are noted to have proven expertise and intellectual acumen in the chosen epic. The equilibrium of comprehensiveness addresses the selection of reference text in such a fashion that it exhibits a healthy equilibrium between brevity of message, description, and all-encompassing coverage of the crux of the epic storyline. It is in line with our focus on achieving a well-balanced knowledge model of the epic as opposed to any model which is constrained, pragmatically, due to detailed literary analysis. Lastly, scholarly currency of the selected reference text should be considered. Ideally the text should be in vogue amongst relevant scholarly communities. Details such as edition number and reprint number can be good markers in assessing the scholarly currency of such a text.

3.3. Conceptual Familiarization of Epics

The next requirement is for an inceptive, conceptual familiarization (see Fig. 1) of the epic. Such an exercise should involve a preliminary understanding of the storyline at the idea plane (Ranganathan, 1967). It should also involve a preliminary understanding of the activities involved in the epic in a non-rigid fashion and intellection of the implications that might arise while giving shape to its knowledge model. The primary cardinal aspect which informs the current architectural component is exploratory intuition. It skims through the chosen reference text in order to arrive at holistic, rapid insights about various entities and activities they shape and inform, as embedded within the plot of the epic. The other cardinal aspect is a reflective discussion that involves discussing, reflecting, and coalescing individual insights gained to arrive at a shared understanding. This will be the pivotal groundwork from the requirements perspective as the architecture steps into its more formal components.

The recommendation is to model the conceptual familiarization effort as exhaustive sets of Competency Questions (CQs) (Grüninger & Fox, 1995b), possibly in collaboration with potential users of varying scholarly expertise in the chosen epic.

3.4. Ontology Availability

Once a shared comprehension of the plot of the epic is attained, it is pertinent to first search for available on-

 https://doi.org/10.1633/JISTaP.2023.11.1.5

tological formalizations of the same epic broadly aligned with intermediate research needs. This should be in conformance with, but not exclusively, the ontology reuse paradigm (Carriero et al., 2020; Pinto & Martins, 2000; Simperl, 2009). The sources considered for searching the availability of such formal ontologies can be from amongst:

- Cross-domain ontology repositories which might host ontologies on epics appropriately exposed via requisite metadata annotations, and/or
- Relevant research paper(s), which might outline the skeletal framework and developmental nuances of an ontology for the selected epic, and/or
- Research content platforms, which might be websites hosting such ontologies developed within the scope of a specific research project or research grant.

Depending on the availability of the ontology (i.e., available or not available) there can be only two outcomes of this decision block. The first is the availability of ontologies (YES). This means that one or more ontological formalizations of the epic are available in conformance to our overall research requirements as understood from conceptual familiarization, i.e. as further detailed in Section 3.4.1. The second outcome is availability of ontologies (NO). This means that no formal ontological representation of the chosen epic is found; in that case we start activity as further detailed in Section 3.4.2.

3.4.1. Availability of Ontologies (YES)

In case one or more formal ontologies related to the chosen epic are available and found, the methodology prescribes the following steps sequentially:

3.4.1.1. Selection of Ontology.

This step concerns the selection of an ontology from amongst the available ontologies found, to be considered as an input to the next step. The methodology does not stipulate any convention for the selection at this stage (in case of multiple available ontologies). It might be chosen at random or as per moving criterion set collaboratively by the PDT. In the case of the availability of a single ontology relating to the chosen epic, the task is trivial. The motivation for this step and related subsequent steps is to model our ontology following the direct reuse paradigm in ontology reuse (Carriero et al., 2020).

3.4.1.2. Ontological Review.

After selecting the ontology from the previous step, this step reviews the selected ontology. The principal intuition behind this step is to check whether the selected ontology conforms to established standards in formal ontology engineering and exhibits requisite ontological commitment (Guarino & Welty, 2002). It also involves a general qualitative examination of the ontology to ascertain whether it is suitable for the application for which it is developed. There are two guiding norms for a comprehensive review, namely, ontological norms and norms of application satisfiability.

The former is the first set of guidelines. It is a mix of the OntoClean best practices (Guarino & Welty, 2002) and Ranganathan's classificatory principles (Ranganathan, 1967). Ontological norms are for reviewing sound ontological conformance. During the ontology review, firstly, we examine the classes and properties of the selected ontology as per their context, such as whether a particular axiomatization of an object property should really be such and not a data property specific to the current requirements, etc. Secondly, with respect to OntoClean, a lightweight ontological analysis can be carried out by examining the ontology elements, namely, classes, object properties, and data properties. This analysis should be from the perspective of essence, rigidity, identity, and unity as appropriate. Thirdly, and most importantly, a guided validation of the subsumption hierarchies in the ontology following Ranganathan's principles is highly recommended. Finally, the exact determination of the mix depends on the requirements of the PDT with respect to the targeted level of ontological conformance.

The latter is the second set of guidelines prescribing an assessment of the selected ontology's suitability for the application scenario for which it is modelled or enhanced. The concrete norms for this step are best left to be articulated by the PDT. Some general norms for such an activity which can be reused and improved upon are:

- Whether the considered ontological schema is amenable to be utilized as a classification ontology (Zaihrayeu et al., 2007), a descriptive ontology for metadata management applications (Satija et al., 2020), or a domain-linguistic ontology, accounting in the objectives of the application scenario;
- Whether the ontology is aligned with any standard, upper ontology in the specific instance of data integration (Giunchiglia et al., 2021a) as an application scenario;

- Whether it reuses concepts from any general-purpose, core, or domain-specific ontological schemas which are considered standard or de facto standard in the ontology engineering research community.

3.4.1.3. Satisfactory Examination.

After the review of intermediate ontology, the methodological flow now concentrates on an active assessment of its suitability to the project requirements (Gómez-Pérez, 2004). To that end, there can be three possible outcomes, which are Yes, No, and Needs Enrichment.

If the reviewed ontology satisfies the examination, i.e., case of Yes, the intermediate ontology is fully aligned with the ontological commitment as expected from the project requirements (modelled as CQs). In this case, the intermediate ontology is: (i) rechristened as the Formalized Epic Knowledge Resource (FEKR) (see 3.5), and (ii) it is production-ready, to be exploited in applications such as management of semantic annotations.

If the reviewed ontology does not satisfy the examination, i.e., case of No, the intermediate ontology fails mostly in its alignment to the ontological commitment as expected from the project requirements. In this case, there are two routes prescribed by the GOMME methodology. The first is, in case of multiple available ontologies of the epic, the methodological flow loops back to (3.4.1.1) to consider the suitability of the next ontology as decided by the PDT. The second is, in case this ontology was the single artefact considered in (3.4.1.1), the methodological flow shifts to the conceptual analysis step, which is a crucial step to design a new ontology of the chosen epic as per the expressed ontological commitments.

If the reviewed ontology partially satisfies the examination, i.e., in case of needing enrichment, the intermediate ontology is satisfactorily aligned as expected from the project requirements but still needs to be enriched in order to attain maximal coverage. The recommended modes of enrichment are external and internal.

- External enrichment is where an in-depth conceptual familiarization of the epic is collaboratively arrived at and compared with both the ontological commitment modelled in the intermediate ontology and the project requirements. Subsequently, the difference in ontological commitment in terms of classes, object properties, data properties, and other relevant axioms are modelled and integrated in the intermediate ontology.
- Internal enrichment is where possible modelling bugs in the ontology such as mispositioning of classes, object properties and data properties, overloading of property restrictions, or *is-a* overloading are resolved.

Finally, the inclusion of external and/or internal enrichment results in an enriched ontological model which again undergoes the suitability test. The iteration continues till the ontological commitment of the enriched model maximally converges with that of the project requirements, as deemed appropriate by the PDT. Once the enriched model achieves maximal convergence, it is deemed as a FEKR which is application ready.

Most importantly, the assessment as to whether the intermediate ontology is satisfactory or not can be determined by evaluating the model against CQs (defined in Section 3), formalized as SPARQL queries (DuCharme, 2013) or even via appropriate sampling survey approaches such as stratified sampling (Parsons, 2014). The strategy is to keep enriching the ontology till it achieves maximal coverage of the project requirements. The exact threshold leading to the determination of maximal, satisfactory, or unsatisfactory alignment as from above is in essence an experiential learning issue for each PDT and is still an open question to be stated generically, if at all. We plan to widely adopt and popularise GOMME in the coming years towards iterative formal modelling of different epics and literary works, each of whose thresholds we expect to be different as of now.

3.4.2. Availability of Ontologies (NO)

Orthogonal to (3.4.1), if no formal ontology related to the chosen epic was found, or, in the case where a single ontology was found but proved unsuitable, the GOMME methodology prescribes the following sequential approach (see Fig. 1) to design an ontology from scratch for the chosen epic:

3.4.2.1. Conceptual Analysis.

The essence of this step is to build a detailed, in-depth, and formally documented shared conceptual understanding of the epic on top of the initial lightweight conceptual familiarization. We find literary justification for such in-depth analysis for knowledge modelling from the following canonical norms for epics which we derived from the genre-theoretic interpretation of epics (Greene, 1961): expansiveness, distinctions, tonality, and feeling.

Expansiveness is the literary quality of an epic to "extend its own luminosity in ever-widening circles" (Greene,

 https://doi.org/10.1633/JISTaP.2023.11.1.5

1961). While distinctions are what exist between the "director and executor of action" (Greene, 1961), in other words, not only factoring in actions but also actors and actants (Ranganathan, 1967) specific for knowledge modelling are considered. Tonality and feeling are needed to capture the "heroic energy, superabundant vitality which charges character and image and action" (Greene, 1961), and the fact remains that causation is much less important for epic modelling than tonality or feeling.

For concretely performing the conceptual analysis, the work focuses on the following two dimensions. First is elicitation via Generative Questions (GQs), which is the conceptual extraction of ontology elements, viz. classes, object properties, and data properties (subsuming actors, actants, and actions) from an in-depth and shared understanding of the epic. Evidently, this approach goes beyond a mere preliminary familiarization with its storyline. The recommendation is to record the information elicited as formal documentation which is compatible with the PDT members. It is important to note that to achieve true conceptual elicitation grounded in expansiveness, distinctions, tonality, and feeling of the epic, we exploit GQs (Vosniadou, 2002), first proposed in knowledge modelling by Bagchi (2021a). GQs reflect conceptualization rooted in human cognition (Giunchiglia & Bagchi, 2021) and cognitive theory of mental models (Johnson-Laird, 1983). These are questions directed at the incremental, dynamic understanding of knowledge domains being modelled, and thus "cannot elicit ready-made solutions" (Bagchi, 2021a). At each iteration of knowledge modelling, GQs "challenge the existent mental model of the domain and improve it" (Bagchi, 2021a). The crucial distinction between GQs with CQs (Grüninger & Fox, 1995b) is the fact that while the latter is focused on adjudicating the competency of the final knowledge model, the former is focused on the iterative conceptual mental model development of the knowledge domain being modelled, and consequently, also of the final knowledge model.

In addition to grounding the conceptual analysis in the canonical norms for epics and GQs, GOMME also recommends utilising and extending the same formal documentation by modelling ontology elements as per standard top-level ontological distinctions. This renders the subsequent formal ontological design and commitment more concrete. For example, the conceptual entities can be characterised as endurants or perdurants, depending on whether the entity is wholly present in time or not (Gangemi et al., 2002). This becomes non-trivial, for instance, when we wish to capture the different facets of an individual in an epic via the roles (Masolo et al., 2004) they play in different sub-contexts of the epic.

3.4.2.2. Knowledge Representation Template.

Alongside the documentation of conceptual analysis, the GOMME methodology prescribes the design of a KR template (see Fig. 2). A KR template aspectually captures the conceptual analysis recorded structurally. It records the actors, actants, and actions via the properties encoded in the contextual recording of actors and actants. The design of such a template should mandatorily (but not only) include the following:

- An enumeration of all the relevant characters of the epic which the PDT wants to model and integrate with the final FEKR, cumulatively grounded in their incremental generation via GQs and inclusion via CQs from the competency perspective.
- The definition, primary, filial, or contextual identity of the characters enumerated in the above step as aptly determined by the PDT.
- The secondary relations among each such character in the epic, thus recording a major proportion of the actions on which the epic is built. A crucial observation is the fact that secondary relations among characters can be potentially many, but the primary definition is only one (all being determined in sync with the objectives of the PDT).

3.4.2.3. Ontology Construction.

Once the design and population of the KR template are completed, the methodological flow concentrates on developing the formal ontological schema from the documented conceptual analysis. The key observation here is that the ontological schema, at this level, focuses on modelling the top-level classes and properties. It postpones the population of the schema with entities from the KR tem-

Sl. No:	Characters	Primary Definitions (of each character)	Secondary Relation (among characters)	
48	Dhritarashtra	kingOf Hastinapur	sonOf Vichitravirya and Ambika	performed VaishnavaSacrifice
49	Draupadi	wifeOf Pandavas	daughterOf Drupada	
50	Drona	teacherOf Kauravas and Pandavas	killed Abhimanyu	
51	Drupada	fatherOf Draupadi	kingOf Panchala	
52	Duhsasana	brother of Duryodhana	killedBy Bhima	
53	Durdhara	brotherOf Duryodhana	killedBy Bhima	

Fig. 2. An example of a fragment of a KR template. KR, knowledge representation.

plate to the later stage to keep distinct the schema layer and data layer at the methodological level itself (Giunchiglia et al., 2021b). The ontological schema is constructed via any state-of-the-art open source ontology editor (such as Protégé) and adheres to the generic ontology development steps of designing the formal hierarchies following:

- Class subsumption hierarchy, wherein the top-level classes for the ontological schema of the epic are modelled,
- Object property hierarchy, wherein the properties interrelating the classes are modelled, with the constraint that each such object property should mandatorily have a domain and range axiomatized,
- Data property hierarchy, wherein the attributes appropriate to be captured and axiomatized for classes are modelled, each having a class as its domain and a datatype as its range.

3.4.2.4. Ontology Population.

Given the availability of the KR template and the formal ontological schema, the current step maps the entity-level data of the KR template to the formal ontological schema developed. There are three options depending upon the nature of the ontology developed and the scope within which the PDT is working, which can be leveraged to perform ontology population. The first option is the manual mapping of data from the KR template to the ontology. This is apt for epics which are limited in scope from the perspective of entities. If the epic is considerably expanded in its scope in terms of data or entities, the second option can be to exploit appropriate plugins (such as cellfie for Protégé) to map the data to their corresponding ontology concepts. This mapping can be via definition of appropriate transformation rules. In the case of an epic/literary work of extensive magnitude (big data, in popular parlance) which requires extended man-hours of work and very complex transformation rules for data mapping, state-of-the-art semi-automatic data mapping tools such as Karma (Knoblock et al., 2012) can be leveraged as the third option. The concrete output of this step is a unified knowledge model encoding both the (semantic) ontological schema and the epic entities mapped to their semantically corresponding concepts in the schema.

3.5. Formalized Epic Knowledge Resource

When the knowledge model as output (via 3.4.1.3) undergoes ontological review and subsequent to the review, there can be three possibilities: (i) the model passes the satisfactory examination and is rechristened as FEKR (case of direct reuse), or (ii) the model undergoes ontological enrichment and is re-sent for ontological review (case of reuse via enrichment), or, (iii) in the worst case, if the designed model is majorly misaligned with the expressed ontological commitment (which we envision to be rare), the methodological flow loops back to conceptual analysis (see 3.4.2.1; see also Fig. 1). Further, as a crucial aspect of the flexibility of GOMME, the PDT can venture for further iterations of (suitable fragments of) GOMME with respect to developing a more fine-grained representational model of the FEKR, which would not only entail modelling more fine-grained CQs but will also depend on the practical requirements and updates to the project that the PDT is contributing to.

3.6. Application

The output of the GOMME methodology - the FEKR - can be exploited and extended in a variety of application scenarios in primarily, but not only, social sciences and digital humanities. Given the crucial characteristics of architectural agility, repurposing, and extensibility of the GOMME methodology, the FEKR developed for a specific epic can be taken as the starting point for developing an all-encompassing Knowledge Organization Ecosystem (Bagchi, 2021b) for narrative and literary domains (Varadarajan & Dutta, 2021a, 2021b) for intricacies specific to modelling narration which innately maps to GOMME. Further, FEKRs modelled via the GOMME methodology can be exploited in the back-end semantic infrastructure (Bagchi, 2019) of chatbots (Bagchi, 2020) built, for instance, for digital humanities research and popularisation or folk literature.

4. THE MAHABHARATA CASE STUDY - BRIEF HIGHLIGHTS

The principal focus of this paper is to introduce the first dedicated methodology - GOMME - for knowledge modelling for epics and more generally for digital humanities. Additionally, this section presents brief highlights of the project on (classificatory) modelling of the Indian epic *Mahabharata* (Rajagopalachari, 1970) and a first implementation of the methodology. A fuller implementational discussion of the Mahabharata Ontology modelled following GOMME (henceforth referred to as MO-GOMME) remains the subject of a future paper.

We first brief the inceptual phase of the GOMME methodology for modelling the *Mahabharata*, namely

 https://doi.org/10.1633/JISTaP.2023.11.1.5

the steps of initial scoping exercise, selection of the epic and reference text, conceptual familiarization of the epic, and investigation of ontology availability as from Section 3 above. The interdisciplinary PDT comprised four individuals accommodating diverse a priori competencies such as literary expertise, ontology methodology development, involvement in hands-on knowledge organization and representation projects, and expertise in qualitative

Table 2. Sample competency questions

Sl. No.	Competency questions (randomly ordered)	Questions by
1	How many sons did Arjuna have?	AP
2	What was the name of the last son of Ganga?	AP
3	Who was Abhimanyu's teacher?	AP
4	Which Rishi cursed King Pandu?	AP
5	What was the name of Arjuna during Agyatvasa (exile)?	KMP
6	Who was Gandhari?	KMP
7	Of which kingdom was Drupad the king?	KMP
8	Who wrote the *Mahabharata*?	KMP
9	What was the name of Arjuna's Bow?	KPT
10	What was the name of Pandu's second wife?	KPT
11	Who was Shakuni?	KPT
12	What was the name of the Lord Krishna's Conch?	KPT
13	Which Yagna was organised by Yudhishthira?	KPT
14	Who was Keechak?	MP
15	Who killed Ghatotkacha?	MP
16	Who designed the Chakravyuha (circular maze)?	MP
17	Who was the second commander of Kauravas?	MP
18	Who was the first commander of the Kaurava Army?	RA
19	Who was the first commander of the Pandava Army?	RA
20	For how many days was the Mahabharata war fought?	RA
21	Who killed Duryodhana?	RA
22	Who was Kripacharya?	RA
23	In which kingdom did Pandavas spend their Agyatvasa?	RPU
24	What was the name of Krishna's Armour?	RPU
25	Who killed Keechak?	RPU
26	Who was the teacher-trainer of Karna?	RPU
27	Whose name was Sairandhri?	VKM
28	Who was the father of Karna?	VKM
29	Where was the Mahabharata war fought?	VKM
30	Who was the teacher-trainer of Kauravas and Pandavas?	VKM

and quantitative evaluation. The initial scoping exercise, over five dedicated two-hour sessions, produced two guiding design insights. Firstly, the PDT conducted a survey of major existing knowledge modelling methodologies for epic, narrative, and literary works and also examined the concrete conceptual/knowledge models produced (summarised in Section 2). The PDT converged on the highlight that there was no dedicated methodology for modelling epics and GOMME was a proposed candidate solution. Secondly, the PDT decided to adopt GOMME not only as it is targeted to model epics, but also to test the intricacies concerning the key features of architectural agility and feasibility of repurposing and extensibility that it highlights.

Regarding the selection of the epic and the reference text, the PDT, after three dedicated two-hour sessions, naturally converged on the idea to model *Mahabharata* given the pre-eminent standing of the epic in the literary and folk cultural tradition of the entire Indian subcontinent and as the baseline for countless other literary works of repute. The PDT chose the one-volume *Mahabharata* by C. R. Rajagopalachari (1970) in English as the reference text, broadly following the norms listed in step 3.2 of Section 3 above. The PDT discussed the reference text over ten dedicated two-hour sessions for the conceptual familiarization step and set up controlled interactions with three categories of participants (seven participants in total) - novice, aware, and expert - to elicit the kinds of questions they would wish the *Mahabharata* ontology modelled following GOMME to answer. The concrete output was a manually documented list of their questions modelled as CQs after minimal manual pre-processing (only when appropriate) for abiding by scholarly norms (see Table 2). Table 2 enlists a random snapshot of thirty CQs and abbreviations of the survey participants' names, i.e., AP, KMP, KPT, MP, RA, RPU, and VKM. The abbreviations are used for anonymizing the survey practitioners' identities. The same abbreviations are used in Table 3 as well. Some of the questions asked, being descriptive and/or repeated by participants, were filtered out. Finally, the PDT, over two dedicated two-hour sessions, decided to reuse the *Mahabharata* ontology (https://sites.google.com/site/ontoworks/ontologies) developed by the Ontology-based Research Group of Indian Institute of Technology (IIT), Madras (henceforth referred to as MO-IITM), following the recommended grounding of the GOMME methodology in the ontology reuse paradigm.

We now briefly elucidate the ontological review and the satisfactory examination process that the PDT performed with respect to MO-IITM. The step concerning the selection of the reused ontology (step 3.4.1.1) was trivial for the discussed project as MO-IITM was the only existing ontology concerning the epic *Mahabharata*. The ontological review following our proposed ontological norms of MO-IITM was conducted by the PDT over five dedicated two-hour sessions, and the exercise yielded many crucial highlights. We only cite three general observations given the focus of the paper. The first observation is the systematic and pervasive violation of tried-and-tested classificatory principles (Ranganathan, 1967) in many sub-fragments of

Table 3. Satisfactory assessment (via CQs) for MO-IITM

Questions by	Questions	Repetition	Unique questions	Answered	% Answered	Unanswered
MP	67	1	66	31	46.9697	35
KPT	37	5	32	18	56.25	14
AP	23	1	22	8	36.3636	14
RPU	17	3	14	5	35.7143	9
RA	20	1	19	7	36.8421	12
KMP	19	2	17	8	47.0588	9
VKM	20	2	18	10	55.5556	8
Total	203	15	188	87	46.2765	101
Descriptive questions	12	0	12	0	0	12
Considered	191	15	176	87	49.4318	89

CQs, competency questions.

 https://doi.org/10.1633/JISTaP.2023.11.1.5

the ontology at different levels of abstraction, thus rendering the overall axiomatization of the model not only ontologically unwell-founded (Guizzardi et al., 2002) but also classificatorily ungrounded (Giunchiglia & Bagchi, 2021; Satija, 2017) to a certain extent. The second observation is the fact that MO-IITM was not supported by openly available documentation of any kind such as a research paper or webpage. As a result, it is not grounded in the state-of-the-art philosophically inspired and ontologically well-founded best practices such as rigidity (Guarino & Welty, 2002). Finally, as a deviation from the two aforementioned critical observations, the PDT post-ontological review was of a unanimous opinion that MO-IITM could be clearly classified as a classification ontology (Zaihrayeu et al., 2007) given its axiomatization.

After the completion of the ontological review, the PDT, over six dedicated two-hour sessions, actively engaged to determine the satisfiability of MO-IITM and refine it, if and as necessary. The assessment of the satisfiability was a curated process in which the coverage of the ontological commitment targeted for MO-GOMME (modelled via CQs as requirements) was examined vis-à-vis the axiomatization of the ontological commitment formalized in MO-IITM. Table 3 displays the results of the satisfactory assessment. It incorporates the statistical summary of the questions asked by the three categories of survey participants: novices, aware, and experts. From all the participants a total of 203 questions were asked out of which 12 questions were descriptive and 15 were repeated. Therefore, 203-(12+15)=176 unique questions were considered for the ontology appraisal process. Each of the questions was manually checked against MO-IITM. The results show that the ontology could answer 87 questions, i.e. 49.4% of the questions, that was considerably close to the rough threshold of 50% that the PDT assumed. The PDT decided to proceed with the reuse and enrichment of the MO-IITM.

The enrichment, in accordance with GOMME (see 3.4.1.3), was performed via two distinct modes of external enrichment and internal enrichment. External enrichment concentrated on integrating, within MO-IITM, the missing classes (for example, Role), object properties (for example, hasSpouse, hasSibling) and data properties (for example, hasSkills, gender, etc.). This would mitigate the difference between the axiomatization already encoded in the MO-IITM and the ontological commitment (as explicated by the CQs) that it should encode. The internal enrichment was performed to fix modelling bugs within MO-IITM which the PDT found to be pervasive (a detailed description is omitted given the scope of the paper).

During the enrichment process, certain observations were made. Firstly, we cite the case of rigidity from OntoClean (Guarino & Welty, 2002) as also outlined in the ontological review. The MO-IITM (as also most state-of-the-art ontologies) did not ground its axiomatization in ontology best practices, and thus failed to distinguish between roles such as charioteer, which is non-rigid and spatio-temporally boxed (Masolo et al., 2004), from the objects that play such roles in a particular spatio-temporal frame. Secondly, we cite a major classification mistake in the MO-IITM (as from Ranganathan, 1967), exemplified in Giunchiglia and Bagchi (2021). The canon of modulation states, quoting Ranganathan (1967), is "...one class of each and every order that lies between the orders of the first link and the last link of the chain" (p. 176), which enforces the rule that a classification chain "shouldn't have any missing link" (Giunchiglia & Bagchi, 2021). The PDT found missing links to be pervasive, especially in the class and object property hierarchies of the MO-IITM. A case in point is that of the inconsistent modelling of relationships in the MO-IITM, wherein father and mother were classified into the hypernym *parent*, whereas husband and wife remained as free-floating object properties without any hypernym. The PDT refined the inconsistency by modelling a novel relationship matrix (see Fig. 3) following the canon of modulation in spirit. Thirdly and unexpectedly, the PDT also found the MO-IITM to be inconsistent in another major knowledge modelling best practice, that of mandatory domain-range assertion in ob-

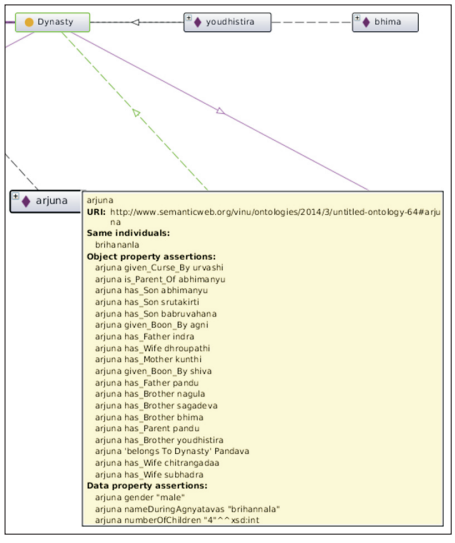


Fig. 3. An example of a relationship matrix.

ject properties, which were subsequently corrected by the PDT as appropriate.

Finally, subsequent to the enrichment process, the PDT conducted a second satisfactory examination of the now enriched MO-IITM, and consequently rechecked the outcome as from Table 3. Table 4 presents the results from the re-checking exercise. The exercise included 203 questions and after elimination of 21 repeated and 14 descriptive questions, 168 questions were considered for the assessment. The PDT found that the enriched MO-IITM could answer 149 questions in total, i.e. around 89% of the CQs, thus exhibiting a major leap from the raw MO-IITM. Based on the t-test of the above sample, the PDT concluded that the curated enrichment is statistically significant. The t-value -5.40 and its corresponding *p*-value of 0.00083 meant that even at a significance level of 0.00083, the curated result is significant. Thus, the PDT decided to rechristen the enriched MO-IITM as the outcome (FEKR) of the case study and renamed it MO-GOMME. Fig. 4 presents a snapshot of MO-GOMME, the developed FEKR. It illustrates a glimpse of the classes, object properties, data properties, and instances of the enriched *Mahabharata* ontology MO-GOMME. As a further note, given the stress of the case study on practically

Table 4. Satisfactory assessment (via CQs) for MO-GOMME

Questions by	Questions	Repetition	Unique questions	Answered	% Answered	Unanswered
MP	65	5	60	48	80	12
KPT	39	3	36	29	80.5556	7
AP	24	2	22	17	77.2727	5
RPU	19	2	17	14	82.3529	3
RA	18	4	14	11	78.5714	3
KMP	17	3	14	13	92.857	1
VKM	21	2	19	17	89.4737	2
Total	203	21	182	149	81.8681	33
Descriptive questions	14	0	14	0	0	14
Considered	189	21	168	149	88.6905	19

CQs, competency questions.

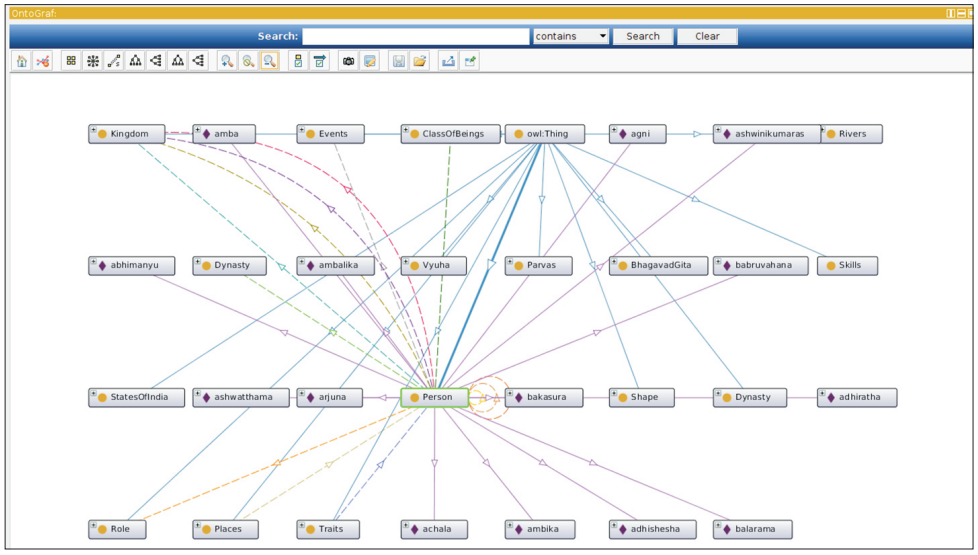


Fig. 4. Snapshot of MO-GOMME.

 https://doi.org/10.1633/JISTaP.2023.11.1.5

implementing GOMME, the PDT did not venture for further iterations of GOMME (which would have obviously required modelling more fine-grained CQs).

5. RESEARCH IMPLICATIONS

The principal research implication of GOMME is that it is the first dedicated methodology for modelling epics in particular, and literary works in general, from the perspective of digital humanities (Berry, 2012; Gold, 2012). All of the existing methodologies for modelling works in digital humanities focus on creating knowledge models such as ontologies and XML schemas from scratch. One of the crucial practical implications of this work is in linking entities in the developed FEKR to the corresponding texts of the epics. For example, we can link entities of the developed FEKR such as characters and places to open knowledge graphs such as DBpedia (Auer et al., 2007) or Wikidata (Vrandečić & Krötzsch, 2014), based on their dereferenceable availability. A case in point is the *Mahabharata* character Arjuna, who can be unambiguously linked to the Wikidata ID Q27131348. Subsequently, we can link such entities in the FEKR with the relevant text/documents/papers that mention the entities. Similar work was done for Book I of the *Odyssey* (Khan et al., 2016). The current work also contributes to information retrieval on epics of any scale, region, or religion by exploiting the concepts of the developed FEKR as elements of a base metadata scheme for epics. For example, the class Event is a metadata element that describes events which happen in an epic. Object properties such as main_Character_In is a metadata element that describes the characters involved in the event. This, we envision, is an initiation towards creating a dedicated metadata schema for annotating epics.

The methodology addresses the likes of researchers in diverse research arenas such as computer science, theology, library and information science, and literary studies. The knowledge model such as FEKR, in the backend of the question answering system and the metadata for the epics, will help in querying and retrieving the information concerning their domain. A line of research in computer science can be to develop a semi-automatic system that helps in classifying the entities of an epic and generating the FEKR. Semi-automation is not only a limitation of the current work but also extremely difficult for epics due to the lack of dedicated state-of-the-art frameworks in the corresponding arena of research in multilingual natural language processing.

6. CONCLUSION AND FUTURE WORK

The current work contributes towards a novel methodology by not only allowing the PDT to model epics/literary works from scratch but also by factoring in the possibility of reusing via enrichment of existing conceptual or ontological formalizations of epics. Apart from the mere fact that GOMME adds to the basket of existing methodologies, it is also, as far as our know-how, the only theoretically enhanced methodology grounded in transdisciplinarity. This is due to the crucial grounding of GOMME in the state-of-the-art norms of epics, ontological well-founded best practices, and tried-and-tested classificatory principles. It is also the first methodology to uniquely accommodate the flexible combination of CQs and GQs to model ontological commitments. It illustrates the feasibility and advantages of the methodology via a first brief implementation of modelling the Indian epic *Mahabharata*. Although the current version of GOMME is focused on modelling epics, in future its potential extensibility for modelling works in a wide variety of other related disciplines such as different arenas of digital humanities, literature, religion, and social sciences will be explored. We envision three immediate future streams of research which we will pursue in upcoming dedicated research papers: (i) the development of a novel theory of knowledge representation grounded in the design foundations which are implicit in GOMME, (ii) modelling different works in different arenas of digital humanities via GOMME, and (iii) potentially creating a dedicated methodology suite for modelling literary knowledge in different arenas of digital humanities and social sciences. In addition, real-world oriented solutions, such as question-answering systems dedicated to epics grounded in (semantic) data management plans (Gajbe et al., 2021), entity disambiguation, and linking in the wild for digital humanities (Frontini et al., 2015; Plirdpring & Ruangrajitpakorn, 2022), are much valued in the domain of digital humanities. GOMME, via its flexible knowledge management infrastructure, can partially (as of now) facilitate their realization.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the late Mr. Keshav Prasad Tiwari and Mr. Mahendra Pandey for their efforts in CQs formulation. Additionally, we wish to thank the editor, copy-editor, and anonymous reviewers for their constructive comments and suggestions.

CONFLICTS OF INTEREST

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

REFERENCES

Abello, J., Broadwell, P., & Tangherlini, T. R. (2012). Computational folkloristics. *Communications of the ACM,* 55(7), 60-70. https://doi.org/10.1145/2209249.2209267

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., & Ives, Z. (2007, November 11-15). DBpedia: A nucleus for a web of open data. In K. Aberer, K. S. Choi, N. Noy, D. Allemang, K. I. Lee, L. Nixon, J. Golbeck, P. Mika, D. Maynard, R. Mizoguchi, G. Schreiber, & P. Cudré-Mauroux (Eds.), *The Semantic Web: 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference* (pp. 722-735). Springer.

Bagchi, M. (2019). *A knowledge architecture using knowledge graphs.* (Master's thesis). Indian Statistical Institute, Bangalore, India.

Bagchi, M. (2020). Conceptualising a library chatbot using Open source conversational artificial intelligence. *DESIDOC Journal of Library & Information Technology,* 40(6), 329-333. https://doi.org/10.14429/djlit.40.6.15611

Bagchi, M. (2021a). A large scale, knowledge intensive domain development methodology. *Knowledge Organization,* 48(1), 8-23. https://doi.org/10.5771/0943-7444-2021-1-8

Bagchi, M. (2021b). Towards knowledge organization ecosystem (KOE). *Cataloging & Classification Quarterly,* 59(8), 740-756. https://doi.org/10.1080/01639374.2021.1998282

Bagchi, M., & Madalli, D. P. (2019, February 26-28). Domain visualization using knowledge cartography in the big data era: A knowledge graph based alternative. In K. R. Patnaik, & K. Satpathi (Eds.), *International Conference on "Future of Libraries"* (pp. 1-22). Indian Institute of Management.

Berry, D. M. (2012). Introduction: Understanding the digital humanities. In D. M. Berry (Ed.), *Understanding digital humanities* (pp. 1-20). Palgrave Macmillan.

Branch, F., Arias, T., Kennah, J., Phillips, R., Windleharth, T., & Lee, J. H. (2017). Representing transmedia fictional worlds through ontology. *Journal of the Association for Information Science and Technology,* 68(12), 2771-2782. https://doi.org/10.1002/asi.23886

Carriero, V. A., Daquino, M., Gangemi, A., Nuzzolese, A. G., Peroni, S., Presutti, V., & Tomasi, F. (2020). The landscape of ontology reuse approaches. In G. Cota, M. Daquino, & G. L. Pozzato (Eds.), *Applications and practices in ontology design, extraction, and reasoning* (pp. 21-38). IOS Press.

Cooren, F. (2015). In medias res: Communication, existence, and materiality. *Communication Research and Practice,* 1(4), 307-321. https://doi.org/10.1080/22041451.2015.1110075

Damiano, R., & Lieto, A. (2013). Ontological representations of narratives: A case study on stories and actions. In M. A. Finlayson, B. Fisseni, B. Löwe, & J. C. Meister (Eds.), *2013 Workshop on Computational Models of Narrative* (pp. 76-93). Schloss Dagstuhl.

Damiano, R., Lombardo, V., & Pizzo, A. (2019). The ontology of drama. *Applied Ontology,* 14(1), 79-118. https://doi.org/10.3233/AO-190204

Das, S., & Roy, S. (2016). Faceted ontological model for brain tumour study. *Knowledge Organization,* 43(1), 3-12. https://doi.org/10.5771/0943-7444-2016-1-3

DuCharme, B. (2013). *Learning SPARQL: Querying and updating with SPARQL 1.1.* O'Reilly.

Dutta, B., Chatterjee, U., & Madalli, D. P. (2015). YAMO: Yet another methodology for large-scale faceted ontology construction. *Journal of Knowledge Management,* 19(1), 6-24. https://doi.org/10.1108/JKM-10-2014-0439

Ehrlinger, L., & Wöß, W. (2016, September 12-15). Towards a definition of knowledge graphs. In M. Martin, M. Cuquet, & E. Folmer (Eds.), *SEMANTICS Leipzig 2016.* Committee of SEMANTICS 2016.

Espinoza-Arias, P., Poveda-Villalón, M., García-Castro, R., & Corcho, O. (2019). Ontological representation of smart city data: From devices to cities. *Applied Sciences,* 9(1), 32. https://doi.org/10.3390/app9010032

Fernández-López, M., Gómez-Pérez, A., & Juristo, N. (1997, March 24-26). Methontology: From ontological art towards ontological engineering. *Proceedings of the Ontological Engineering AAAI-97 Spring Symposium series* (pp. 33-40). AAAI.

Frontini, F., Brando, C., & Ganascia, J. G. (2015, June 1). Semantic web based named entity linking for digital humanities and heritage texts. In A. Zucker, I. Draelants, C. F. Zucker, & A. Monnin (Eds.), *Proceedings of the First International Workshop Semantic Web for Scientific Heritage at the 12th ESWC 2015 Conference* (pp. 77-88). HAL Open Science.

Gajbe, S. B., Tiwari, A., Gopalji., & Singh, R. K. (2021). Evaluation and analysis of data management plan tools: A parametric approach. *Information Processing & Management,* 58(3), 102480. https://doi.org/10.1016/j.ipm.2020.102480

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., & Schneider, L. (2002, October 1-4). Sweetening ontologies with DOLCE. In A. Gómez-Pérez, & V. R. Benjamins (Eds.), *EKAW 2002: Knowledge engineering and knowledge management: ontologies and the semantic web* (pp. 166-181).

Springer.

Giunchiglia, F., & Bagchi, M. (2021, September 11-18). Millikan + Ranganathan – From perception to classification. In E. M. Sanfilippo, O. Kutz, N. Troquard, T. Hahmann, C. Masolo, R. Hoehndorf, & R. Vita (Eds.), *CAOS 2021: 5th Workshop on Cognition And OntologieS*. Springer.

Giunchiglia, F., Bocca, S., Fumagalli, M., Bagchi, M., & Zamboni, A. (2021a). iTelos -- purpose driven knowledge graph generation. *arXiv,* arXiv:2105.09418. https://doi.org/10.48550/arXiv.2105.09418

Giunchiglia, F., Dutta, B., & Maltese, V. (2014). From knowledge organization to knowledge representation. *Knowledge Organization,* 41(1), 44-56. https://doi.org/10.5771/0943-7444-2014-1-44

Giunchiglia, F., Zamboni, A., Bagchi, M., & Bocca, S. (2021b, June 6). Stratified data integration. In D. Chaves-Fraga, A. Dimou, P. Heyvaert, F. Priyatna, & J. Sequeda (Eds.), *Knowledge Graph Construction 2021*. CEUR.

Gold, M. K. (2012). *Debates in the digital humanities.* University of Minnesota Press.

Gómez-Pérez, A. (2004). Ontology evaluation. In S. Staab, & R. Studer (Eds.), *Handbook on ontologies* (pp. 251-273). Springer.

Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2004). *Ontological Engineering: With examples from the areas of knowledge management, e-commerce and the semantic web*. Springer.

Greene, T. (1961). The norms of epic. *Comparative Literature,* 13(3), 193-207. https://doi.org/10.2307/1768997

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies,* 43(5-6), 907-928. https://doi.org/10.1006/ijhc.1995.1081

Grüninger, M., & Fox, M. S. (1995a). *Methodology for the design and evaluation of ontologies*. Paper presented at IJCAI'95, Montreal, Canada.

Grüninger, M., & Fox, M. S. (1995b). The role of competency questions in enterprise engineering. In A. Rolstadås (Ed.), *Benchmarking — Theory and practice* (pp. 22-31). Springer.

Guarino, N., & Welty, C. (2002). Evaluating ontological decisions with Ontoclean. *Communications of the ACM,* 45(2), 61-65. https://doi.org/10.1145/503124.503150

Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology? In S. Staab, & R. Studer (Eds.), *Handbook on ontologies* (pp. 1-17). Springer.

Guizzardi, G., Herre, H., & Wagner, G. (2002, October 7-11). On the general ontological foundations of conceptual modeling. In S. Spaccapietra, S. T. March, & Y. Kambayashi (Eds.), *Conceptual modeling — ER 2002* (pp. 65-78). Springer.

Hyvönen, E., Valo, A., Komulainen, V., Seppälä, K., Kauppinen, T., Ruotsalo, T., Salminen, M., & Ylisalmi, A. (2005). Finnish national ontologies for the semantic web -- Towards a content and service infrastructure. *International Conference on Dublin Core and Metadata Applications* (pp. 219-222). Dublin Core.

Iqbal, R., Mustapha, A., & Yusoff, Z. M. (2013). An experience of developing Quran ontology with contextual information support. *Multicultural Education & Technology Journal,* 7(4), 333-343. https://doi.org/10.1108/METJ-03-2013-0009

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Harvard University Press.

Kent, R. E. (2000, July 10-13). The information flow foundation for conceptual knowledge organization. In C. Beghtol, L. Howarth, & N. J. Williamson (Eds.), *Dynamism and stability in knowledge organization: Proceedings of the Sixth International ISKO Conference* (pp. 111-117). International Society for Knowledge Organization.

Khan, A. F., Bellandi, A., Benotto, G., Frontini, F., Giovannetti, E., & Reboul, M. (2016, September 18-20). Leveraging a narrative ontology to query a literary text. In B. Miller, A. Lieto, R. Ronfard, S. G. Ware, & M. A. Finlayson (Eds.), *The 7th Workshop on Computational Models of Narrative (CMN 2016)* (pp. 10:1-10:10). Schloss Dagstuhl.

Klarer, M. (2013). *An introduction to literary studies*. (3rd ed.). Routledge.

Knoblock, C. A., Szekely, P., Ambite, J. L., Goel, A., Gupta, S., Lerman, K., Muslea, M., Taheriyan, M., & Mallick, P. (2012, May 27-31). Semi-automatically mapping structured sources into the semantic web. In E. Simperl, P. Cimiano, A. Polleres, O. Corcho, & V. Presutti (Eds.), *ESWC 2012: The semantic web: research and applications* (pp. 375-390). Springer.

Lassila, O., & McGuinness, D. (2001). *The role of frame-based representation on the semantic web*. Linköping University Electronic Press.

Lehmann, F. (1992). Semantic networks. *Computers & Mathematics with Applications,* 23(2-5), 1-50. https://doi.org/10.1016/0898-1221(92)90135-5

Lochtefeld, J. G. (2002). *The illustrated encyclopedia of hinduism*. Rosen Publishing.

Mahesh, K., Helmreich, S., & Wilson, L. (1996). *Ontology development for machine translation: Ideology and methodology*. New Mexico State University.

Masolo, C., Vieu, L., Bottazzi, E., Catenacci, C., Ferrario, R., Gangemi, A., & Guarino, N. (2004, June 2-5). Social roles

and their descriptions. In D. Dubois, C. Welty, & M. A. Williams (Eds.), *Proceedings of the 9th International Conference on the Principles of Knowledge Representation and Reasoning* (pp. 267-277). AAAI Press.

Mulholland, P., Collins, T., & Zdrahal, Z. (2004, January 13-16). Story fountain: Intelligent support for story research and exploration. In J. Vanderdonckt (Ed.), *Proceedings of the 9th International Conference on Intelligent User Interfaces* (pp. 62-69). ACM.

Nakasone, A., & Ishizuka, M. (2006, December 18-22). Storytelling ontology model using RST. In T. Nishida (Ed.), *2006 IEEE/WIC/ACM International Conference on Intelligent Agent Technology* (pp. 163-169). IEEE.

Noy, N. F., & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*. Stanford University.

Parsons, V. (2014). Stratified sampling. In N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri, & J. Teugels (Eds.), *Wiley StatsRef: Statistics reference online.* John Wiley & Sons.

Phoksawat, K., Mahmuddin, M., & Ta'a, A. (2019). Intercropping in rubber plantation ontology for a decision support system. *Journal of Information Science Theory and Practice,* 7(4), 56-64. https://doi.org/10.1633/JISTaP.2019.7.4.5

Pinto, H. S., & Martins, J. P. (2000, March 20-22). Reusing ontologies. *AAAI 2000 Spring Symposium on Bringing Knowledge to Business Processes* (pp. 77-84). AAAI.

Plirdpring, P., & Ruangrajitpakorn, T. (2022). Using ontology to represent cultural aspects of local products for supporting local community enterprise in Thailand. *Journal of Information Science Theory and Practice,* 10(1), 45-58. https://doi.org/10.1633/JISTaP.2022.10.1.4

Rajagopalachari, C. (1970). *Mahabharata.* Diamond Pocket Books.

Ranganathan, S. R. (1967). *Prolegomena to library classification.* (3rd ed.). Ess Ess Publications.

Satija, M. P. (2017). Colon Classification (CC). *Knowledge Organization,* 44(4), 291-307. https://doi.org/10.5771/0943-7444-2017-4-291

Satija, M. P., Bagchi, M., & Martínez-Ávila, D. (2020). Metadata management and application. *Library Herald,* 58(4), 84-107. https://doi.org/10.5958/0976-2469.2020.00030.2

Simperl, E. (2009). Reusing ontologies on the semantic web: A feasibility study. *Data & Knowledge Engineering,* 68(10), 905-925. https://doi.org/10.1016/j.datak.2009.02.002

Suárez-Figueroa, M. C., Gómez-Pérez, A., & Fernández-López, M. (2015). The NeOn methodology framework: A scenario-based methodology for ontology development. *Applied Ontology,* 10(2), 107-145. https://doi.org/10.3233/AO-150145

Swartjes, I., & Theune, M. (2006). A fabula model for emergent narrative. In S. Göbel, R. Malkewitz, & I. Iurgel (Eds.), *Technologies for interactive digital storytelling and entertainment* (pp. 49-60). Springer.

Syamili, C., & Rekha, R. V. (2018). Developing an ontology for Greek mythology. *Electronic Library,* 36(1), 119-132. https://doi.org/10.1108/EL-02-2017-0030

Tiwari, A., & Madalli, D. P. (2021). Maturity models in LIS study and practice. *Library & Information Science Research,* 43(1), 101069. https://doi.org/10.1016/j.lisr.2020.101069

Uschold, M., & King, M. (1995). *Towards a methodology for building ontologies*. University of Edinburgh.

Varadarajan, U., & Dutta, B. (2021a, September 18-20). Towards development of knowledge graph for narrative information in medicine. In B. Villazón-Terrazas, F. Ortiz-Rodríguez, S. Tiwari, A. Goyal, & M. Jabbar (Eds.), *KGSWC 2021: Knowledge graphs and semantic web* (pp. 290-307). Springer.

Varadarajan, U., & Dutta, B. (2021b). Models for narrative information: A study. *arXiv,* arXiv:2110.02084. https://doi.org/10.48550/arXiv.2110.02084

Vosniadou, S. (2002). Mental models in conceptual development. In L. Magnani, & N. J. Nersessian (Eds.), *Model-based reasoning: Science, technology, values* (pp. 353-368). Springer.

Vrandečić, D., & Krötzsch, M. (2014). Wikidata: A free collaborative knowledge base. *Communications of the ACM,* 57(10), 78-85. https://doi.org/10.1145/2629489

Vrandečić, D., Pinto, S., Tempich, C., & Sure, Y. (2005). The DILIGENT knowledge processes. *Journal of Knowledge Management,* 9(5), 85-96. https://doi.org/10.1108/13673270510622474

Yeh, J. H. (2017, May 13-17). StoryTeller: An event-based story ontology composition system for biographical history. In T. H. Meen, S. D. Prior, & A. D. K. T. Lam (Eds.), *Proceeding of the 2017 International Conference on Applied System Innovation (ICASI)* (pp. 1934-1937). IEEE.

Zaihrayeu, I., Sun, L., Giunchiglia, F., Pan, W., Ju, Q., Chi, M., & Huang, X. (2007, November 11-15). From web directories to ontologies: Natural language processing challenges. In K. Aberer, K. S. Choi, N. Noy, D. Allemang, K. I. Lee, L. Nixon, J. Golbeck, P. Mika, D. Maynard, R. Mizoguchi, G. Schreiber, & P. Cudré-Mauroux (Eds.), *The semantic web: 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference* (pp. 623-636). Springer.

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

RESEARCH PAPER
J Inf Sci Theory Pract 11(1): 79-88, 2023

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2023.11.1.6

Development of Ontology for Thai Country Songs

Thunyaluk Jaitiang
Department of Information Science, Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand
E-mail: thanya02523@gmail.com

Wirapong Chansanam*
Department of Information Science, Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand
E-mail: wirach@kku.ac.th

Malee Kabmala
Department of Information Science, Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand
E-mail: malee_ka@kku.ac.th

ABSTRACT

This study aimed to develop an ontology for Thai country songs by using the seven steps of an ontology development process. Hozo-Ontology Editor software and Ontology Application Management Framework were tools used in this study. Nine classes of ontology were identified: song, singer, emotion, author, language used, language type, song style, original, and content, and it was found that the song class had a relationship with all of the other classes. The developed ontology was evaluated by seeking opinions from experts in the field of Thai country songs, who agreed that the ontology was highly effective. Additionally, the evaluation employed the knowledge retrieval concept, and the precision, recall, and overall effectiveness were measured, with a precision of 92.59%, a recall of 86.21%, and an overall effectiveness (F-measure) of 89.28%. These results indicate that the developed ontology is highly effective in describing the scope of knowledge of Thai country songs.

Keywords: ontology, Thai country songs, semantic web, semantic search, digital humanities

Received: December 8, 2022 **Revised:** January 15, 2023
Accepted: January 31, 2023 **Published:** March 30, 2023

***Corresponding Author:** Wirapong Chansanam
https://orcid.org/0000-0001-5546-8485
E-mail: wirach@kku.ac.th

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far 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.

© Thunyaluk Jaitiang, Malee Kabmala, Wirapong Chansanam, 2023

1. INTRODUCTION

The Thai country song tradition or *Phleng Luk thung* is considered a Thai cultural heritage information resource. The music contains key messages regarding the evolution of the country and wisdom in each era (Sirisamphan & Virulrak, 2010). Thai country songs are an intangible cultural heritage retelling various stories. The music portrays knowledge of culture, community, values, and narratives. Moreover, Thai country songs play an important role in preserving culture (UNESCO, 2016) and also vividly reflect particular ways of life and traditions of people through simple language. The music delivers messages and moods such as love, sorrow, joy, pride, courage, or happiness to listeners directly and impactfully. Therefore, the Thai country song tradition is well known to everyone (Khuajaroen, 2014; Srinok, 2010; Thanasilangkun, 2001).

Rapid technological change makes it more accessible for people to listen to music through digital formats (Kongsomsawaeng et al., 2021); for instance, people often listen to music through streaming services such as Spotify, Deezer, JOOX, LINE MUSIC, Apple Music, and YouTube Music (Viklin, 2023). This has led to a major rise in popularity of Thai country music, created with various melodies, lyrics, and styles. The key messages of the music have been illustrated in terms of society, culture, religion, traditions, rituals, ways of life, love, loneliness, work life, localization, and courage (Chaichana, 2016; Kentorrapak, 2021; Powell, 2022). Recently, people tend to listen to Thai country songs increasingly through digital platforms: Facebook, YouTube, and websites. They often search for music from keywords, for instance, title, singer, songwriter, and some parts of lyrics; yet this might lead to inaccurate and insufficient retrieval for people to get the right songs they search for, since the information is not interconnected. As a result, people might need to search on many platforms, and it might take time to find a single song to listen to; therefore, a Semantic Web or information networking system could assist users in getting information faster and more conveniently. Therefore, digital preservation is a research trend in the digital humanities domain (Chansanam et al., 2022).

The purpose of this study was to work on developing an ontology for Thai country songs. In this study, a synthetic separation method was used to analyze and classify the content of modern Thai country songs. The knowledge structure needed for this study was collected to understand contemporary Thai country songs better. Ontology is an alternative knowledge-based platform that shows a concept's scope and categorizes scattering data into an ontology (Antoniou & van Harmelen, 2004). The researchers acknowledged the insufficiency regarding Thai country song retrieval; consequently, an ontology is needed to provide precise keywords, semantics, properties, characteristics, and relationships of all related Thai country song concepts. Therefore, ontology development is related to the semantic tools that facilitate users' searches for needed information. Furthermore, the researchers employ the Hozo-Ontology Editor program to develop the Thai country song ontology, and the Ontology Application Management (OAM) Framework to develop a Semantic Web to store the developed ontology and allow users to access all Thai country song information.

2. LITERATURE REVIEW

There has been a trend of listening to Thai country music and mixing Thai country songs with new rhythmic music. With the advancement of technology and the digital age, Thais are becoming more interested in listening to country music from the old cassette tape era and moving to the CD or DVD era (Kongsomsawaeng et al., 2021). Music can be listened to more easily from online media, especially the YouTube website, resulting in a new type of Thai country song called contemporary Thai country songs, which are published during the present period and can reflect the way of life, ideological conditions, and Thai culture through their current content. A combination of country, local, folk, and life songs have emerged (Kentorrapak, 2021). As a result, Thai country songs have a variety of contents, pique the interest of scholars from various fields, and lead to different views on Thai country songs as follows: the use of language, overlapping words, double words in foreign languages, and the number of rhetorical images used by the content of Thai country songs for the benefit of dictionaries, vocabulary lists, and taxonomies (Chutapanish, 2004).

Scholars in the mass communication field have looked at Thai country music; their content is about creating artists, production processes, singers, and communication strategies, by applying the contents of Thai country songs for the benefit of information dissemination and public relations for better practical results (Tohombut, 2015). Thai country songs are believed to be an intermediary for conveying political concepts of the time, social criticisms, parody, and satire. The contents of Thai country songs have been used for the purpose of communicating government policies (Fyeted, 2005). Thai studies scholars

 https://doi.org/10.1633/JISTaP.2023.11.1.6

have thought that Thai country songs have different kinds of content, such as social history, language, and culture, and that the songs' content helps people understand Thai society and culture and contributes to Thai society's creation (Yaikoom, 2004). Therefore, Thai country songs can reflect beliefs about merit, sin, karma, hell, heaven, and doctrines. Thai country songs can serve as a guideline for living a good life by adhering to Buddhist teachings and cultivating desirable morals and ethics for youth in educational institutions (Srinok, 2010). In information science, it is clear that Thai country songs are a type of information resource that contains a variety of content about life, culture, and society, and *Luk thung* is a medium used to convey meaning to listeners so that they understand the contents that appear (Kabmala, 2020). Thai country songs, according to historians, are a record of the society's history at the time. Scholars use their knowledge of Thai country songs to study history based on historical events (Krobthong, 1998).

Knowledge organization serves as a foundation of knowledge structure development and ontology development, which contributes to effective information searching and access to knowledge. Apart from studies for cultural knowledge organization, there exists an approach to cultural information management, referred to as digital humanities research. Ontology provides a fundamental framework for the development of the Semantic Web. A knowledge organization system is the key element of knowledge engineering (Bouihi & Bahaj, 2019; Broughton et al., 2005; Tuamsuk et al., 2018a). Knowledge organization, ontology development, and semantic search development are interrelated approaches. In the development of semantic search, its basic data is derived from the data structure design with ontology. Still, ontology development requires analyzing and synthesizing data and knowledge about a particular topic, as well as determining the concepts and relations of knowledge systematically (Noy & McGuinness, 2001).

An ontology is a commonly used tool to explain the representation of knowledge. The advantages of an ontology include its interoperability to share a common understanding among people or software agents, enabling the reuse of domain knowledge and facilitating explicit conventions (Chansanam et al., 2021). We have adapted some of the data from previous ontologies to fit the scope of this study. Ontologies can be shared between applications or reused because an ontology is the keyword source used to describe domain concepts. The ontology can copy the knowledge base and use it in new applications. There may be an improvement in the structure or the addition of more knowledge (Gašević et al., 2009; Tuamsuk et al., 2018b).

Here contemporary music content is analyzed and classified into various aspects of Thai country music content, so that all fields of study can be used appropriately. From the academic point of view, scholars in each field of study use contemporary Thai country song content differently according to their intended use, and the importance of Thai country song content varies in terms of subject areas. Therefore, we studied the content analysis and classification of contemporary Thai country songs using ontology techniques.

3. METHODS

3.1. Ontology Development Process

The study employed seven steps of ontology development from Noy and McGuinness (2001) as follows:

3.1.1. Determine Scope

This study determined the scope of song, singer, emotion, author, language used, language type, song style, original, and content.

3.1.2. Consider Reuse

This study employed concepts in ontology development to create a prototype of a semantic website containing Thai country song information since the music has been launched on websites.

3.1.3. Enumerate Terms and Properties

The study listed the terms and properties of each class as in Table 1.

Table 1. List of terms and properties in a song's class

Attribute	Kind	Class constraint	Label
song_id	a/o	integer	song ID
song_name	a/o	string	song title
singer	p/o	string	singer
mood	p/o	string	mood
author	p/o	string	songwriter
language_style	p/o	string	language style
content	p/o	string	message

a/o, attribute of; p/o, part of.

3.1.4. Define Class

There are nine classes, namely song, singer, emotion, author, language used (word choices, stylistic devices, idioms), language type (pronouns, time, settings, dialects, connotation, foreign language, compound, figurative language, reduplication, colloquial, trendy words, emotion words, synonym, rhetorical questions, paradox, symbol, onomatope, hyperbole, synesthesia, analogy, simile, metaphor, modern idioms, old idioms, metonymy, personification, allusion), song style (pop, dance-pop, *Phleng phuea chiwit* or Thai country song), originality (contests, recommendations, digital platforms, own debuts), and content (society, beliefs, nature). The details of the nine classes are presented for each class in Table 2.

3.1.5. Define Properties of Classes

- Concept relationship: "is-a," e.g., family is-a social community.
- Property relationship: "part-of" or "p/o," e.g., mood is part-of a song.
- Attribute relationship: "attribute-of" or "a/o," e.g., song ID and song title are "attribute-of" a song.

3.1.6. Define Constraints

Constraints add conditions in classes; e.g., in a class of language style, the constraints were addressed according to word choices, stylistic devices, and idioms only.

3.1.7. Create Instances

Instances address examples; e.g., in a class of singer, if a singer's name is "Tai Orathai," the instance would be singer: Tai Orathai.

3.2. Application of the Thai Country Song Ontology to a Semantic Web

According to the ontology development process, the researchers gathered information regarding Thai country songs from related resources: 1) the Office of Information Technology Administration for Educational Development (http://tdc.thailis.or.th/tdc) (Office of Information Technology Administration for Educational Development: UniNet, 2000); 2) Thai Journals Online (https://www.tci-thaijo.org) (Thailand Research Fund, King Mongkut's University of Technology Thonburi, Thammasat University, National Electronics and Computer Technology Center, Thai-Journal Citation Index Centre, 2001); and 3) Thai country songs released between 2008 and 2020. This information was gathered as part of the Semantic Web development. The researchers used the OAM Framework as a primary program to apply all data from the ontology since this program could take RDF and OWL from publishing and the developed databases, as shown in Fig. 1 (Buranarach, 2017). The researchers only mapped existing data with the developed ontology by importing the .owl file from the ontology and data from MySQL to the ap-

Table 2. List of main classes and definitions

Class/Concept	Definition
Song	Thai country songs published today are also known as contemporary Thai country songs where those composers, and the singer, are alive. The content reflects the way of life, culture, society, living, family, politics and government, economy, occupation, education, values, beliefs, religion, language, sexuality, and the beauty of nature with melodies, vocals, and accents, using simple, straightforward language and playing with patterns and uniqueness in each locality and region in Thailand.
Singer	A person who sells sound and sound services and sings professionally in Thai country songs. There are many types of singers according to music genres. A good singer has a beautiful voice that attracts listeners to enjoy the music.
Emotion	The emotion that needs to be communicated to the listener to recognize or feel the content of the Thai country songs.
Author	The composer of lyrics and melody or composer in Thai country songs.
Language used	The use of language in the songwriting of Thai country songs.
Language type	The type of language used in the songwriting of Thai country songs.
Song style	A genre of songs.
Originality	The beginning before becoming a Thai country song singer.
Content	Music content that reflects the way of life, social conditions, ideals, and Thai culture.

 https://doi.org/10.1633/JISTaP.2023.11.1.6

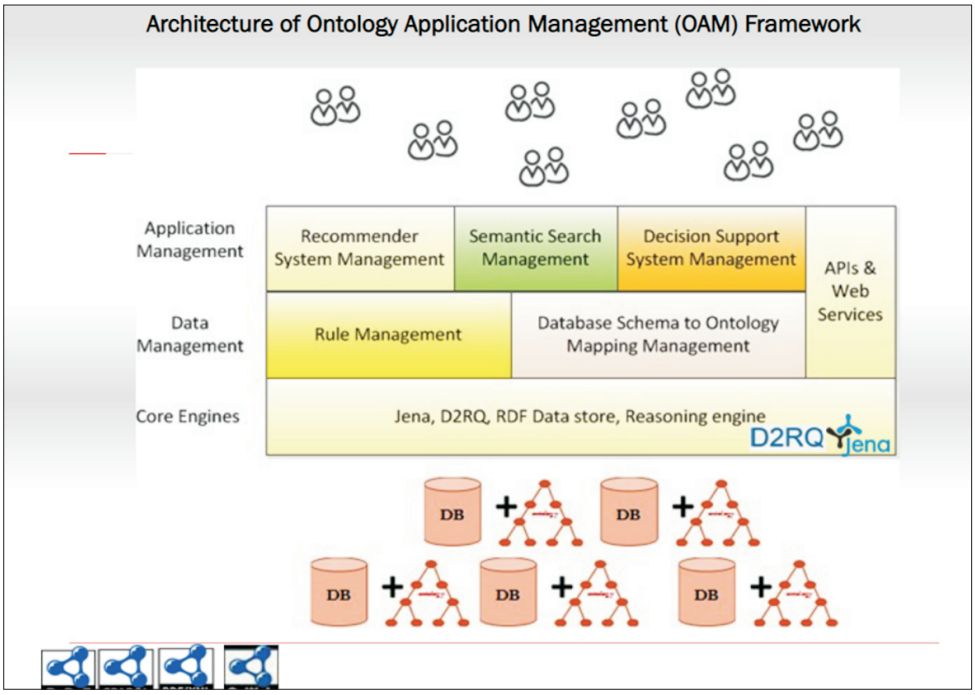


Fig. 1. Development of a Semantic Web using Ontology Application Management (OAM) Framework. *Source*: Buranarach, 2017.

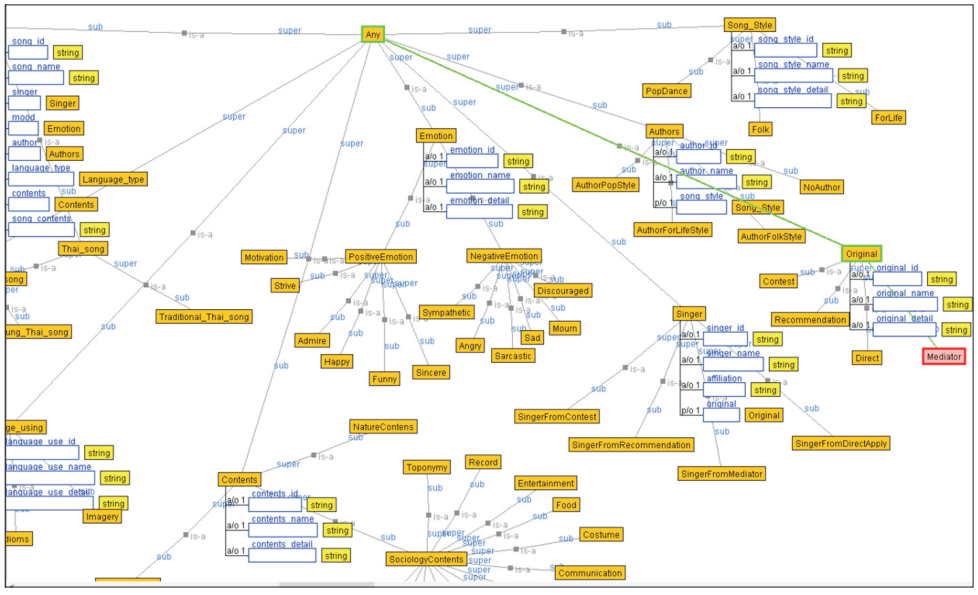


Fig. 2. The ontology of Thai country songs using HoZo-Ontology Editor.

plication template; moreover, the researchers coordinated with APIs & Web Service to make the Semantic Web compatible with other search engines (Buranarach et al., 2013) to reduce the complexity of program development.

Regarding keyword entry and results, the researchers developed a web application with Java where users can search and get all the information through Web API. The web application will send the information to the Semantic Web of the OAM Framework and show the search results via the web application.

4. FINDINGS

4.1. Ontology Development

The data on Thai country songs were collected from three sources: ThaiLIS, ThaiJO, and all songs released between 2008 and 2020. The researchers analyzed and developed an ontology consisting of nine classes: song, singer,

emotion, author, language used, language type, song style, originality, and content. Then, the knowledge structure in Fig. 2 was developed using HoZo-Ontology Editor (Kozaki et al., 2002).

4.2. Semantic Web Development

After the ontology development, the data was converted to OWL format, as in Fig. 3.

The researchers created a MySQL database to store the data. Then, the OAM Framework was used to link the data from MySQL with the data from the ontology, and the data from two sources were converted to OWL (Fig. 3) for a Database-Ontology Mapping process, as in Fig. 4. Fig. 5 shows relationship configurations from those two sources. The relationship types that OAM Framework can map are one-to-one, one-to-many, and many-to-many. These relationships are for linking with subclasses where vocabulary mapping is processed. Fig. 5 shows relationships between columns and subclasses.

The user interface is in the form of a webpage written in Java and JavaScript. First, users can use keyword search and advanced search functions, and then the search will reach semantic search via Web API & Web Service, as in Fig. 6.

To use the developed Semantic Web, users can use it through a web browser by entering this URL: http://localhost:8080/searching. There are two semantic search forms, as follows:

a) Keyword search. Users can search according to classes or concepts, such as songwriter, mood, language

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
        xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
        xmlns:owl="http://www.w3.org/2002/07/owl#"
        xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
        xmlns="http://www.hozo.jp/owl/Song.owl#"
        xml:base="http://www.hozo.jp/owl/Song.owl#">

<owl:Ontology rdf:about="">
    <rdfs:comment>
        HOZO:OWL Export
    </rdfs:comment>
</owl:Ontology>

<owl:Class rdf:ID="RelationalConcept">
    <rdfs:label>RelationalConcept</rdfs:label>
</owl:Class>
<owl:ObjectProperty rdf:ID="hasPart">
    <rdfs:label>hasPart</rdfs:label>
</owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="hasAttribute">
    <rdfs:label>hasAttribute</rdfs:label>
</owl:ObjectProperty>
<owl:Class rdf:ID="Any">
    <rdfs:label>Any</rdfs:label>
</owl:Class>
<owl:Class rdf:ID="Song">
    <rdfs:label>Song</rdfs:label>
    <rdfs:subClassOf rdf:resource="#Any" />
    <rdfs:subClassOf>
        <owl:Restriction>
            <owl:cardinality
rdf:datatype="http://www.w3.org/2001/XMLSchema#nonNegativeInteger
">1</owl:cardinality>
            <owl:onProperty rdf:resource="#has_song_id" />
        </owl:Restriction>
    </rdfs:subClassOf>
```

Fig. 3. Ontology data in an OWL format.

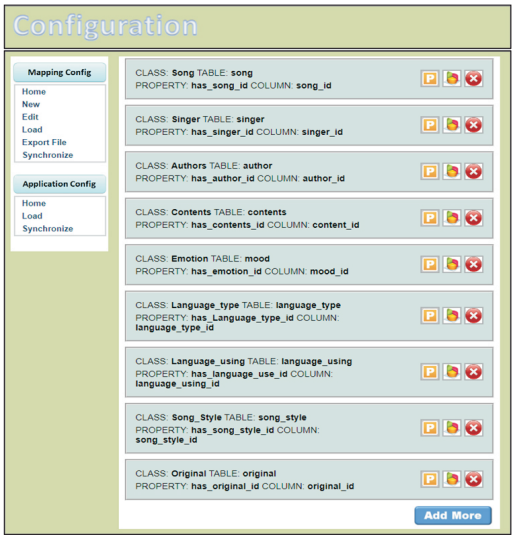


Fig. 4. Class relationship settings.

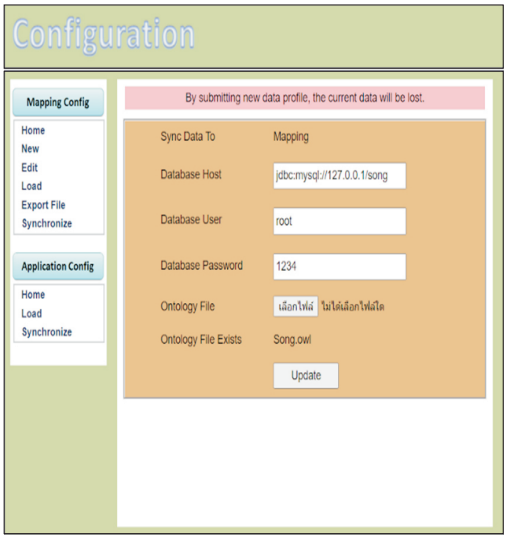


Fig. 5. Columns and relationships in subclasses of ontology and data from the ontology of MySQL.



Fig. 6. API setting for data search in the ontology database.

 https://doi.org/10.1633/JISTaP.2023.11.1.6

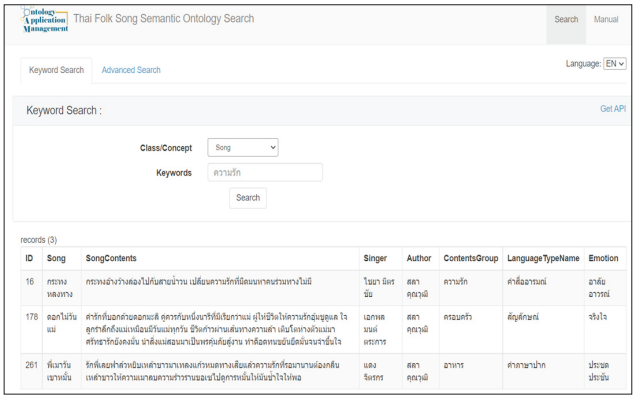


Fig. 7. Keyword search form.

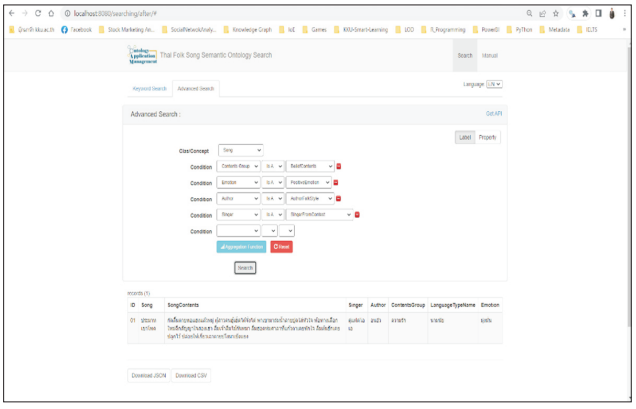


Fig. 8. Advanced search form.

used, language style, singer, song title, genre, background, and content. Fig. 7 shows an example of a keyword search from the song's class using "love" as a keyword. There are three entries found from 150 sample entries with details.

b) Advanced search. Users can give conditions to a search by selecting a class and filling in conditions accordingly. Fig. 8 shows an example of an advanced search using conditions in belief, positive mood, country song songwriter, and singer's background from contest. There is one entry found from the sample song list.

5. ONTOLOGY EVALUATION

5.1. Ontology Evaluation by Ontology Engineering and Thai Language Experts

The researchers inquired with three experts from the ontology and language field to evaluate the developed ontology. The results show that the overall effectiveness of the Semantic Web is high ($\bar{x}$=4.29). For the results of each aspect, it was found that the content aspect is ($\bar{x}$=4.38), the accuracy aspect is ($\bar{x}$=4.35), the format aspect is ($\bar{x}$=4.15), the ease of use aspect is ($\bar{x}$=4.22), the timeliness aspect is ($\bar{x}$=4.10), and the perceived usefulness aspect is ($\bar{x}$=4.51). The experts suggested that the system should provide the full lyrics. All of the results are in Table 3.

Table 3. Effectiveness evaluation from the experts

	Items	Average	Effectiveness degree
1	Content	4.38	High
2	Accuracy	4.35	High
3	Format	4.15	High
4	Ease of use	4.22	High
5	Timeliness	4.10	High
6	Perceived usefulness	4.51	High
Total average		4.29	High

5.2. Effectiveness Evaluations from Users

The evaluations employed a concept of knowledge retrieval from Yao et al. (2007) in which effectiveness refers to the outcomes from the search results. This implies that the accuracy and completeness of retrieval could be interpreted from the F-measure (Manning et al., 2008; Vanadic et al., 2012). The researchers asked two scholars from the Thai language department and a songwriter to evaluate the system. The researchers gathered 150 Thai country songs as a sample data set. Each evaluator conducted a keyword search with five random keywords in each class. Every evaluator entered 45 keywords and recorded all the search results. It was found that the evaluators could retrieve songs relevant to their needs by using 125 keywords, but the other ten words were not matched and not relevant. All 135 keywords were calculated as follows:

When relevant = number of all related documents
Retrieved = number of related documents which were retrieved

$$\text{Precision} \quad 125 = \frac{0.9259}{135}$$

$$\text{Recall} \quad 125 = \frac{0.8621}{0 + 135}$$

$$\text{F-measure} \quad \frac{2 \times (0.9259 \times 0.8621)}{0.9259+0.8624} = 89.28$$

The precision degree is 92.59 percent, the recall degree is

86.21 percent, and the overall effectiveness (F-measure) is 89.28 percent. According to the results, it shows that the Semantic Web is practical and effective.

6. DISCUSSION

The researchers developed an ontology of Thai country songs consisting of nine classes. The main feature of this ontology is that users can search for songs according to provided classes: song, singer, emotion, author, language used, language type, song style, originality, and content. The system was evaluated by domain experts and gained a precision degree of 92.59 percent, a recall degree of 86.21 percent, and an overall effectiveness degree (F-measure) of 89.28 percent. The high effectiveness degree demonstrates that the system was processed carefully regarding the needs analysis and development phase. Consequently, the system is suitable for scholars interested in applying Thai country songs to teaching and writing articles. It is also for songwriters who intend to do further studies in songwriting.

Furthermore, scholars and songwriters supported that this ontology could assist users in retrieving relevant and accurate information. This claim is related to a study by Chompookham et al. (2020) demonstrating that an ontology could help users expand their keyword searches and be developed and applied to be semantic tools for a song's meaning, sound, or local folk song notes in a digital library (Kaewboonma & Tuamsuk, 2018). An ontology could help create flexibility, collaboration, retrieval, and semantic retrieval for tourism (Nisheva-Pavlova & Lilev Pavlov, 2011; Prantner et al., 2007). Singthong and Kritpolviman (2017) revealed that they also applied an ontology and developed a Thai country song semantic search with Java and the Hozo-Ontology Editor program. Then, they further developed a Semantic Web with an OAM Framework where users can search for detailed and relevant examples.

Additionally, it is worth mentioning that the use of ontologies in digital libraries aligns with previous findings in the field, which have shown the potential for ontologies to expand keyword searches and serve as semantic tools for understanding the meaning, sound, and local characteristics of songs. This study also highlights the benefits of utilizing ontologies in digital libraries, such as adaptability, collaboration, and the ability to retrieve contextually relevant content.

However, the study recognizes the limitation of ontology in terms of the lack of diverse data sources. This suggests that future research could focus on incorporating additional sources in order to broaden the ontology and enhance its usefulness. Another avenue for future research could be the evaluation of ontologies through technology, which could reduce the workload on specialists.

7. CONCLUSION

The construction and evaluation of an ontology for Thai country songs, which consists of nine classes and allows users to search for songs based on a variety of parameters, was carried out here. This provides evidence that the system was well-designed and is suitable for use by academics and songwriters for the purposes of conducting research and instructing students. This ontology is consistent with the findings of other studies, which found that ontologies can be useful for expanding keyword searches and as semantic tools for understanding the meaning, sound, and local characteristics of songs located in a digital library. In addition, using ontologies in digital libraries improves flexibility, cooperation, and makes retrieval of relevant content for a given context easier.

On the other hand, the text draws attention to a shortcoming of the Thai country song ontology, notably the lack of a number of different data sources. For the advancement of investigation, it has been suggested that additional sources should be mapped in order to broaden the ontology. Additionally, research on evaluating ontologies through technology to lessen the amount of work that requires specialists might be an option to take into consideration for a future strategy. Overall, the present study provides a concise and extensive account of the construction and evaluation of a Thai country song ontology, and underlines the potentiality of the ontology for scholars and songwriters working on the subject.

ACKNOWLEDGEMENTS

This study was funded by the National Research Council of Thailand (NRCT) in the fiscal year 2020.

CONFLICTS OF INTEREST

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

REFERENCES

Antoniou, G., & van Harmelen, F. (2004). *A semantic web*

primer. MIT Press.

Bouihi, B., & Bahaj, M. (2019). Ontology and rule-based recommender system for e-learning applications. *International Journal of Emerging Technologies in Learning*, 14(15), 4-13. https://doi.org/10.3991/ijet.v14i15.10566

Broughton, V., Hansson, J., Hjørland, B., & López-Huertas, M. J. (2005). Knowledge organization. In L. Kajberg, & L. Lørring (Eds.), *European curriculum reflections on library and information science education* (pp. 133-148). Royal School of Library and Information Science.

Buranarach, M. (2017). *Introduction to ontology application development*. https://slideplayer.in.th/slide/10537452

Buranarach, M., Thein, Y. M., & Supnithi, T. (2013, December 2-4). A community-driven approach to development of an ontology-based application management framework. In H. Takeda, Y. Qu, R. Mizoguchi, & Y. Kitamura (Eds.), *Semantic technology* (pp. 306-312). Springer.

Chaichana, T. (2016). *[Roles of ordinary women in contemporary folk songs]*. (Master's thesis). Chiang Mai Rajabhat University, Chiang Mai, Thiland. Thai.

Chansanam, W., Ahmad, A. R., & Li, C. (2022). Contemporary and future research of digital humanities: A scientometric analysis. *Bulletin of Electrical Engineering and Informatics*, 11(2), 1143-1156. https://doi.org/10.11591/eei.v11i2.3596

Chansanam, W., Tuamsuk, K., Kwiecien, K., Sutthiprapa, K., & Supnithi, T. (2021). Thai tattoo wisdom's representation of knowledge by ontology. *Informatics*, 8(1), 3. https://doi.org/10.3390/informatics8010003

Chompookham, T., Limpinan, P., Sangkatip, W., & Yuensuk, T. (2020). [An ontology model for semantic information retrieval on children's infectious diseases]. *Journal of Applied Information Technology*, 6(2), 73-81. Thai. https://opacj.bsru.ac.th/cgi-bin/koha/opac-detail.pl?biblionumber=25328

Chutapanish, R. (2004). *[The research on famous Thai folk songs during 2001-2002: A study of contents, languages and social appearances]*. (Master's thesis). University of the Thai Chamber of Commerce, Bangkok, Thiland. Thai.

Fyeted, K. (2005). *[Political communication in Thai country songs: 1964-2004]*. (Doctoral dissertation). Thammasat University, Bangkok, Thiland. Thai.

Gašević, D., Djurić, D., & Devedžić, V. (2009). *Model driven engineering and ontology development* (2nd ed.). Springer.

Kabmala, M. (2020). [Developing Isanness conceptual metaphor ontology in Isan folksong (Isan Luk Thung)]. *Journal of Information Science*, 38(1), 59-87. Thai. https://doi.org/10.14456/jiskku.2020.4

Kaewboonma, N., & Tuamsuk, K. (2018). Ontology of folk songs in the Greater Mekong Subregion (GMS). *Knowledge Organization*, 45(1), 33-42. https://doi.org/10.5771/0943-7444-2018-1-33

Kentorrapak, P. (2021). *[Romantic discourse in contemporary Luk Thung Isan music]*. (Master's thesis). Mahasarakham University, Kantharawichai, Thiland. Thai.

Khuajaroen, P. (2014). *[Mistress: The representation of marginal women through Thai folk songs]*. (Unpublished master's thesis). Khon Kaen University, Khon Kaen, Thiland. Thai.

Kongsomsawaeng, M., Thanon, P., & Vipasrinimit, P. (2021). [Communication strategies of Isan country music digital album: Monkaen Kaenkoon]. *Dhurakij Pundit Communication Arts Journal*, 15(2), 33-73. Thai. https://so01.tci-thaijo.org/index.php/dpuca/article/view/248709/168924

Kozaki, K., Kitamura, Y., Ikeda, M., & Mizoguchi, R. (2002, October 1-4). Hozo: An environment for building/using ontologies based on a fundamental consideration of "role" and "relationship". In A. Gómez-Pérez, & V. R. Benjamins (Eds.), *Knowledge engineering and knowledge management: Ontologies and the semantic web* (pp. 213-218). Springer.

Krobthong, S. (1998). *[The evolution of Thai country songs in Thai society, 1938-1992]*. (Master's thesis). Chulalongkorn University, Bangkok, Thiland. Thai.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.

Nisheva-Pavlova, M., & Pavlov, P. (2011). Ontology-based search and document retrieval in a digital library with folk songs. *Information Services & Use*, 31(3-4), 157-166. https://doi.org/10.3233/ISU-2012-0645

Noy, N. F., & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*. http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html

Powell, C. (2022). *6 Characteristics of country music*. https://www.musicindustryhowto.com/characteristics-of-country-music/

Prantner, K., Ding, Y., Luger, M., Yan, Z., & Herzog, C. (2007, October 5-8). Tourism ontology and semantic management system: State-of-the-arts analysis. In P. Isaías, M. B. Nunes, & J. Barroso (Eds.), *IADIS International Conference: WWW/Internet 2007* (pp. 111-115). IADIS.

Singthong, S., & Kritpolviman, K. (2017, November 24). [Semantic web search using ontologyfor natural dye plants]. *Proceedings of the 7th STOU National Research Conference* (pp. 1190-1205). Thammasat University. Thai.

Sirisamphan, J., & Virulrak, S. (2010). [Lukthung song in Thai junior encyclopedia foundation]. In Thai Encyclopedia for Youth Project (Ed.), *Thai encyclopedia for youth* (Vol. 15, pp. 45-46). Thai Encyclopedia for Youth Project. Thai.

Srinok, S. (2010). *[The beliefs and factors influencing the be-*

liefs in puñña And pâpa in the contents in the north eastern contemporary Thai folk songs]. (Doctoral dissertation). Mahachulalongkornrajavidyalaya University, City, Thiland. Thai.

Thailand Research Fund, King Mongkut's University of Technology Thonburi, Thammasat University, National Electronics and Computer Technology Center, & Thai-Journal Citation Index Centre. (2001). *ThaiJO*. https://www.tci-thaijo.org/

Thanasilangkun, P. (2001). *[An analysis of the song literary works by Sala Khunnawutthi]*. (Master's thesis). Mahasarakham University, Kantharawichai, Thiland. Thai.

Tohombut, P. (2015). *[The role of folk songs for life by Mike Piromporn in the communication of culture and way of life of the workers in Bangkok]*. (Master's thesis). Dhurakij Pundit University, Bangkok, Thiland. Thai.

Tuamsuk, K., Chansanam, W., Chaikhambung, J., & Kaewboonma, N. (2018a). *[Digital humanities research]*. Klangnanawittaya Printing. Thai.

Tuamsuk, K., Chansanam, W., & Kaewboonma, N. (2018b). Ontology of folktales in the Greater Mekong Subregion. *International Journal of Metadata, Semantics and Ontologies*, 13(1), 57-67. https://doi.org/10.1504/IJMSO.2018.096454

UNESCO. (2016). *What is intangible cultural heritage?* https://ich.unesco.org/en/what-is-intangible-heritage-00003

UniNet. (2000). *What is TDC or Thai digital collection*. https://tdc.thailis.or.th/

Vandic, D., van Dam, J. W., & Frasincar, F. (2012). Faceted product search powered by the semantic web. *Decision Support Systems*, 53(3), 425-437. https://doi.org/10.1016/j.dss.2012.02.010

Viklin. (2023). *Best free online music streaming apps on Android and iOS (2023)*. https://naijaknowhow.net/free-online-music-streaming-applications/

Yaikoom, J. (2004). *[Thai lifestyle as reflected in country songs]*. (Master's thesis). Ramkhamhaeng University, Bangkok, Thiland. Thai.

Yao, Y., Zeng, Y., Zhong, N., & Huang, X. (2007, November 2-5). Knowledge retrieval (KR). In T. Y. Lin, L. Haas, J. Kacprzyk, R. Motwani, A. Broder & H. Ho (Eds.), *Proceeding of the 2007 IEEE/WIC/ACM International Conference on Web Intelligence* (pp. 729-735). IEEE.

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? Seonhee Lee (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 run under the 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 Article Contribution Management System (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 and authors' permission. Open access articles are automatically archived in the Korea Institute of Science and Technology Information (KISTI)'s Open Access repository (KPubS).

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, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly 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
Affilliation, Postal Address. E-mail
Author2
Affilliation, 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

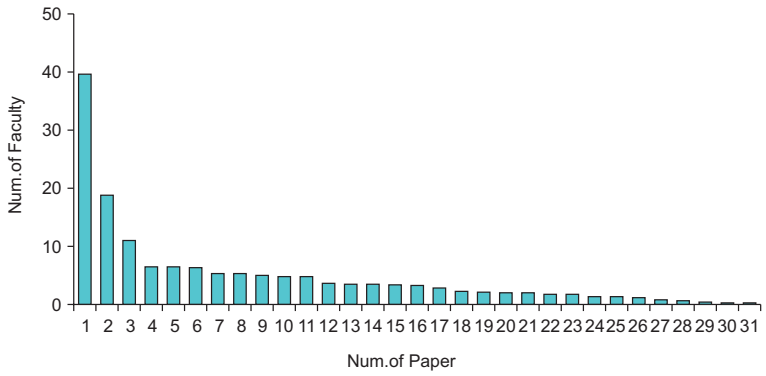


Fig. 1. Distribution of authors over publication count.

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

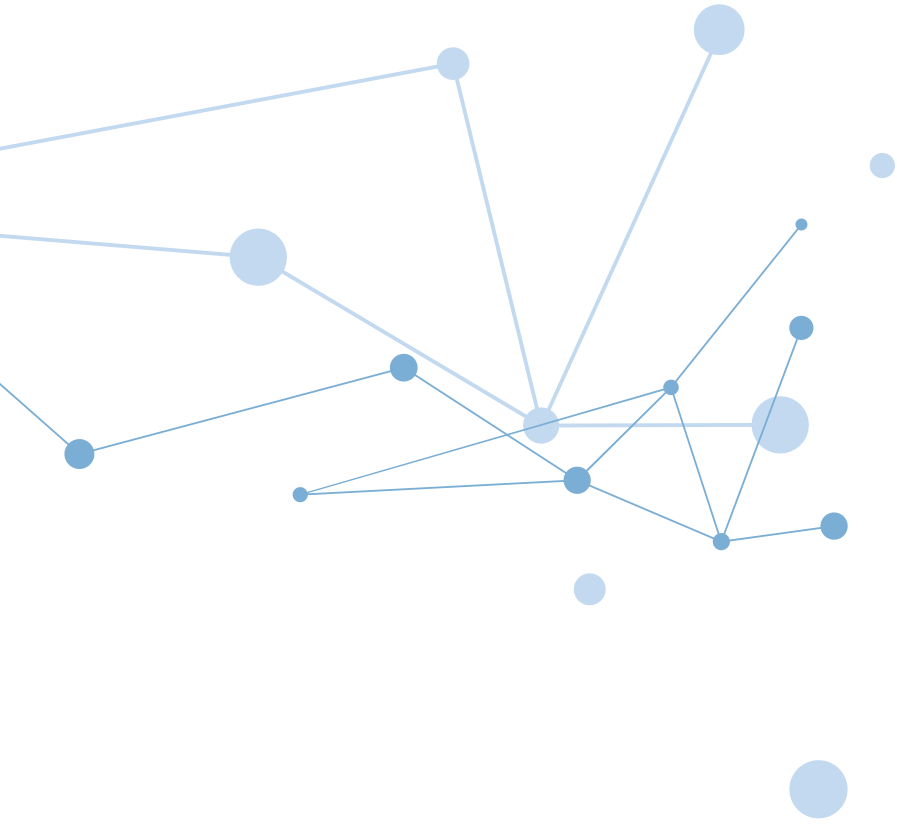
- Author, A., & Author, B. (Year). Chapter title. In A. Editor, B. Editor, & C. Editor (Eds.), *Book title* (pp. xx-xx). Publisher Name.
- Smith, K. & Brown, M. (2012). Author impact factor by weighted citation counts. In G. Martin (Ed.), *Bibliometric approach to quality assessment* (pp. 101-121). Springer.

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
http://www.jistap.org