eISSN : 2287-4577
pISSN : 2287-9099
https://www.jistap.org
Vol. 13 No. 3 September 30, 2025

# Journal of Information Science Theory and Practice

Journal of Information Science Theory and Practice

JISTaP

01 Predicting Workforce Needs in Thailand's Digital Industry: A Machine Learning Approach (2023-2024)
17 The Roles of Information Professionals in Community Engagement for Developing a Learning Community in Longan Farming
36 Developing Essential Datasets for a Community Cultural Capital Information System
55 The Effect of Institutional Pressures on the Intention to Adopt Environmental Disposal of Electronic Waste: Attitude as the Mediating Variable
72 Emoji Votes: Predicting Kathmandu's 2022 Mayoral Election via Facebook Sentiment
85 Exploring the Knowledge Structure and Thematic Evolution of NTIS-Related Research in South Korea (2008-2024): A Bibliometric Analysis

Vol. 13 I No. 3 I September 30, 2025

KiSTi
www.kisti.re.kr

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

JISTaP
Journal of Information Science Theory and Practice

# General Information

## Aims and Scope

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

## Indexed/Covered by

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

## Publisher

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

## Managing Editor: Hyekyong Hwang, Eungi Kim

## Copy Editor: Ken Eckert

## Design & Printing Company: MEDrang Inc.

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

## Open Access and Creative Commons License Statement

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

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

2025 Copyright © Korea Institute of Science and Technology Information

JISTaP
Journal of Information Science Theory and Practice

# Editorial Board


## Co-Editors-in-Chief

**Gary Marchionini**
University of North Carolina, USA

**Dong-Geun Oh**
Keimyung University, Korea

## Co-Associate Editors

**Kiduk Yang**
Youngnam Old Document Archive Center, Korea

**Hyejin Lee**
Korea Institute of Science and Technology Information, Korea

## Co-Managing Editors

**Hyekyong Hwang**
Korea Institute of Science and Technology Information, Korea

**Eungi Kim**
Keimyung University, Korea

## Editorial Board

**Dan Albertson**
University at Buffalo, SUNY, USA

**Daniel Martínez Ávila**
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

**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

**Kulthida Tuamsuk**
Khon Kaen University, Thailand

**Lihong Zhou**
Wuhan University, China

**Xiao-Zhi Gao**
University of Eastern Finland, Finland

**Wole Michael Olatokun**
University of Ibadan, Nigeria

## Consulting Editors

**Reysa Alenzuela**
Oriental Institute, ASCR, Czech Republic

**Wayne Buente**
University of Hawaii, USA

**Sujin Butdisuwan**
Mahasarakham University, Thailand

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

**M. Krishnamurthy**
DRTC, Indian Statistical Institute, India

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

**P. Rajendran**
SRM University, India


2025 Copyright © Korea Institute of Science and Technology Information

JISTaP
Journal of Information Science Theory and Practice

# Table of Contents

**Vol. 13 No. 3 September 30, 2025**
Journal of Information Science Theory and Practice • https://www.jistap.org

## Articles

01 Predicting Workforce Needs in Thailand's Digital Industry: A Machine Learning Approach (2023-2024)
Rachawit Tipsena, Wirapong Chansanam, Kulthida Tuamsuk, Paiboon Manorom

17 The Roles of Information Professionals in Community Engagement for Developing a Learning Community in Longan Farming
Chanchira Laorach, Suwannee Hoaihongthong

36 Developing Essential Datasets for a Community Cultural Capital Information System
Chatkamon Anontachai, Kulthida Tuamsuk, Wirapong Chansanam

55 The Effect of Institutional Pressures on the Intention to Adopt Environmental Disposal of Electronic Waste: Attitude as the Mediating Variable
Ghaith Abdulridha Mubdir, Sharizal Hashim, Abdul Aladheem Jabbar Alziady, Sabah Hussein Enayah

72 Emoji Votes: Predicting Kathmandu's 2022 Mayoral Election via Facebook Sentiment
Kapil Gurung, Shreejana Gnawali

85 Exploring the Knowledge Structure and Thematic Evolution of NTIS-Related Research in South Korea (2008-2024): A Bibliometric Analysis
Moon-Koo Kim, Byung-Hwan Hyun

105 **Call for Paper**

106 **Information for Authors**

2025 Copyright © Korea Institute of Science and Technology Information

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

RESEARCH PAPER
J Inf Sci Theory Pract 13(3): 1-16, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.3.1

# Predicting Workforce Needs in Thailand's Digital Industry: A Machine Learning Approach (2023-2024)

**Rachawit Tipsena**
Department of Information Science, Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand
E-mail: rachawit.t@kkumail.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

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

**Paiboon Manorom**
Department of Information Science, Faculty of Humanities and Social Sciences, Khon Kaen University, Khon Kaen, Thailand
E-mail: paiboon.m@kkumail.com

## ABSTRACT

Thailand's digital industry is experiencing rapid growth, driven by government initiatives and the widespread adoption of digital technologies. In response to this dynamic development, the present study proposes a predictive model to forecast future workforce requirements across five key segments of the digital industry: hardware and smart devices, software and software services, digital services, digital content, and telecommunications. Data were collected via web scraping from ten job advertisement websites between 2023 and 2024, resulting in a dataset comprising 24,494 job positions. The collected data underwent comprehensive preprocessing through natural language processing techniques—including text cleaning, punctuation removal, tokenization, stop word removal, and feature extraction using term frequency-inverse document frequency—to prepare it for analysis. Four supervised machine learning models (logistic regression, decision tree, k-nearest neighbors, and Naïve Bayes) were constructed and evaluated using performance metrics such as accuracy, precision, recall, and F1-score, alongside receiver operating characteristic-area under the curve analysis. The results demonstrate that the k-nearest neighbors model outperformed the other methods, achieving an accuracy (AUC) of 0.792, precision of 0.793, recall of 0.731, and an F1-score of 0.751, with all models yielding AUC values greater than 0.5. These findings indicate an upward trend in digital service job demand and underscore the model's potential utility in guiding workforce planning and human resource strategies in the digital sector. The study offers a robust, data-driven framework that can be adapted to forecast workforce needs in other rapidly evolving industries.

**Keywords:** predicting, workforce, digital industry, Thailand, machine learning

**Received:** October 2, 2024 **Revised:** May 23, 2025
**Accepted:** June 12, 2025 **Published:** September 30, 2025

***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. Open Access articles are automatically archived in the Korea Institute of Science and Technology Information (KISTI)'s Open Access repository (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

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

## 1. INTRODUCTION

### 1.1. Background

The digital industry was initially perceived as a new wave of industry that underwent constant changes in response to the dynamics of technology. The information and communications technology industry's expansion shaped the definition and scope of the digital sector. As for the promotion of the digital economy, the definition of this type of industry is quite broad and includes various manufacturing and service industries that involve digital technology or the use of digital technology. Five main categories divide the digital industry: a) the hardware and smart device industry; b) the software and software service industry; c) the digital service industry; d) the digital content industry; and e) telecommunications. Even the overall view of Thailand's digital industry has continued to grow every year, but there are still limitations to growth. Small and medium-sized entrepreneurs lack the capital to expand their businesses and cannot easily apply for a loan because the majority of digital businesses are considered too high-risk for banks to accept. Furthermore, the industry struggled to invest in innovation for high-value products and services, especially the linkage of digital research and development to the industrial sectors (depa, n.d.).

### 1.2. Digital Industry Context in Thailand

The digital industry job market must prioritize hiring individuals who meet the digital competency framework. These individuals must be knowledgeable, skilled, and capable of using digital technology effectively, appropriately, safely, creatively, ethically, and with freedom, aligning with the digital competency framework for Thai citizens in the twenty-first century. In this era, citizens must have the knowledge and capacity to survive the rapid changes, including a digital competency framework for future careers. In the Thai context, this phenomenon necessitates that citizens possess functional and hard skills that align with digital competencies and understanding. These skills include problem-solving with digital tools and adapting to digital transformations locally and internationally. In 2017, the United Nations Educational, Scientific and Cultural Organization compiled and analyzed a digital competency framework to establish guidelines for developing and educating digital citizenship. This led to introducing a digital competency framework, a digital curriculum, and digital education outcomes from leading organizations, all aimed at assessing and enhancing digital competencies for global digital citizens. However, companies in the industrial business sector can benefit from this framework. The human resource development and the human resource management sections of industrial businesses can apply the competency framework as a guideline for developing employees' competencies and skills, enabling them to achieve the desired competencies. Employers can use the competency framework to select and evaluate employee qualifications (BDE Digital Competence Test, n.d.). Therefore, the competency framework is an essential factor in studying the needs of the digital industry job market. The data can be analyzed to find trends and needs for future workforce requirements in the digital industry. This will help workers about to enter the job market in this area to prepare.

### 1.3. Problem Statement

We employed data analysis processes to predict future workforce requirements in the digital industry. These processes analyzed past and present data to identify trends in specific competency requirements necessary to meet the needs of the digital industry business sector. Artificial intelligence (AI) techniques can develop a predictive model as a statistical model to analyze data for forecasting efficiently. Generally, computer programs process data analysis for prediction through machine learning from big data, making it more precise and accurate than traditional workforce planning methods.

### 1.4. Previous Studies and Techniques

This study employed the k-nearest neighbor (KNN) technique (Okoye et al., 2024), a supervised learning technique, to predict graduation and measure the model performance by determining accuracy, precision, recall, and F1-score. Studies by Alnuaimi and Tabbara (2023) and Tadla et al. (2023) employed the decision tree technique to create a predictive model for sustainable careers. The results showed that we can use the machine learning process to predict careers, allocate human resources, and plan for a sustainable workforce. Next, the study by Khalaf et al. (2023) employed logistic regression techniques to analyze the decline and growth of careers, revealing important characteristics for career prediction and future career changes. Research by Rajath and Pandita (2022) employed logistic regression techniques to predict absenteeism for workforce planning and overtime cost reduction. The studies of Arromrit et al. (2023) and Chakraborty et al. (2021) also used the Naïve Bayes method to create a model that could predict when an employee would quit. This helped researchers figure out what causes people to

 https://doi.org/10.1633/JISTaP.2025.13.3.1

quit, such as the chance of job changes, so they could plan and measure ways to keep employees from quitting in the future. Additionally, machine learning techniques, such as Random Forest, Adaboost, XGBoost, Support Vector Machines, and Gradient Boosting, are highly effective in creating prediction models from big data to analyze human resource needs in organizations. The work by Puknilchai (2014) includes related factors, as does research by Gopinath and Appavu alias Balamurugan (2023), Lee et al. (2023), and Percia David et al. (2023). A predictive model has been developed to examine the needs for a suitable workload in the workforce. The results conclude that the number of human resources with appropriate proportions according to the workload still lacks prediction of organizational needs in other areas, such as specific job qualifications, skills, and abilities.

In addition to serving as a tool for forecasting future employment trends, predictive modeling is critical in guiding career development and aligning educational outcomes with labor market demands. Palee (2020) examined the concept of digital professional competency and proposed a career prediction model to prepare graduates for the evolving digital workforce. The findings indicated that students predominantly aspired to careers in four areas: content and design, technology services, business services, and product development—fields that align closely with the competencies expected of digital professionals. However, the study's scope was limited to these specific digital careers and did not address all five key sectors of the broader digital industry. To achieve comprehensive and reliable forecasting, it is essential to incorporate detailed labor market data across diverse occupational domains. This is crucial for executive decision-making, as accurate workforce predictions inform strategic planning, operational alignment, and supply chain forecasting (Bunsongkha, 2018). Accordingly, the present study aims to develop a predictive model to estimate future workforce requirements across the digital industry and to identify emerging trends in human resource demand within this sector.

## 2. OBJECTIVE

### 2.1. Development of a Machine Learning–Based Predictive Model for Workforce Estimation

To develop a machine learning-based predictive model for estimating future workforce requirements in Thailand's digital industry.

### 2.2. Model Performance Evaluation

To evaluate the model's performance using standard classification metrics, including accuracy, precision, recall, F1-score, and area under the receiver operating characteristic (ROC) curve (area under the curve [AUC]).

## 3. METHODOLOGY

This study employs a research methodology focused on developing a predictive model to forecast future workforce requirements in Thailand's digital industry. The methodology involved several key stages, beginning with the collection of job data through web scraping from various job advertisement websites. After gathering the data, natural language processing (NLP) techniques, such as text cleaning, tokenization, and stop word removal, were applied to prepare it for further analysis. The textual data was then transformed into a format suitable for model training using feature extraction techniques such as term frequency-inverse document frequency (TF-IDF). Various supervised learning models, including logistic regression, decision tree, KNN, and Naïve Bayes, were applied to build the predictive model. These models were evaluated using metrics such as accuracy, precision, recall, and F1-score to determine the best-performing model. The research methodology is illustrated in Fig. 1.

Fig. 1 is a framework that illustrates a systematic data pipeline for developing an educational application through five sequential phases. The process is initiated with web scraping-based data collection from job posting websites, suggesting a focus on labor market analytics.

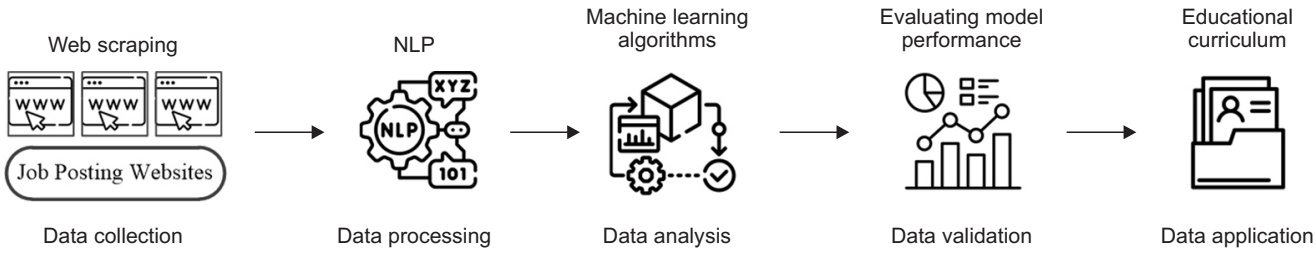



**Fig. 1.** Framework of the proposed method. NLP, natural language processing.

Subsequently, NLP techniques are employed to structure and clean the gathered information.

The data analysis phase incorporates machine learning algorithms to extract meaningful patterns and insights. The framework then progresses to data validation, where model performance is evaluated through metrics and statistical analysis to ensure reliability and accuracy. The final phase, data application, culminates in an educational curriculum or learning content, indicating the framework's ultimate goal of translating labor market data into educational outcomes.

This sequential methodology framework demonstrates an evidence-based approach to curriculum development. It leverages contemporary data science techniques to align educational content with real-world employment demands. The framework represents a holistic integration of web technologies, AI, and educational design principles.

### 3.1. Data Collection

The data collection was conducted on ten job advertisement and job recruitment websites in Thailand, with the following selection criteria: 1) The website featured job positions in the digital industry across five categories, including hardware and smart devices, software and services, digital services, digital content, and telecommunication; 2) It was dependable and verifiable; and 3) The data on the website was current and consistently displayed new job opportunities throughout the data collection period. To ensure methodological rigor in the data collection process, we implemented the following measures: defining the scope and data sources, planning and designing data collection, validating data accuracy, managing missing or duplicate data, ensuring ethical and legal compliance, and maintaining reproducibility. In this study, we selected data from job positions from 2023 to 2024 using a text mining process (Chansanama & Tuamsuk, 2020), web scraping, and extracting information from a large amount of data (Monika et al., 2024). We also conducted commands in Python using the *"Beautiful Soup library"* (Abodayeh et al., 2023), which is a web scraping library for extracting data from HTML and XML documents from Google Colab. Web display programs can write and execute these commands, extracting information as text on the target web page. Then, as shown in Table 1, we exported the data in plain text for 24,494 positions in CSV file format.

To ensure methodological rigor, data collection followed a structured process, including defining selection criteria, validating data accuracy through cross-checking with multiple job sources, and managing missing data using imputation techniques. Data integrity was maintained by removing duplicate records and standardizing job titles across different postings. These preprocessing steps were implemented using Python libraries such as *"Pandas"* and *"NumPy"* to enhance data reliability for subsequent analysis.

The job advertisements analyzed in this study predominantly represent workforce demand at the entry to mid-career levels, typically targeting recent graduates and professionals with up to ten years of experience. These roles constitute the majority of postings on recruitment platforms within Thailand's digital industry, reflecting the current labor market's emphasis on early-career talent acquisition.

### 3.2. Data Processing

NLP is a process that enables computers to interpret and understand the language used in human communication, with the aim of analyzing and extracting knowledge

**Table 1.** Descriptive attributes of job postings data

| No | Filed | Description |
|---|---|---|
| 1 | JobTitle | The specific title of the job being offered, clearly indicating the role, such as software engineer, project manager, or digital marketing specialist. This helps applicants understand the exact nature of the position being advertised |
| 2 | JobRequirements | A detailed list of the qualifications and skills necessary for the job. This could include technical skills, years of experience in a particular field, or additional competencies specified by the company, such as proficiency in specific software tools or fluency in a foreign language like English |
| 3 | JobWebsite | The domain or website where the job listing is posted. It provides a source where applicants can visit to explore more detailed information about the job position or the company itself |
| 4 | JobLabel | The classification indicates the specific category within the digital industry that the job belongs to. Includes 5 groups |
| 5 | JobDate | The date when a job post was published |

 https://doi.org/10.1633/JISTaP.2025.13.3.1

from large datasets (Rubio-Martín et al., 2024; SCB 10X, 2023). In this study, the NLP process involved the following crucial steps:

### 3.2.1. Text Cleaning

Text cleaning is the process of checking basic information to reduce errors, such as deleting redundant and unrelated words, correcting mistakes, handling lost text, or deleting HTML tags. This process can be governed by manual commands and organization.

### 3.2.2. Punctuation Removal

Punctuation removal is the process of making the text more standard by removing punctuation marks, symbols, or other special characters that may affect data processing, such as (!"#$%&'*+,-./:;<=>?@[]^_`{ |}), etc. This process is completed by importing the *"re"* and "NLTK" libraries in Python as a deletion tool.

### 3.2.3. Word Tokenization

The module utilized *"pythainlp.tokenize"* from the *PyThaiNLP* library to divide the text into words (Chotisarn et al., 2022). It then used the package *"word_tokenize"* for word segmentation (PyThaiNLP, n.d. -a), resulting in sentences being segmented into words, as shown in Table 2.

### 3.2.4. Stop Word Removal

The process of stop word removal involves eliminating frequently occurring words from documents or sentences that lack significance for analysis or knowledge extraction. To remove redundant words, we ran the *"pythainlp.corpus"* module in the *"PyThaiNLP"* library, followed by the *"thai_stopwords"* package (PyThaiNLP, n.d. -b) in Python, as illustrated in Table 3.

### 3.2.5. Feature Extraction (TF-IDF)

By configuring TF-IDF, feature extraction assigns features to words for analysis through machine learning processes. In this process, we ran the module *"sklearn.feature_extraction.text"* and the *"TfidfVectorizer"* and *"CountVectorizer"* package in the Python language (Chaudhary, 2020; Wendland et al., 2021) to find TF-IDF values, as shown in Table 4.

Table 4 presents the TF-IDF weighting for words extracted from job position descriptions, showing their importance in different job postings. Each row represents a job position ($P_1$ to $P_{24,494}$), while each column represents a specific word feature (e.g., "bachelor," "experience," "engineering," "database," and "knowledge"), with corresponding TF-IDF scores indicating the significance of each term in the respective job posting. The final column, "label," categorizes the job position into one of five groups (G1-G5).

From the Table 4, higher TF-IDF values suggest that a word is more relevant to a particular job description. For example, "engineering" has a high weight (0.4523) in $P_4$,

**Table 2.** Word tokenization

| | | | | | |
|---|---|---|---|---|---|
| bachelor | degree | engineering | science | fields | experience |
| system | administrator | Unix | Sun | Solaris | operating |
| computer | Sybase | Oracle | database | networking | preferable |
| thesis | educate | relate | database | Administrator | manage |
| data | SQL | JAVA | experience | knowledge | framework |
| software | dynamics | implementation | good | english | skills |
| configuration | monitor | performance | Exceptional | frontend | support |

**Table 3.** Thai stop words

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| têe | ที | mâi' | ไม่ | mee | มี | rĕu | หรือ | læ' | และ |
| jà' | จะ | dhæ̀ | แต่ | dâi | ได้ | dee | ดี | bhai' | ไป |
| sêung' | ซึง | an' | อัน | doy | โดย | dûay | ด้วย | à'-yâng | อย่าง |
| gà' | กะ | sŏong | สูง | bhen' | เป็น | kêun' | ขึ้น | næ' | แนะ |
| gan | การ | kăw | ขอ | nai' | ใน | bon' | บน | chên' | เช่น |
| lǽo | แล้ว | ma | มา | à'-yâ | อย่า | nam' | นำ | gwà | กว่า |

**Table 4.** Term frequency-inverse document frequency weighting word vectors

| Job position | $Bachelor_1$ | $Experience_2$ | $Engineering_3$ | $Database_4$ | $Knowledge_5$ | ... | $Word_{100}$ | Label |
|---|---|---|---|---|---|---|---|---|
| $P_1$ | 0.2377 | 0.2168 | 0.3763 | 0.0000 | 0.2584 | ... | 0.5549 | G2 |
| $P_2$ | 0.0779 | 0.0711 | 0.0000 | 0.0000 | 0.3270 | ... | 0.0000 | G1 |
| $P_3$ | 0.3171 | 0.1446 | 0.0000 | 0.1735 | 0.4528 | ... | 0.0000 | G3 |
| $P_4$ | 0.2858 | 0.2607 | 0.4523 | 0.0000 | 0.1834 | ... | 0.0000 | G3 |
| $P_5$ | 0.0620 | 0.0566 | 0.0000 | 0.0000 | 0.2386 | ... | 0.1448 | G4 |
| $P_6$ | 0.0000 | 0.1232 | 0.0000 | 0.0000 | 0.4922 | ... | 0.3070 | G5 |
| $P_7$ | 0.0000 | 0.3117 | 0.0000 | 0.0000 | 0.0000 | ... | 0.0000 | G4 |
| $P_8$ | 0.1577 | 0.0719 | 0.2496 | 0.2639 | 0.1766 | ... | 0.3586 | G1 |
| $P_9$ | 0.1916 | 0.0874 | 0.0000 | 0.0000 | 0.3125 | ... | 0.4356 | G5 |
| $P_{10}$ | 0.1053 | 0.1928 | 0.2041 | 0.3123 | 0.0000 | ... | 0.0000 | G5 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| $P_{24,494}$ | 0.1682 | 0.1534 | 0.2662 | 0.2814 | 0.2841 | ... | 0.5549 | G1 |

indicating its strong association with that job position. Similarly, "knowledge" has a high weight (0.4922) in $P_6$, showing its importance for that specific job.

## 3.3. Data Analysis

To construct a predictive model, the researchers separated the data into two primary categories: trained data and test data. They defined k-fold cross-validation and divided the data into four groups: k=10, k=15, k=20, and k=25. They then imported the *"sklearn.model_selection," "KFold," "cross_val_predict,"* and *"train_test_split"* libraries in Python. After that, the researchers used Python to create a trained data set, as follows:

### *3.3.1. Logistic Regression Model*

Create a logistic regression model, import the *"sklearn.linear_model"* library, and run the *"LogisticRegression"* package. Designed to find the sigmoid function for probability prediction, the logistic regression model typically is shaped like a curve on a graph with values ranging from 0 to 1. Based on the sigmoid equation, if the probability value is greater than or equal to 0.5, the model will assume the result to be 1, and if the value is less than 0.5, the model will assume the result to be 0 (Kanade, 2022).

$$f(x) = \frac{1}{1+e^{-x}} \tag{1}$$

where: $f(x)$ is the output of the logistic function, $x$ is the input to the function, $e$ base of natural logarithms.

### *3.3.2. Decision Tree Model*

Create a decision tree model, import the *"sklearn.tree"* library, and run the *"DecisionTreeClassifier"* package. By creating a hierarchical tree diagram with the root node as the top node and the starting point of the decision process, the decision tree model predicts successive outcomes (regression). The top node, known as the parent node, can split into two child nodes. The child node is the result of the parent node, while the leaf node represents the final decision. Every node must have a rule for decision-making, such as yes or no (Charbuty & Abdulazeez, 2021; Thanapatpisarn, 2022).

### *3.3.3. K-Nearest Neighbors (K-NN) Model*

Construct a model with KNN to import the *"sklearn.neighborslibrary"* and run the *"KNeighborsClassifier"* package. The KNN model is a set of k-values to study the closest member data, such as k=5, which means only the first five data points that are close to the targeted point. Then, using the following equation, calculate the distance between the sample data of interest and other data using the Euclidean distance (Ninthanom, 2021).

$$D_{Euclidean}(X_iY_i) = \sqrt{\sum_{k=1}^{n}(X_{i,k} - X_{j,k})^2} \tag{2}$$

where: $D_{Euclidean}(X_iY_i)$ is the Euclidean distance between the points $X_i$ and $Y_i$, $X_i$ and $Y_i$ are the two points in n-dimensional space, $X_{i,k}$ and $Y_{i,k}$ are the k-th coordinates of the points $X_i$ and $Y_i$, respectively, $n$ is the number of dimensions.

 https://doi.org/10.1633/JISTaP.2025.13.3.1

#### 3.3.4. Naïve Bayes Model

Construct a model with Naïve Bayes, import the *"sklearn.naive_bayes"* library, and run the *"MultinomialNB"* or *"GaussianNB"* package. The Naïve Bayes model is designed to investigate probability theory and statistics based on Bayes' theorem to predict the label of new data (Ray, 2025). The probability value can be calculated based on the following equation:

$$P(A|B) = \frac{P(B|A)\times P(A)}{P(B)} \quad (3)$$

where: $P(A|B)$ is the conditional probability of event $A$ occurring given that event $B$ has occurred, $P(B|A)$ is the conditional probability of event $B$ occurring given that event $A$ has occurred, $P(A)$ is the marginal probability of event $A$ occurring, $P(B)$ is the marginal probability of event $B$ occurring.

### 3.4. Model Evaluation

#### 3.4.1. Performance Evaluation Metrics

The study used accuracy, precision, recall, and the F1-score (Kundur & Mallikarjuna, 2022) as key performance indicators to evaluate and rate the predictive efficiency of the model across various dimensions. These metrics provide a comprehensive understanding of the model's capabilities in terms of correct predictions, handling imbalances in the dataset, and maintaining reliability in diverse scenarios. The following equations show how the efficiency was measured.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

$$F1-score = 2 \times \frac{Precision\times Recall}{Precision+Recall} \quad (7)$$

where: TP (true positives), the number of correctly predicted positive observations, TN (true negatives), the number of correctly predicted negative observations, FP (false positives), the number of incorrectly predicted positive observations, FN (false negatives), the number of incorrectly predicted negative observations.

#### 3.4.2. External Validation

In addition to evaluating performance through metrics such as accuracy, precision, recall, and the F1-score, which constitute internal validation, this study also includes external validation (Manorom et al., 2024). Utilizing data not included in the model development process, the external validation process evaluates the accuracy of the model. Alternative datasets source this data, representing new information absent from the original dataset. The purpose is to examine whether the model performs effectively across diverse scenarios. In this step, we test the developed model using data from one website against data from other websites and industries. By doing so, this process mitigates the risk of overfitting. External validation guarantees the model's generalizability and applicability to diverse contexts, without excessive customization to specific datasets.

### 3.5. Model Validation

To ensure the robustness and reliability of the predictive model, this study implements a structured validation process. Model validation is a crucial step in machine learning, distinguishing between development-phase validation and final model evaluation. The following validation techniques were applied:

#### 3.5.1. K-Fold Cross-Validation

To assess model generalization and reduce variance, k-fold cross-validation was performed. The dataset was split into k subsets (k=10, 15, 20, and 25), where each fold was used as a test set while the remaining k-1 folds were used for training. The average performance across all folds provided a reliable estimate of the model's predictive capability. This approach helps mitigate overfitting by ensuring the model is tested on multiple data partitions.

#### 3.5.2. Hyperparameter Tuning

To optimize model performance, hyperparameter tuning was conducted using grid search and random search techniques. The *"GridSearchCV"* function from *"scikit-learn"* was employed to systematically evaluate different hyperparameter combinations, selecting the best-performing configuration based on accuracy, precision, recall, and F1-score. Randomized search was also used to explore a broader range of hyperparameter values efficiently.

#### 3.5.3. Data Quality Verification

Before training, the dataset underwent integrity checks and preprocessing verification to ensure data consistency. This included:

- Handling missing values by applying imputation

techniques or removing incomplete entries.
- Eliminating duplicate records to prevent data leakage.
- Normalizing text data to standardize formats across job postings.
- Ensuring label distribution balance to prevent model bias toward overrepresented classes.

By implementing these validation steps, this study ensures that the predictive model is well-optimized, generalizes effectively, and maintains methodological rigor in machine learning-based workforce prediction.

### 3.6. Data Application

This section outlines the process of utilizing the processed and analyzed data to derive actionable insights and implement practical solutions. We applied the results of the predictive models to two primary domains: educational curriculum development and digital industry workforce planning.

The educational implications of this study are particularly relevant to programs in information science and other digital-related disciplines. The predictive findings offer valuable insights for aligning academic curricula with emerging industry demands in data science, AI, digital marketing, user experience/user interface (UX/UI) design, and software development. These areas are increasingly critical to the growth and competitiveness of Thailand's digital economy.

The findings from this study can be used by educational institutions to integrate the identified skills and competencies into their curricula, aligning academic programs with industry demands. This guarantees that graduates possess the essential skills to fulfill the workforce demands of the digital economy. The insights can inform the design of specialized courses and certifications to address emerging trends in digital services, such as data analysis, AI, and UX design.

The digital industry can apply the data to support strategic workforce planning. Organizations can leverage the predictions the predictions to identify future skill gaps, enabling them to implement targeted training programs and talent acquisition strategies. The findings also can also facilitate industry-wide initiatives to prepare for the increasing demand for digital service professionals by fostering collaborations between private sector organizations and educational institutions.

The application of data derived from this study bridges the gap between academia and industry, contributing to a more skilled and adaptable workforce and supporting sustainable growth in Thailand's digital economy.

This study is structured as a case study on workforce prediction in Thailand's digital industry, utilizing job posting data from 2023 to 2024. To ensure clarity in scope, the dataset limitations and case boundaries are explicitly stated in the methodology section. Additionally, the industry analysis has been incorporated into the methodology to provide contextual background, while the conclusion focuses on model performance, implementation constraints, and future research directions for broader workforce applications.

## 4. STUDY RESULTS

In this section, we present the findings from our study, detailing the development and evaluation of models using various techniques. We rigorously tested and analyzed each model's performance to determine its effectiveness.

### 4.1. Study Results with Model Development

This section presents the outcomes of the model development process using four different techniques. We partitioned the data using k-fold cross-validation to ensure robust model testing and evaluation. Specifically, the dataset was split into 80% for training and 20% for testing, with a random state of 5 to ensure reproducibility. We calculated key performance metrics such as accuracy, precision, recall, and F1-score to evaluate each model's effectiveness. Table 5 summarizes the detailed results.

The results demonstrate that the KNN model, with k set to 5, outperforms other predictive models, achieving the highest k-fold number of 25. The model demonstrates an accuracy value of 0.792, a precision value of 0.793, a recall value of 0.731, and an F1-value score of 0.751.

From Table 5, it is evident that KNN consistently achieves the highest performance across all k-fold values, with notable improvements in accuracy, precision, recall, and F1-score as k increases. The decision tree model performs well, ranking second in most cases, but falls short compared to KNN. Logistic regression exhibits moderate performance but struggles with recall, indicating weaker capability in identifying all relevant instances. Naïve Bayes consistently underperforms across all metrics, showing particularly low recall and F1-score, suggesting limited predictive reliability. These results confirm that KNN is the most effective model for workforce prediction in this study.

 https://doi.org/10.1633/JISTaP.2025.13.3.1

**Table 5.** Predictive model performance evaluation results

| k-fold | Algorithm | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|---|
| 10-fold | Logistic regression | 0.624 | 0.715 | 0.504 | 0.507 |
| | Decision tree | 0.707 | 0.668 | 0.665 | 0.666 |
| | K-nearest neighbor | 0.790 | 0.787 | 0.727 | 0.747 |
| | Naïve Bayes | 0.545 | 0.555 | 0.489 | 0.474 |
| 15-fold | Logistic regression | 0.625 | 0.513 | 0.507 | 0.508 |
| | Decision tree | 0.707 | 0.663 | 0.662 | 0.663 |
| | K-nearest neighbor | 0.791 | 0.789 | 0.729 | 0.749 |
| | Naïve Bayes | 0.545 | 0.553 | 0.489 | 0.474 |
| 20-fold | Logistic regression | 0.628 | 0.677 | 0.509 | 0.512 |
| | Decision tree | 0.712 | 0.672 | 0.670 | 0.671 |
| | K-nearest neighbor | 0.792 | 0.790 | 0.729 | 0.749 |
| | Naïve Bayes | 0.546 | 0.554 | 0.489 | 0.474 |
| 25-fold | Logistic regression | 0.628 | 0.715 | 0.510 | 0.511 |
| | Decision tree | 0.714 | 0.676 | 0.671 | 0.673 |
| | K-nearest neighbor | 0.792 | 0.793 | 0.731 | 0.751 |
| | Naïve Bayes | 0.545 | 0.553 | 0.489 | 0.474 |

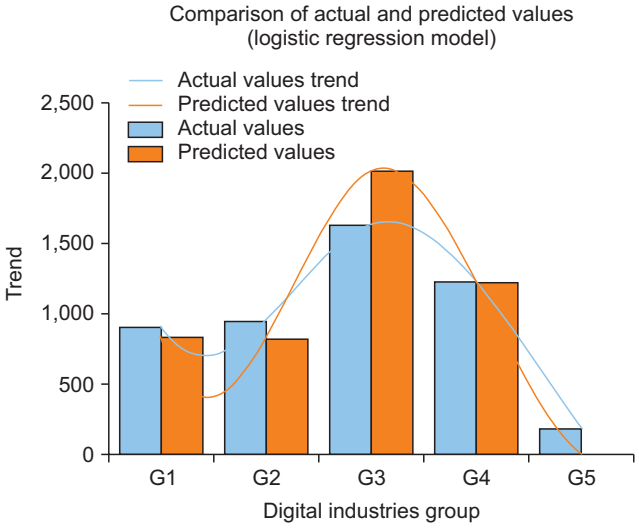



**Fig. 2.** Comparison of predicted and actual values using logistic regression model.

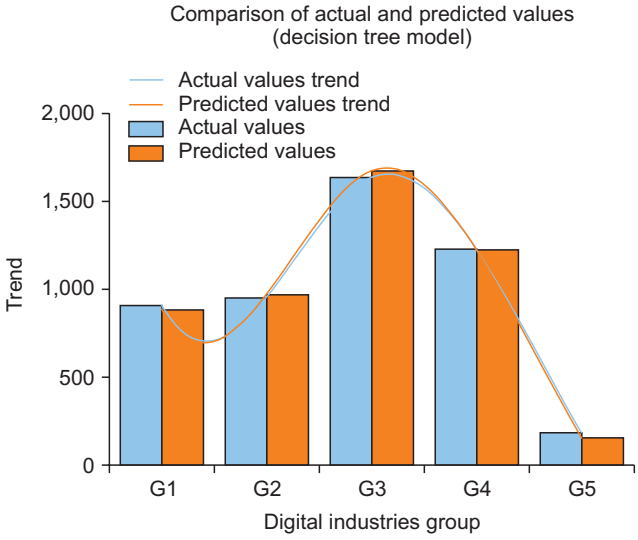



**Fig. 3.** Comparison of predicted and actual values using decision tree model.

## 4.2. Results of Prediction

When four predictive models were applied to predict future workforce requirements in the digital industry, three of the models consistently indicated that the digital service sector would demand the highest number of jobs. This outcome is illustrated in Figs. 2-5, which demonstrate that the digital service industry is expected to experience the greatest increase in job demand across the models.

Fig. 2 illustrates a comparison between actual and predicted job demand values using the logistic regression model across five digital industry groups (G1-G5). The actual values (blue bars) and predicted values (red bars) generally follow a similar trend, with the model effectively capturing the increasing demand in G3, which shows

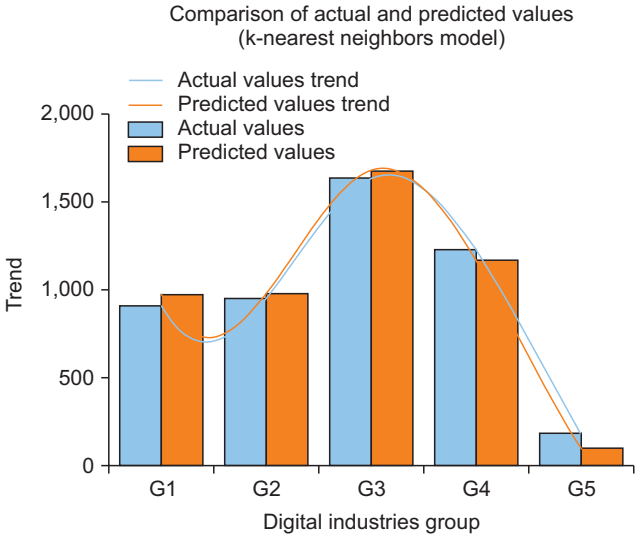



**Fig. 4.** Comparison of predicted and actual values using k-nearest neighbor model.

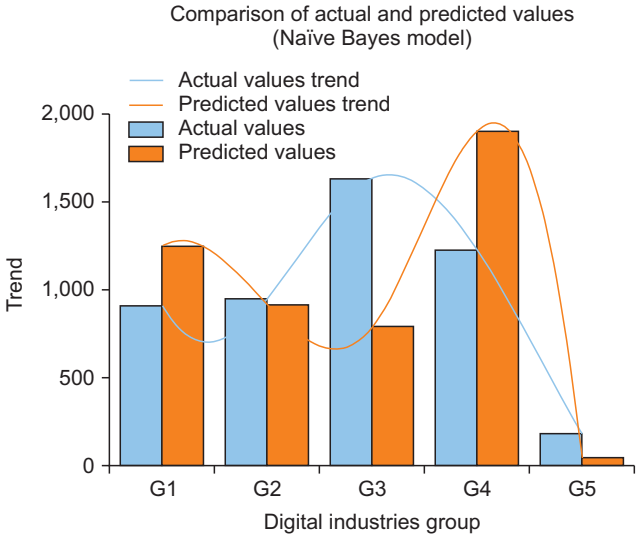



**Fig. 5.** Comparison of predicted and actual values using Naïve Bayes model.

the highest job demand. The trend lines indicate that the model aligns well with actual trends in G1 to G3, while slight underestimations occur in G4 and G5. Overall, the model successfully predicts workforce demand patterns in the digital industry, with minor deviations suggesting potential improvements through further optimization or additional data sources.

Fig. 3 illustrates a comparison between actual and predicted job demand values using the decision tree model across five digital industry groups (G1-G5). The actual values (blue bars) and predicted values (red bars) generally follow a similar pattern, with the model effectively capturing the peak demand in G3. The trend lines indicate that the model aligns well with actual trends, especially in G1 to G4, but shows a slight overestimation in G3 and G4, while G5 exhibits minor underestimation. Overall, the decision tree model demonstrates strong predictive capability in workforce demand estimation, with some discrepancies suggesting possible refinements in feature selection and model tuning for improved accuracy.

Fig. 4 presents a comparison of actual and predicted job demand values using the KNN model across five digital industry groups (G1-G5). The actual values (blue bars) and predicted values (red bars) show a similar distribution, with the model accurately capturing job demand trends in G1 to G3. However, minor discrepancies appear in G4 and G5, where the model slightly underestimates demand. The trend lines indicate that the KNN model follows the general shape of actual demand but tends to smooth out variations, leading to slight overestimations in some groups. Overall, the KNN model performs well in workforce prediction but could benefit from further tuning to refine accuracy in lower-demand categories.

Fig. 5 presents a comparison of actual and predicted job demand values using the Naïve Bayes model across five digital industry groups (G1-G5). While the model captures some trends, notable discrepancies exist, particularly in G3 and G4, where the predicted values (red bars) significantly deviate from actual values (blue bars). The trend lines show that the model struggles to align with actual job demand patterns, especially overestimating demand in G4 and underestimating in G3 and G5. This suggests that the Naïve Bayes model may not be as effective in workforce prediction due to its assumption of feature independence, which may not fully capture complex job market patterns. Further refinements, such as feature engineering or hybrid modeling approaches, could improve its predictive performance.

### 4.3. ROC Curve and AUC Analysis of Predictive Models

The ROC curve and the AUC for each model, as presented in Figs. 6-9, provide a comprehensive evaluation of their performance. The ROC curve visually represents the trade-off between the TP rate (sensitivity) and the FP rate (1-specificity) at various threshold settings. The AUC quantifies the overall ability of the model to discriminate between the positive and negative classes, with a value of 1 indicating perfect classification and a value of 0.5 representing random guessing. Together, these metrics offer

 https://doi.org/10.1633/JISTaP.2025.13.3.1

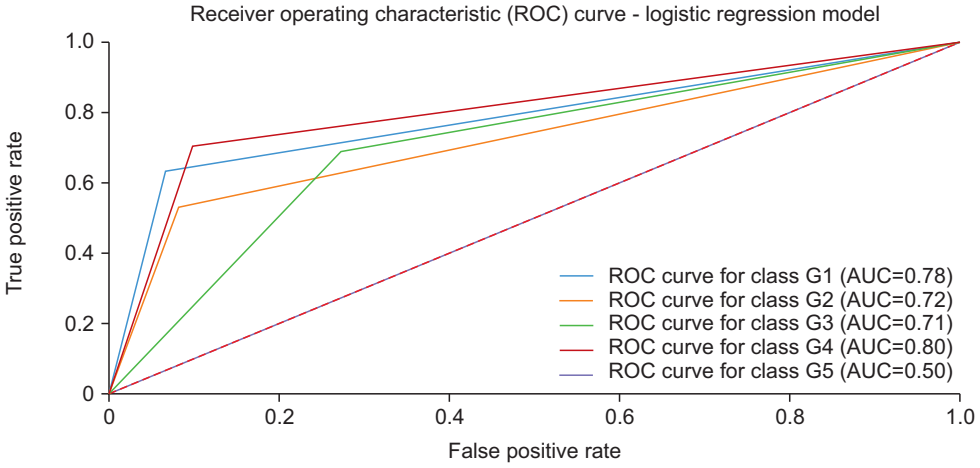



**Fig. 6.** Receiver operating characteristic (ROC) curve values are displayed using a logistic regression model. AUC, area under the curve.

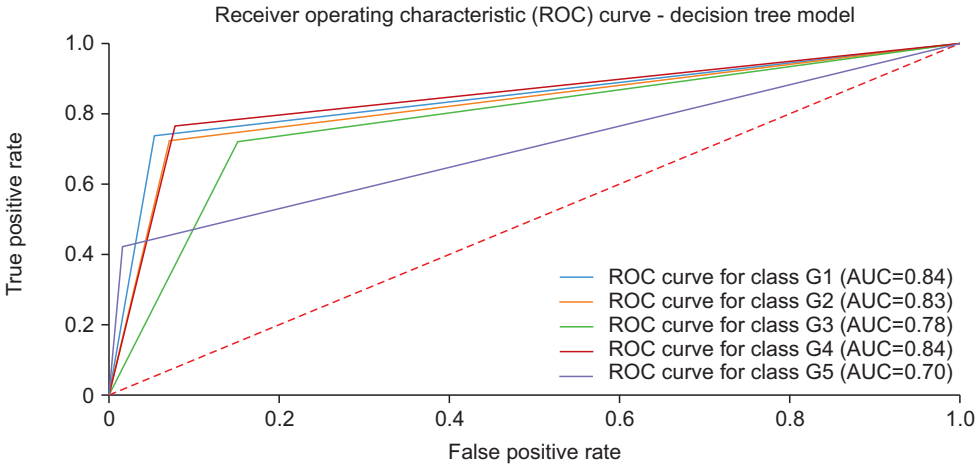



**Fig. 7.** Receiver operating characteristic (ROC) curve values are displayed using a decision tree model. AUC, area under the curve.

a clearer and more complete understanding of how effectively each model distinguishes between the five classes.

Fig. 6 presents the ROC curve for the logistic regression model, illustrating its classification performance across five digital industry groups (G1-G5). The AUC values indicate the model's ability to distinguish between classes, with G4 achieving the highest AUC (0.80), followed by G1 (0.78), suggesting strong predictive performance in these categories. G2 (0.72) and G3 (0.71) show moderate classification capability, while G5 (0.50) indicates poor distinction, performing at a random classification level. Overall, the model performs well for certain groups but may require further optimization, particularly for improving classification in G5.

Fig. 7 shows the ROC curve for the decision tree model, illustrating its classification performance across five digital industry groups (G1-G5). The AUC values indicate that G1 (0.84) and G4 (0.83) exhibit relatively strong predictive capability, while G2 (0.83) and G3 (0.78) show moderate performance. G5 (0.70) demonstrates the lowest distinction, suggesting that further model refinement may be necessary to enhance its classification accuracy in this group.

Fig. 8 presents the ROC curve for the KNN model, illustrating its classification performance across five digital industry groups (G1-G5). The AUC values indicate that G1 and G4 achieve the highest classification performance with an AUC of 0.88, followed by G2 (0.86) and G3 (0.84), showing strong but slightly lower predictive capability. G5 has the lowest AUC (0.72), suggesting weaker model performance in distinguishing this category. Overall, the KNN model demonstrates robust classification accuracy for most groups, with some limitations in predicting G5, which may require further optimization or feature engineering.

Fig. 9 presents the ROC curve for the Naïve Bayes model, illustrating its classification performance across five digital industry groups (G1-G5). The AUC values

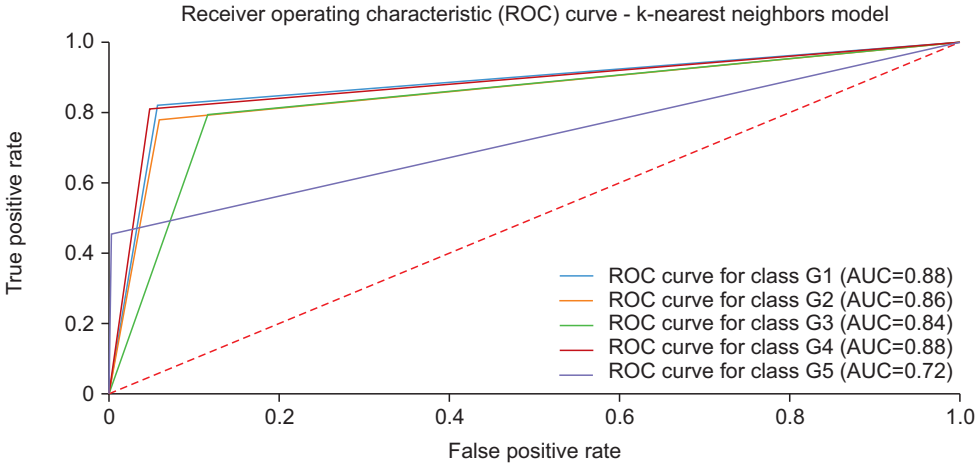



**Fig. 8.** Receiver operating characteristic (ROC) curve values are displayed using a k-nearest neighbor model. AUC, area under the curve.

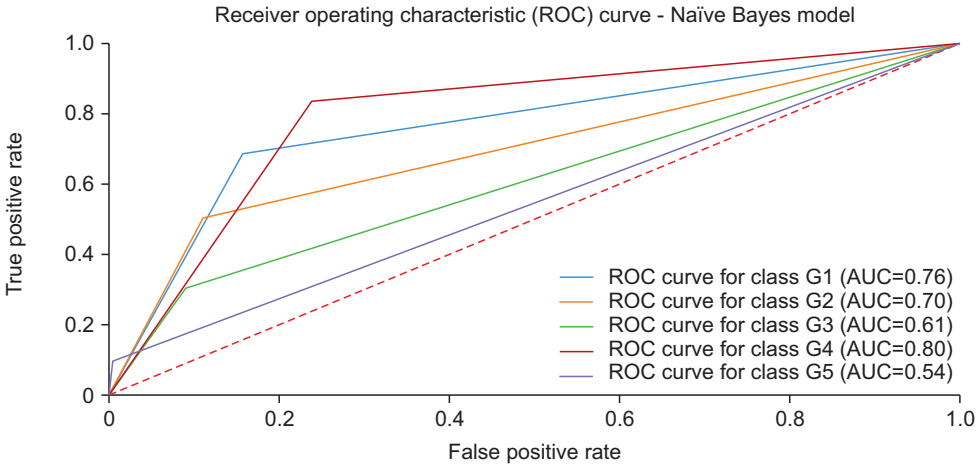



**Fig. 9.** Receiver operating characteristic (ROC) curve values are displayed using a Naïve Bayes model. AUC, area under the curve.

indicate that G4 achieves the highest predictive capability (AUC=0.80), followed by G1 (0.76) and G2 (0.70), showing moderate performance. However, G3 (0.61) and G5 (0.54) exhibit lower AUC scores, suggesting weaker classification accuracy, particularly for G5, which is close to random prediction. Overall, while the Naïve Bayes model performs reasonably well for certain groups, its assumptions of feature independence may limit its effectiveness in capturing complex patterns, necessitating further optimization or alternative modeling approaches.

Table 6 presents the AUC values for each predictive model across five different groups (G1-G5). These values indicate each model's performance in terms of distinguishing between classes, with higher values suggesting better model accuracy in predicting workforce requirements. The comparison between models, including logistic regression, decision tree, KNN, and Naïve Bayes, reveals significant variations in performance across different groups.

From the Table 6, the KNN model demonstrates the highest overall performance, achieving the highest AUC values across most groups, particularly in G1 (0.88) and G4 (0.88). The decision tree model also performs well, with AUC values exceeding 0.80 in all groups except G3 and G5. logistic regression shows moderate performance, particularly struggling in G5 (0.50), indicating poor classification ability for this group. The Naïve Bayes model exhibits the weakest performance, with an AUC as low as

**Table 6.** Area under the curve values in each predictive model

| Model | Area under the curve | | | | |
|---|---|---|---|---|---|
| | G1 | G2 | G3 | G4 | G5 |
| Logistic regression | 0.78 | 0.72 | 0.71 | 0.80 | 0.50 |
| Decision tree | 0.84 | 0.83 | 0.78 | 0.84 | 0.70 |
| K-nearest neighbor | 0.88 | 0.86 | 0.84 | 0.88 | 0.72 |
| Naïve Bayes | 0.76 | 0.70 | 0.61 | 0.80 | 0.54 |

 https://doi.org/10.1633/JISTaP.2025.13.3.1

0.54 for G5 and 0.61 for G3, suggesting difficulty in distinguishing these categories effectively. Overall, the KNN and decision tree models outperform the others, making them more reliable for workforce prediction tasks.

## 5. DISCUSSION

Using four machine-learning techniques, this study developed a predictive framework to forecast future workforce requirements in Thailand's digital industry. Notably, the KNN model outperformed logistic regression, decision tree, and Naïve Bayes models, achieving an accuracy of 0.792, precision of 0.793, recall of 0.731, and an F1-score of 0.751 under a 25-fold cross-validation setup. While the KNN model demonstrated marked improvements in accuracy, precision, and F1-score compared to previous studies (Okoye et al., 2024), the recall metric remained consistent, suggesting that specific data characteristics may influence model sensitivity.

Our findings reinforce the growing body of research that highlights the efficacy of machine learning in workforce prediction. Similar to the work of Okoye et al. (2024), the superior performance of the KNN model underlines its potential for accurately forecasting labor market trends based on job posting data. Additionally, the observed upward trend in digital service job demand is consistent with industry reports (6Wresearch, 2025; TheReporterAsia, 2024) and is likely driven by rapid digital transformation, government initiatives, and evolving consumer behavior (Dharmaraj, 2024).

Despite these promising results, several limitations should be noted. First, the dataset is restricted to job postings from 2023 to 2024 within Thailand's digital industry, potentially limiting the model's generalizability to broader or multi-regional contexts. Second, the reliance on text-based job advertisement data may introduce biases due to variability in job description formats and potential underrepresentation of specific roles. Future research should consider integrating cross-industry, longitudinal, and multi-regional data to enhance the robustness and applicability of the predictive model.

The insights derived from this study have significant practical implications. For industry stakeholders, the ability to predict workforce requirements can inform strategic human resource planning, targeted training initiatives, and proactive talent acquisition strategies. For educational institutions, these findings provide a data-driven basis for curriculum design—ensuring academic programs align with emerging digital competencies and market demands. The study supports a more responsive and future-ready workforce by bridging the gap between labor market intelligence and educational outcomes.

The introduction outlined the rapid evolution of Thailand's digital industry and the need for agile workforce planning to address emerging skill gaps. The results presented here confirm these trends and offer a quantitative validation of the potential for machine learning to enhance predictive accuracy in workforce planning. The strong performance of the KNN model, in particular, validates the initial hypothesis that advanced data analytics can provide actionable insights into the future needs of the digital service sector.

It is important to recognize that the predictive results primarily reflect workforce demand for early-career professionals. This limitation arises from the dataset, which was derived mainly from job postings targeting recent graduates and individuals with a few years of professional experience. Future research should incorporate data sources that capture senior-level and executive positions to obtain a more comprehensive view of labor market trends.

In summary, this research demonstrates that machine learning, particularly the KNN model, can be a powerful tool in forecasting workforce requirements within the digital industry. The integration of robust validation techniques—such as k-fold cross-validation, hyperparameter tuning, and comprehensive data quality checks—ensures the model is reliable and adaptable. These findings advance the theoretical framework of labor market intelligence and provide a practical foundation for strategic workforce planning and curriculum development. Ultimately, the study highlights the critical role of data-driven decision-making in preparing for the evolving challenges of the digital economy.

## 6. CONCLUSION

This study developed a comprehensive predictive model to forecast the future workforce requirements of Thailand's rapidly evolving digital industry. The data collection process involved gathering information from multiple sources, including job advertisements, recruitment websites, and relevant web pages, ensuring a diverse and extensive dataset for analysis. After data collection, we utilized NLP techniques to clean, tokenize, and extract meaningful features from the text, transforming it into a suitable format for predictive modeling.

Four machine learning algorithms were employed to

construct the predictive models: logistic regression, decision tree, KNN, and Naïve Bayes. Each model was evaluated using several performance metrics, including accuracy, precision, recall, F1-score, and AUC, to gauge its capacity to forecast workforce trends in the digital industry. After constructing the models, a thorough evaluation was conducted to re-examine their prediction performance.

The results of the predictive models indicate a continuous increase in workforce requirements within the digital industry, which is growing at a faster pace compared to other sectors. Specifically, the KNN model demonstrated superior prediction accuracy, particularly in the digital services sector. These findings provide valuable insights into the future skill requirements and workforce trends, which can be highly beneficial for the development of educational curricula. Additionally, the insights from these predictions can serve as a strategic tool for digital industry organizations to better plan for the workforce they will need in the future.

Educational institutions can align their programs with industry needs and equip graduates with the necessary skills for the future workforce by analyzing the forecasted trends in job demand and required skill sets. The following can be guided by the study's findings:

- Curriculum design: By incorporating the identified key skills and competencies into their courses, educational institutions can equip students for high-demand roles in the digital industry, especially in digital service roles as predicted by the predictive model. For instance, academic programs can emphasize areas such as UX/UI design, e-commerce specialists, data analysts, social media managers, or digital marketing managers if the prediction indicates a growing need for expertise in these areas.
- Course offerings: Based on the study's predictions, universities and vocational training centers can introduce specialized courses or certifications focused on emerging technologies, ensuring that the workforce is ready for upcoming digital trends. This would enable organizations in the digital sector to hire graduates who possess the specific skills required to adapt to industry shifts.
- Skills gap analysis: By comparing the study's workforce predictions with the current curriculum, institutions can identify gaps in their existing offerings. This analysis helps institutions make informed decisions about adapting or expanding their programs to better meet industry requirements. For digital organizations, this alignment ensures a smoother transition from education to employment, as graduates will be more immediately employable in the rapidly evolving job market.
- Long-term planning: Educational institutions can use the study's results to inform long-term curriculum planning, ensuring that academic programs remain relevant and responsive to the evolving digital job market. For the digital industry, this allows for a more predictable pipeline of skilled talent, helping organizations plan for future recruitment needs with greater accuracy.
- Making strategic decisions: Predictive data aids educational institutions and organizations in making informed strategic decisions about curriculum development and workforce training, ensuring that the right skills are being developed to meet future requirements.

This study presents a case study on workforce prediction in Thailand's digital industry using job posting data from 2023 to 2024. The findings demonstrate the effectiveness of machine learning models in predicting workforce requirements; however, the study's scope is limited by geographic, temporal, and industry constraints. To enhance generalizability, future research should incorporate multi-regional and cross-industry datasets, extend the study period, and explore additional data sources beyond job postings. These steps would help broaden the applicability of the proposed framework for workforce planning on a larger scale.

In conclusion, the predictive models collectively indicate a continuous increase in workforce requirements within Thailand's rapidly growing digital industry, particularly in digital services. Expected to outpace other sectors, this surge underscores the importance of preparing for the rising need for skilled professionals in the digital economy. The study's findings offer significant insights for modifying educational programs and curricula to better meet future workforce requirements, equipping students with the necessary skills for digital service roles in the digital industry. Additionally, these insights are crucial for digital industry organizations in strategically planning their workforce needs and ensuring that their teams are well-prepared to meet the challenges of an ever-evolving market. Strategic workforce planning and curriculum development are critical to addressing this growing demand and meeting the evolving needs of both the education sector and the industry, ultimately contributing to a more

 https://doi.org/10.1633/JISTaP.2025.13.3.1

skilled and industry-ready workforce.

Accordingly, the educational reforms proposed in this study are specifically tailored to the context of information science and digital technology programs rather than being broadly generalized. These reforms are designed to address the disconnect between academic training and industry requirements, by ensuring that graduates are equipped with the relevant competencies demanded in key areas such as digital services, marketing technology, and intelligent systems.

## CONFLICTS OF INTEREST

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

## ACKNOWLEDGMENTS

The researchers would like to thank the job advertisement website for giving them the opportunity to use the data in their research. This work has received a scholarship under the Research Assistant Program from Khon Kaen University, Thailand (Grant No. RA2566-12).

## REFERENCES

6Wresearch. (2025). *Thailand digital market (2025-2031) outlook | Trends, value, revenue, growth, analysis, share, forecast, companies, size & industry*. https://www.6wresearch.com/industry-report/thailand-digital-market

Abodayeh, A., Hejazi, R., Najjar, W., Shihadeh, L., & Latif, R. (2023, March 14-15). Web scraping for data analytics: A BeautifulSoup implementation. In A. Rehman, T. Saba, N. S. Jamail, & R. Latif (Eds.), *2023 Sixth International Conference of Women in Data Science at Prince Sultan University (WiDS PSU)* (pp. 65-69). IEEE. https://doi.org/10.1109/WiDS-PSU57071.2023.00025

Alnuaimi, N. H., & Tabbara, H. (2023, December 10-11). Predicting the sustainable career paths of working women in the UAE Private sector using machine learning. *2023 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT)* (pp. 174-178). IEEE. https://doi.org/10.1109/GCAIoT61060.2023.10385133

Arromrit, T., Srisakaew, K., Roswhan, N., & Mahikul, W. (2023, August 25-27). A supervised machine learning method for predicting the employees turnover decisions. *2023 IEEE 8th International Conference on Software Engineering and Computer Systems (ICSECS)* (pp. 122-127). IEEE. https://doi.org/10.1109/ICSECS58457.2023.10256357

BDE Digital Competence Test. (n.d.). *Digital competency framework for Thai citizens*. https://www.ondedigitalct.online/competenceframe

Bunsongkha, A. (2018). *Eye classes demand forecasting case study: RAK WAN shop* (master's thesis). Srinakharinwirot University, Bangkok, Thailand.

Chakraborty, R., Mridha, K., Shaw, R. N., & Ghosh, A. (2021, September 24-26). Study and prediction analysis of the employee turnover using machine learning approaches. *2021 IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)* (pp. 65-69). IEEE. https://doi.org/10.1109/GUCON50781.2021.9573759

Chansanama, W., & Tuamsuk, K. (2020). Thai Twitter sentiment analysis: Performance monitoring of politics in Thailand using text mining techniques. *International Journal of Innovation, Creativity and Change*, 11(12), 436-452. https://www.ijicc.net/images/vol11iss12/111227_Chansanam_2020_E_R.pdf

Charbuty, B., & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, 2(01), 20-28. https://doi.org/10.38094/jastt20165

Chaudhary, M. (2020). *TF-IDF Vectorizer scikit-learn*. https://medium.com/@cmukesh8688/tf-idf-vectorizer-scikit-learn-dbc0244a911a

Chotisarn, N., Siriphongwatana, P., Thanisuwiphat, P., Gulyanon, S., & Nadee, W. (2022). *Finding customer behavior insights for content creation in material and product sourcing using specialized topic analysis*. Paper presented at The International Conference on Electronic Business, Bangkok, Thailand. https://aisel.aisnet.org/iceb2022/17

depa. (n.d.). *Thailand digital economy glance*. https://www.depa.or.th/th/article-view/thailand-digital-economy-glance

Dharmaraj, S. (2024). *Thailand's digital boom: Economy, cybersecurity, and AI on the rise*. https://archive.opengovasia.com/2024/10/15/thailands-digital-boom-economy-cybersecurity-and-ai-on-the-rise/?c=kr

Gopinath, K., & Appavu alias Balamurugan, S. (2023, December 15-16). Human resource analytics: Leveraging machine learning for employee attrition prediction. In S. K. Sharma, Y. K. Dwivedi, B. Metri, B. Lal, & A. Elbanna (Eds.), *Transfer, Diffusion and Adoption of Next-Generation Digital Technologies. TDIT 2023. IFIP Advances in Information and Communication Technology* (pp. 137-158). Springer, Cham. https://doi.org/10.1007/978-3-031-50188-3_13

Kanade, V. (2022). *What is logistic regression? Equation, assumptions, types, and best practices*. https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-logistic-regression/

Khalaf, C., Michaud, G., & Jolley, G. J. (2023). Predicting declining and growing occupations using supervised machine learning. *Journal of Computational Social Science*, 6(2), 757-780. https://doi.org/10.1007/s42001-023-00211-0

Kundur, N. C., & Mallikarjuna, P. B. (2022). Deep convolutional neural network architecture for plant seedling classification. *Engineering, Technology & Applied Science Research*, 12(6), 9464-9470. https://doi.org/10.48084/etasr.5282

Lee, Y. M., Bacchi, S., Sia, D., Casson, R. J., & Chan, W. (2023). Optimising vitrectomy operation note coding with machine learning. *Clinical & Experimental Ophthalmology*, 51(6), 577-584. https://doi.org/10.1111/ceo.14257

Manorom, P., Detthamrong, U., & Chansanam, W. (2024). Comparative assessment of fraudulent financial transactions using the machine learning algorithms decision tree, logistic regression, Naïve Bayes, k-nearest neighbor, and random forest. *Engineering, Technology & Applied Science Research*, 14(4), 15676-15680. https://doi.org/10.48084/etasr.7774

Monika, P., Devi Sri, D., & Suneetha, M. (2024). Uncovering insights in agroforestry: A text mining analysis using web scraping and topic modeling. In B. Rama Devi, K. Kumar, M. Raju, K. Srujan Raju, & M. Sellathurai (Eds.), *Proceedings of Fifth International Conference on Computer and Communication Technologies* (pp. 95-104). Springer, Singapore. https://doi.org/10.1007/978-981-99-9704-6_8

Ninthanom, P. (2021). Performance comparison of transformation methods in data mining classification technique. *Thai Journal of Science and Technology*, 10(1), 14-25. https://doi.org/10.14456/tjst.2021.2

Okoye, K., Nganji, J. T., Escamilla, J., & Hosseini, S. (2024). Machine learning model (RG-DMML) and ensemble algorithm for prediction of students' retention and graduation in education. *Computers and Education: Artificial Intelligence*, 6, 100205. https://doi.org/10.1016/j.caeai.2024.100205

Palee, P., Wannapiroon, P., & Nilsook, P. (2020). The architecture of intelligent career prediction system based on the cognitive technology for producing graduates to the digital manpower. *International Journal of Advanced Computer Science and Applications*, 11(12), 115-121. https://dx.doi.org/10.14569/IJACSA.2020.0111214

Percia David, D., Anton Moreno, S., Maréchal, L., Maillart, T., & Mermoud, A. (2023). Forecasting labor needs for digitalization: A bi-partite graph machine learning approach. *World Patent Information*, 73, 102193. https://doi.org/10.1016/j.wpi.2023.102193

Puknilchai, P. (2014). A model for forecasting the appropriate manpower for the workload: A case study of the branches of CP All Public Company Limited. *Panyapiwat Journal*, 5(2), 103-119. https://so05.tci-thaijo.org/index.php/pimjournal/article/view/20129

PyThaiNLP. (n.d. -a). *pythainlp.tokenize*. https://pythainlp.org/dev-docs/api/tokenize.html

PyThaiNLP. (n.d. -b). *pythainlp.corpus*. https://pythainlp.org/dev-docs/api/corpus.html

Rajath, J., & Pandita, D. (2022, October 25-26). Machine learning as a data science tool to predict absenteeism for factory workers. *2022 International Conference on Data Analytics for Business and Industry (ICDABI)* (pp. 261-265). IEEE. https://doi.org/10.1109/ICDABI56818.2022.10041623

Ray, S. (2025). *Naive Bayes classifier explained: Applications and practice problems of Naive Bayes classifier*. https://www.analyticsvidhya.com/blog/2017/09/naive-bayes-explained

Rubio-Martín, S., García-Ordás, M. T., Bayón-Gutiérrez, M., Prieto-Fernández, N., & Benítez-Andrades, J. A. (2024). Enhancing ASD detection accuracy: A combined approach of machine learning and deep learning models with natural language processing. *Health Information Science and Systems*, 12(1), 20. https://doi.org/10.1007/s13755-024-00281-y

SCB 10X. (2023). *Find out more about natural language processing (NLP), a tool that enables computers to comprehend language*. https://www.scb10x.com/en/blog/natural-language-processing

Tadla, V. S., Singh, P. M., Thakkar, K. M., & Adatkar, R. (2023, December 8-9). Campus placement using machine learning: An extensive review and comparative study of machine learning methods. In S. K. Ukarande (Ed.), *2023 6th International Conference on Advances in Science and Technology (ICAST)* (pp. 427-430). IEEE. https://doi.org/10.1109/ICAST59062.2023.10455050

Thanapatpisarn, P. (2022). *Decision tree: A tree model with the complexity of decision making!!! Part 1*. https://datascihaeng.medium.com/decision-tree-part01-47ef24539fba

TheReporterAsia. (2024). *depa - IMC reveals Thai digital industry growth of 3.88% in 2023*. https://thereporter.asia/eng/2024/09/depa-imc-thai-digital-industry/

Wendland, A., Zenere, M., & Niemann, J. (2021). Introduction to text classification: Impact of stemming and comparing TF-IDF and count vectorization as feature extraction technique. In: M. Yilmaz, P. Clarke, R. Messnarz, M. Reiner (Eds.), *Systems, Software and Services Process Improvement* (pp. 289-300). Springer, Cham. https://doi.org/10.1007/978-3-030-85521-5_19

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

RESEARCH PAPER
J Inf Sci Theory Pract 13(3): 17-35, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.3.2

# The Roles of Information Professionals in Community Engagement for Developing a Learning Community in Longan Farming

**Chanchira Laorach**
Social Innovation Program, Faculty of Liberal Arts, Maejo University, Chiang Mai, Thailand
E-mail: chanchira.l@gmaejo.mju.ac.th

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

## ABSTRACT

This study employed a convergent mixed-methods approach to examine the development of learning communities among longan farmers in Chiang Mai Province, Thailand, and the role of information professionals in community engagement. Quantitative and qualitative data were collected simultaneously and synthesized to provide a comprehensive understanding. The qualitative phase involved structured interviews with ten longan farmers, with six participating, and data were analyzed descriptively to assess learning community development. In the quantitative phase, surveys were distributed to 144 stakeholders, with 136 completed responses (94.44%). Descriptive and inferential statistical analyses were conducted using IBM SPSS Statistics 30.0 (IBM Co., Armonk, NY, USA), including the Shapiro-Wilk test, Friedman test, and Wilcoxon Signed-Rank test, to evaluate the significance of community engagement in agricultural development. Findings from the qualitative analysis highlight three key aspects of learning center operations: (1) management, (2) farmer skills development, and (3) knowledge creation and transfer. The quantitative results identify four areas in which information professionals contribute to community engagement: (1) Outreach (Information Providing), (2) Collaboration (Information Evaluation), (3) Involvement (Information Analysis), and (4) Consultation (Information Consulting). Qualitative insights complement the statistical findings, providing a deeper understanding of the learning community development process. This integration clarifies the role of information professionals and supports policy development initiatives aimed at strengthening the agricultural capacity of longan farmers, with potential applicability to other agricultural communities. Furthermore, the study offers a framework for information professionals, librarians, and agricultural specialists to enhance learning communities and contributes to the advancement of curricula in library and information science.

**Keywords:** community engagement, information professional, learning community, longan farming, community management

**Received:** October 14, 2024 **Revised:** April 10, 2025
**Accepted:** April 28, 2025 **Published:** September 30, 2025

***Corresponding Author:** Suwannee Hoaihongthong
https://orcid.org/0000-0002-0554-5945
**E-mail:** suwaho@kku.ac.th

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 (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

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

## 1. INTRODUCTION

Thailand is currently prioritizing the development of its agricultural sector through an integrated approach designed to achieve balanced progress in the economic, social, and environmental dimensions at the household, community, and national levels, with a particular focus on promoting self-reliance (Ministry of Agricultural and Cooperatives, 2023). Strengthening the agricultural production base necessitates adaptation through the establishment of a robust and efficient management system. This process includes enhancing farmers' knowledge and expertise across all facets of production and management while maintaining a balance between economic, social, environmental, and ecological considerations (Ministry of Higher Education, Science, Research and Innovation, 2021). These initiatives align with the broader objective of advancing national development toward sustainability, in accordance with Sustainable Development Goals (SDGs). The SDGs provide an international framework that advocates for an improved quality of life for both farmers and consumers while fostering sustainable interactions with the environment (Ministry of Agricultural and Cooperatives, 2023; Ministry of Higher Education, Science, Research and Innovation, 2021).

Chiang Mai Province encompasses agricultural areas across all 25 districts. According to the Trade Policy and Strategy Office (2023), Thailand ranks among the world's leading exporters of longan, with Chiang Mai being a key region for its production. The processing and export of longan—whether fresh, dried, frozen, or canned—are thus critical to both the province's agricultural and processing sectors. To bolster the agricultural sector, Chiang Mai has established the Agricultural Learning Center (ALCs), which functions as a hub for transferring knowledge on production technology, management, and marketing to farmers. Additionally, the center provides agricultural services, disseminates valuable information, and integrates efforts from various agencies to address agricultural challenges and promote regional development. Presently, there are 25 primary ALCs, categorized by agricultural product, with five centers specifically dedicated to longan (Chiang Mai Agricultural Extension Office, n.d.).

Chiang Mai Province has established a network of agricultural communities, with government agencies acting as key facilitators. This initiative embodies the characteristics of a learning community, where groups of people come together to exchange and share knowledge, skills, and experiences in agriculture with a common goal. However, achieving sustainable development within these agricultural communities may be unattainable through government support alone. For any community to progress, it is essential that the members themselves take an active role in addressing local challenges and identifying their most urgent needs. Consequently, the development of a learning community necessitates fostering critical thinking among residents and encouraging their participation in all stages of development. This participatory approach equips individuals with the skills needed for critical thinking, analysis, problem-solving, and the application of knowledge, contributing not only to their personal development but also to the long-term growth of the community (Nobnop et al., 2021).

The effectiveness of a learning community can be significantly enhanced through the involvement of information professionals. These specialists are proficient in acquiring, collecting, recording, organizing, storing, maintaining, searching, disseminating, and providing information services across various formats, tailored to specific contexts (Greer et al., 2013). Their role in community engagement is pivotal, requiring not only technical expertise but also the capacity to forge positive relationships with the community. Unlike traditional information services, community engagement encompasses strategic activities that necessitate information professionals to collaborate closely with the community toward clearly defined, beneficial objectives aimed at fostering community change (Caspe & Lopez, 2018; Winberry & Bishop, 2021).

The development of longan farming communities in Chiang Mai Province highlights critical challenges faced globally by farmers in developing countries, such as accessing reliable information, fostering effective collaboration within communities, and sustaining knowledge development. These challenges are not unique to rural areas in Thailand but are also prevalent in agricultural communities across Southeast Asia, Africa, and Latin America. These regions often share similar characteristics, including a reliance on agriculture as a primary livelihood, limited access to government support resources, and a disparity between farmers' information needs and the availability of expertise from information professionals (Food and Agriculture Organization of the United Nations [FAO], 2017; Tumwebaze et al., 2025).

A review of the literature reveals that while some information professionals and library and information science specialists engage with communities, their efforts predominantly focus on providing information services within the framework of academic service projects. There

 https://doi.org/10.1633/JISTaP.2025.13.3.2

is a notable deficiency in research addressing community-oriented missions (Caspe & Lopez, 2018). Consequently, there is limited exploration of the roles and responsibilities of information professionals in community engagement, particularly within the agricultural sector. This research article aims to examine the development status of learning communities among longan farmers and the role of community engagement of information professionals in developing a learning community for longan farming in Chiang Mai Province. The goal is to offer guidelines for the work of information professionals, librarians, agricultural academics, and officials involved in developing agricultural learning communities. Particularly, the framework adopted here offers insights that extend beyond the local context and can contribute to understanding knowledge management and community engagement in broader settings. Specifically, the findings have implications for agricultural community development in other developing countries, where holistic approaches emphasizing collaboration and capacity-building are essential for sustainable growth (World Bank, n.d.).

## 2. LITERATURE REVIEW

Information professionals, including librarians and information officers, hold positions across governmental and private sectors, specialized agencies, and independent entities (Greer et al., 2013). In the job market, these professionals are predominantly employed in libraries, information service centers, learning centers, archives, and museums (Mancini, 2012). Traditionally, information professionals are equipped with the knowledge and skills necessary for procuring, collecting, recording, organizing, storing, maintaining, searching, disseminating, and providing information services across various formats, in alignment with relevant objectives and contexts. However, they may lack the specialized knowledge and skills needed to address social issues, engage in action, build trust, make decisions, or solve problems—tasks that often require professionals with specific expertise in these areas. Consequently, information professionals frequently collaborate with others to perform information management duties, which may include communication efforts to enhance understanding within the community (Clinical and Translational Science Awards Consortium et al., 2011).

When engaged in community outreach, information professionals may apply their expertise in information science to support community engagement efforts. The conceptual framework for community engagement typically incorporates theories or practices utilized by academics and practitioners involved in community projects with the aim of effecting community change. Despite this, a review of the literature indicates that there is no specific conceptual framework tailored explicitly for community engagement work. In this context, the authors employ the Community Engagement Continuum model developed by the US National Institutes of Health. This model delineates five distinct work styles for community engagement practitioners: outreach, consulting, involvement, collaboration, and shared leadership. Each style specifies the characteristics of community relations, communication and decision-making processes, levels of participation, the positioning of entities, and anticipated outcomes (Clinical and Translational Science Awards Consortium et al., 2011).

Through an analysis of the general roles of information professionals in relation to the community engagement model, as defined by the concept of community engagement, it is evident that information professionals are responsible for sourcing, collecting, recording, organizing, storing, searching, disseminating, and providing information services (Caspe & Lopez, 2018). However, they may lack the in-depth knowledge and skills necessary to address social issues comprehensively, which are essential for participating in action, building trust, and engaging in decision-making processes aimed at solving various problems. Such tasks often require the expertise of academics and professionals with specialized knowledge pertinent to the issue at hand. Consequently, information professionals must collaborate with other stakeholders to manage information-related duties and may need to engage in communication efforts to enhance understanding within the community (Winberry & Bishop, 2021). In community engagement missions, information professionals can leverage their expertise in information science across four distinct types of community engagement activities: (1) Outreach, by serving as information providers; (2) Consulting, by acting as information consultants; (3) Involvement, by functioning as information analysts; and (4) Collaboration, by serving as information evaluators (Clinical and Translational Science Awards Consortium et al., 2011). It is important to recognize that a single project or activity may encompass multiple roles. All these activities should have well-defined objectives, align with the community's problems and needs, and contribute to positive change within the community.

The literature review reveals that community participation in agricultural projects or activities is generally lim-

ited to basic levels, such as providing information, offering consultations, or participating in activities organized by external agencies. This participation may take the form of ad hoc participation or involvement in formal, structured processes. However, there has been little in-depth research on higher levels of community participation, such as joint decision-making, involvement in design and implementation, or joint learning. Participation at these higher levels not only enhances farmers' knowledge and skills but also helps build confidence and trust in the sustainable development process within the community. The role of information professionals, who can contribute beyond simply providing information or consultations by fostering higher levels of participation and facilitating more effective decision-making, represents a research gap that has not been thoroughly explored in the context of longan agriculture in Thailand.

In recent years, numerous studies have underscored the role of information professionals in school and public libraries in facilitating access to agricultural information, particularly by providing services to farmers in rural areas. Such information is critical for the advancement of agricultural professions. Research has shown that information professionals play a pivotal role in organizing and disseminating data pertaining to agricultural techniques, such as crop cultivation, pest and disease management, and market-related information, all of which farmers can utilize to make informed decisions and optimize agricultural practices. The integration of agricultural information with the needs of farmers is considered a fundamental strategy for promoting the sustainable development of agriculture, as accurate and timely data empowers farmers to adopt effective methods for improving crop yield and quality. Furthermore, it enables them to make strategic decisions regarding investments and manage potential risks arising from external factors, such as natural disasters or market fluctuations. In addition, community libraries play an essential role in enhancing access to agricultural knowledge by organizing training programs, offering consultations, and fostering learning networks within the community. This not only ensures farmers have access to valuable information but also allows them to exchange experiences and best practices for improving agricultural activities in their respective regions (Jarolímek et al., 2024; Singh et al., 2022; Yusuf et al., 2019).

Based on a review of relevant research, the role of information professionals in supporting agriculture across different regions globally is reflected through various activities. In Southeast Asia, such as in Vietnam, librarians and information professionals have developed digital platforms, such as databases, that provide farmers with access to climate information and agricultural technologies. The Smart Farmer Libraries project in Vietnam serves as a knowledge hub, enabling rural farmers to learn and apply agricultural practices (Catelo et al., 2016; Tuan, 2024). In Africa, librarians work in resource-constrained environments, providing agricultural information through community radio in countries such as Kenya and Uganda, making it easily accessible and relevant to local needs. Additionally, they promote access to information for female farmers in natural resource management to reduce gender disparities (Buehren, 2023; Ng'eno & Mutula, 2022). In Europe, librarians utilize advanced technologies such as Big Data and artificial intelligence (AI) to analyze agricultural data. For example, International System for Agricultural Science and Technology (AGRIS), a global agricultural research database, enables farmers to access essential research data quickly. Community libraries in Spain and Italy organize events that foster knowledge exchange on sustainable agriculture (FAO, 2022). In the Americas, Extension Services programs and the Ag Data Commons platform support smallholder farmers' decision-making by providing information on crop management and marketing. Additionally, in Brazil and Mexico, libraries serve as intermediaries, connecting farmers with researchers to support the development and adoption of agricultural technologies (U.S. Department of Agriculture, n.d.).

The concept of lifelong learning has been officially recognized since approximately 1960, with United Nations Educational, Scientific and Cultural Organization (UNESCO) playing a significant role in its advocacy. Numerous global conferences have been held to discuss lifelong learning, which has led to its establishment as a fundamental principle in the educational policies of various countries around the world. Thailand first adopted the notion of lifelong learning officially in its National Education Plan of 1977 (Office of the National Education Commission, 1977). By the late twentieth century, just before entering the twenty-first century, lifelong learning had become a widely adopted approach to educational reform across many nations. This development responded to the rapid and complex changes in global society affecting economic, social, environmental, and communication aspects during the information age, which significantly impacted citizens' lives. Consequently, countries have sought ways to ensure their citizens can engage in continuous learning to adapt to these changes (Akther, 2020).

These events led to the emergence of learning commu-

 https://doi.org/10.1633/JISTaP.2025.13.3.2

nities, which are defined as groups or communities that recognize the importance of learning and actively support the creation, sharing, and application of knowledge for the community's benefit. These communities prioritize local needs and contexts and employ participatory processes that encourage members to think critically and utilize knowledge to solve problems. They promote continuous development of diverse skills in an environment conducive to various learning opportunities that align with individual capacities. When community members possess knowledge and skills, they can enhance their communities' strength and sustainability (Nobnop et al., 2021).

The development of rural learning communities has gained global attention, emphasizing the role of information professionals in enhancing knowledge sharing and fostering sustainable practices. Community informatics, for example, explores how technology and information systems bridge knowledge access gaps in rural areas (Sindakis & Showkat, 2024), with professionals facilitating communication and resource-sharing. This aligns with the concept of learning communities, where collaboration and continuous learning address local challenges (UNESCO International Research and Training Centre for Rural Education, 2014). Rural knowledge management further highlights strategies for capturing and disseminating local knowledge to improve agricultural productivity and livelihoods (Akuku et al., 2021).

Studies in countries such as India and Kenya demonstrate the pivotal role of information professionals in bridging knowledge gaps within farming communities, offering insights relevant to similar efforts in Thailand. Collaborative learning, as emphasized in global research, empowers communities to share knowledge and best practices, promoting sustainable development (Siregar et al., 2024). Applying these frameworks to the Thai context underscores the significance of information professionals in supporting longan farmers, bridging knowledge gaps, and fostering sustainable agricultural practices.

## 3. RESEARCH METHODOLOGY

This study utilized a convergent mixed-methods approach to examine the role of information professionals in promoting the development of learning communities among longan farmers in Chiang Mai Province. Both qualitative and quantitative research methods were employed to collect data simultaneously, synthesize and analyze both phases, and present holistic findings on the role of information professionals in community engagement (Joungtrakul & Ferry, 2020).

### 3.1. Phase 1: Qualitative Research

Qualitative research designed to investigate the status of the development of learning communities among longan farmers in Chiang Mai Province encompasses the following aspects:

#### *3.1.1. Population and Sample*

The initial sample included ten participants: five heads of ALCs specializing in longan farming, with one representative from each center, and five longan farmers who were members of these centers. However, during data collection, four individuals declined to participate, leaving six respondents who contributed to this study. The participant selection process employed purposive sampling to ensure alignment with the study's objectives.

The sample consisted of two main groups: (1) model farmers, who served as heads of ALCs, with one head selected from each of the five learning centers located in different districts of Chiang Mai, and (2) member farmers, who were affiliated with these learning centers. Five member farmers were initially selected based on recommendations from the center heads, ensuring that they had relevant knowledge and experience in agricultural learning and community development. Participants were required to meet specific inclusion criteria: They had to have served as an ALCs head or member for at least one year, be between 18 and 65 years old, be Thai nationals with valid identification, be in good health, have the ability to communicate fluently in Thai, and be willing and available to participate in interviews.

Interviews were conducted at the ALCs in Chiang Mai Province, a key region for longan farming in Thailand. Conducting interviews in this familiar and relevant setting facilitated authentic and context-specific responses. The discussions focused on participants' direct experiences within ALCs, including the operational characteristics of the learning centers, methods of information exchange and sharing, access to agricultural knowledge, and the processes of information development, storage, dissemination, and curriculum design. In-depth interviews were conducted using open-ended questions, allowing participants to express their perspectives freely. Audio recordings and written notes were taken during the interviews to ensure accuracy and completeness in capturing the knowledge development processes within the agricultural community.

#### 3.1.2. Structured Interview Design

The researchers employed structured interviews as the primary data collection tool, which were developed based on a comprehensive review of principles, concepts, theories, academic literature, and related research. The interview questions were designed to address key areas, including the operational characteristics of learning centers, management structures, and the roles of personnel, as well as methods of information exchange and practices of information sharing. Furthermore, the questions explored access to information, considering both enabling factors and constraints, and examined the processes of information development, storage, and dissemination. The design also encompassed the creation of learning curricula tailored to the needs of users. To ensure the questions aligned with the research objectives and maintained clarity and relevance, the instruments were reviewed and validated by experts in the relevant fields.

#### 3.1.3. Data Analysis

The qualitative data in this study were analyzed using thematic analysis, following the six-phase framework proposed by Braun and Clarke (2006). The analysis was conducted manually without the aid of qualitative data analysis software to allow for a more nuanced and context-sensitive interpretation. First, the interview recordings were transcribed verbatim. The researchers engaged in repeated readings of the transcripts to immerse themselves in the data and gain a comprehensive understanding of its depth and breadth. Field notes and reflective memos were incorporated to supplement the transcripts and enrich contextual interpretation.

Data coding was conducted line-by-line using an inductive approach, allowing codes to emerge naturally without relying on a predefined framework. The initial coding was performed manually to ensure that both semantic (explicit) and latent (underlying) patterns were captured. The first researcher conducted the initial coding, and a peer reviewer independently checked the codes to enhance credibility; discrepancies were discussed collaboratively. Significant features of the data were systematically coded across the entire dataset, and similar codes were organized into potential themes that aligned with the research objectives.

Potential themes were iteratively reviewed and refined for coherence and relevance. Themes were merged, separated, or discarded as necessary, and the thematic structure was checked against both the coded extracts and the complete dataset to ensure internal consistency and validity. Once the final themes were established, each theme was clearly defined and named to reflect its core meaning and contribution to the research questions. Detailed descriptions were developed to articulate the essence of each theme.

Throughout the analytic process, strategies such as incorporating field notes and reflective memos, maintaining reflexivity, and monitoring data saturation were employed to enhance the trustworthiness of the findings. Data saturation was considered achieved when no new themes or insights emerged from successive interviews. Diverse participant perspectives, including divergent or minority views, were intentionally preserved and reported to reflect the complexity of participants' experiences.

### 3.2. Phase 2: Quantitative Research

Quantitative research designed to investigate the role of community engagement initiatives undertaken by information professionals in developing learning communities for longan farming in Chiang Mai Province encompasses the following aspects:

#### 3.2.1. Population and Sample

The population in this study comprised 16 librarians or information officers working in public libraries under the Chiang Mai Provincial Learning Promotion Office, 22 school librarians under the Chiang Mai Secondary Education Service Area Office and the Office of the Basic Education Commission, and 106 agricultural academics from the Chiang Mai Agricultural Extension Office involved in promoting and developing agricultural learning

**Table 1.** Populations and samples

| Position | Affiliated organization | Population | Sample |
|---|---|---|---|
| Librarians/information officers | Public libraries | 25 | 16 |
| Librarians/information officers | School libraries | 34 | 22 |
| Agricultural extension officers | Agricultural extension office | 166 | 106 |
| Total | | 225 | 144 |

 https://doi.org/10.1633/JISTaP.2025.13.3.2

communities, as detailed in Table 1. The inclusion criteria required participants to hold positions as librarians, information officers, or agricultural extension officers. Eligible participants were selected based on the following criteria: They could be of either gender, aged between 18 and 65 years, had held their positions for at least one year, demonstrated proficient communication skills in Thai, and expressed willingness to dedicate time to completing the questionnaire.

The sample size was determined using Yamane (1973)'s formula, with a margin of error set at 5%. A stratified sampling method was subsequently employed to ensure proportional representation of each subgroup within the population. The population was categorized into three strata based on professional roles: (1) public library librarians or information officers, (2) school librarians, and (3) agricultural academics. Proportional sampling was conducted within each stratum to maintain consistency with their representation in the total population. This approach ensured that the sample accurately reflected the diversity of the population while upholding statistical validity and minimizing selection bias.

### *3.2.2. Questionnaire Design and Validation*

The researcher analyzed literature and research related to the role of community engagement to gain clear insights that would guide the development of the questionnaire. The research instrument is divided into two parts: general information about the respondents, and information regarding the role of community engagement by information professionals in developing learning communities. Specifically, the questionnaire covers: 1) Outreach, 2) Consultation, 3) Involvement, and 4) Collaboration. Subsequently, the quality and validity of the instrument were assessed using an Index of Item-Objective Congruence evaluation form completed by three experts. Data were then collected from a pilot group of 30 individuals who were not part of the sample population. The instrument's reliability was calculated using Cronbach's alpha coefficient, following Cronbach (1990)'s method, which specifies that reliability should not fall below 0.7. In this study, the instrument demonstrated a reliability score of 0.93. After confirming the instrument's quality, revisions were made to enhance the clarity, conciseness, and overall precision of the questions.

### *3.2.3. Data Collection*

Data were collected using an online questionnaire from individuals involved in the development of the learning community. This included 16 librarians/information officers from public libraries, all of whom returned the questionnaires, 22 librarians/information officers from school libraries, of which 20 returned the questionnaires, and 106 agricultural extension officers from the Chiang Mai Provincial Agricultural Office, with 100 returning the questionnaires. This resulted in a total of 136 respondents, representing a response rate of 94.44 percent.

### *3.2.4. Data Analysis*

The data were analyzed using statistical evaluation and descriptive analysis with IBM SPSS Statistics 30.0 (IBM Co., Armonk, NY, USA). The statistics employed for data analysis included the mean ($\chi$) and standard deviation (S.D.) to interpret opinions regarding the role of community engagement from the sample group. The mean was calculated for each community engagement role (Outreach, Collaboratives, Involves, Consults) to assess the significance of each role in developing the learning community for longan farmers in Chiang Mai.

To interpret the opinions based on the Likert (1932) scale, the following criteria were applied:

4.51-5.00: The most significant role.
3.51-4.50: Very significant role.
2.51-3.50: Moderately significant role.
1.51-2.50: Lesser significant role.
1.00-1.50: Least significant role.

To assess the distribution of the data, the Shapiro-Wilk test was used to check whether the data followed a normal distribution (Shapiro & Wilk, 1965). The results of this test guided the choice of appropriate statistical analysis methods. If the data were found not to be normally distributed, non-parametric methods, such as the Friedman test, would be used instead.

Since the data collected were measurements taken multiple times from the same sample group and did not follow a normal distribution, the Friedman test was used to compare the differences between the four roles of community engagement (Friedman, 1937). The results of this test helped determine whether there were statistically significant differences between the roles.

After conducting the Friedman test to compare the roles, post-hoc analysis (pairwise comparison) was performed using the Wilcoxon Signed-Rank test (Wilcoxon, 1945). The significance of the differences was evaluated by adjusting the $p$-value using Bonferroni correction (Bonferroni, 1936) to control for Type I error due to multiple

comparisons.

## 4. RESULTS

### 4.1. Qualitative Research Results on the Development Status of Learning Communities among Longan Farmers in Chiang Mai Province

#### *4.1.1. Operation of the ALCs*

The ALCs is a project overseen by the Ministry of Agriculture and Cooperatives. The process of developing farmers' potential begins with the Department of Agricultural Extension in each province, which conducts an analysis of the learning center's capabilities to ensure alignment with the local context. Following this, a development plan and corresponding guidelines are formulated in collaboration with the District Agricultural Office. Subsequently, an action plan is created and disseminated to all main centers and network centers.

The primary objective of the learning curriculum for farmers is to enhance their potential by equipping them with the knowledge and skills necessary for sustainable and efficient agricultural development within their communities. Designed to integrate with the Bio-Circular-Green Economy Model Agritech and Innovation Center, the curriculum enables farmers to apply appropriate knowledge and technology suited to their local contexts. It is divided into four main components: The Core Curriculum covers fundamental agricultural topics such as zoning, cost reduction, productivity enhancement, marketing management, product processing, and pest management; the Compulsory Curriculum addresses essential concepts like the sufficiency economy, new theory agriculture, integrated farming, and household accounting; the Supplementary Curriculum expands knowledge in areas such as water management, community enterprise development, business plan preparation, and supplementary careers to boost income; and the Advanced Farmer Development Curriculum focuses on enhancing business management, marketing, networking, and modern agricultural technology skills, including smart water systems and solar energy, empowering farmers to apply and disseminate this knowledge to others.

Furthermore, there is a focus on developing learning bases that serve as centers for delivering curriculum-aligned content and modern technologies. These centers are designed to showcase various innovations that farmers can apply in practice. Knowledge is gathered and cultivated at these learning bases, with integration from various agencies to bring in external resources and expertise. The ultimate goal is to maximize the benefits for farmers, including enhancing the agricultural plots of model farmers to facilitate the transfer of agricultural knowledge to interested parties. The initiative aims to teach cost-reduction methods and improve agricultural production efficiency in line with the principles of the sufficiency economy. Additionally, there is an emphasis on improving planting techniques and integrating new technologies, including the installation of educational signage on learning plots to ensure clear understanding and assist farmers in adapting to changing economic and environmental conditions.

From interviews with longan farmers in the area, it was found that most agricultural learning communities are currently at a medium to advanced level. Approximately 60-70% of farmers consistently participate in the center's activities. Farmers who are members of the center's network stated that "participating in activities helps us gain new and up-to-date knowledge that can be applied to our agricultural plots." However, 30-40% of farmers do not participate regularly, citing concerns that "some training does not address the specific problems in our area, and there is no follow-up after the training." Additionally, some farmers reported limitations in accessing resources, such as a lack of appropriate technological equipment or delays in support from relevant agencies.

These findings highlight gaps in the development of learning centers and training curricula that align more closely with local needs. Designing curricula that address specific regional challenges and establishing long-term support networks can enhance farmers' access to relevant knowledge and resources in a sustainable manner. Furthermore, the development and expansion of digital learning platforms, such as online training modules, could help mitigate disparities in access to agricultural information and training. Additionally, sustained support from relevant agencies, along with systematic follow-up on training outcomes, remains essential for ensuring that learning centers effectively contribute to the long-term capacity building of farmers.

#### *4.1.2. Skill Development of Longan Farmers*

The ALCs have consistently organized training programs aimed at enhancing farmers' knowledge and skills in longan farming, with an emphasis on promoting large-scale farming. The target groups include leading farmers, new farmers, members of large farming plots, and those preparing to join such plots. Following this, an analysis of the area is conducted to identify agricultural issues that

 https://doi.org/10.1633/JISTaP.2025.13.3.2

require development, which are then selected as topics for two training sessions based on the planned curriculum. Additionally, a study visit focusing on relevant technologies and innovations that align with the curriculum is conducted. The training takes place at the Agricultural Extension Office, main centers, or network centers. Upon completion of the learning program, the responsible agencies record farmers' learning outcomes to inform improvements for future training sessions.

In addition, for enhancing the skills of the heads of the ALCs specializing in longan farming, a training session was conducted, addressing key topics such as the production of safe agricultural products, the application of appropriate agricultural innovations and technologies, and the utilization of online learning platforms. Additionally, the training emphasized the connection between knowledge from the ALCs and local farmers. The Agricultural Office retains the flexibility to modify the training format or study tours to accommodate the specific conditions of each area.

ALCs therefore play a crucial role in enhancing farmers' knowledge, with notable success stories demonstrating their impact. For instance, a Learning Center organized a training program on "Water Management for Agriculture," enabling farmers to reduce their water usage by 30% during the dry season. Farmers in the area reported, "After the training on efficient water use, our orchards no longer experience water shortages during the dry season, and we have also been able to significantly reduce costs."

Similarly, a Learning Center implemented a project on "Longan Value Addition," promoting the production of high-quality dried longans. This initiative led to a 15% increase in the average household income of participating farmers. One farmer noted, "Learning about dried longan production has allowed us to sell our longans at a better price. Our income has actually increased, and now we have more money in reserve."

These cases highlight the potential of ALCs to foster sustainable agricultural development and strengthen farming communities. By promoting agricultural innovations tailored to local contexts, such as "Water Management for Agriculture" and "Additional Career Development" programs, ALCs have effectively addressed specific challenges. In areas facing water shortages, these initiatives have significantly reduced production costs and improved efficiency. Some farmers shared their experiences, stating, "After we joined the training, we learned how to use water efficiently and could immediately apply it in our longan orchards," and "Now we have more ways to add value to our products and can sell them at better prices than before."

#### *4.1.3. Creation and Transfer of Knowledge*

The ALCs focus on knowledge generation through systematic analysis of agricultural situations, working with relevant agencies, committees, and network centers. This process covers identifying key agricultural problems in each area and their urgent needs, leading to the development of development plans that are consistent with the identified key problems. In the knowledge collection process, the center focuses on recording detailed data from the analysis and selection of key problems or situations in the area. It also collects data on agricultural innovations for dissemination. This data plays an important role in effective agricultural decision-making and strategy making.

For knowledge transfer, various activities focus on disseminating agricultural innovations, especially in the longan category, such as organizing exhibitions in collaboration with educational institutions, government agencies, and the private sector. These activities help promote knowledge exchange among stakeholders. At the end of the knowledge application, it was found that farmers were able to apply the innovations and knowledge they received to their cultivation and agricultural management effectively.

Last year, approximately 250 longan farmers participated in the project, with 70% of participants being able to effectively apply the innovations they learned. One example is the use of smart water systems that help reduce water usage and production costs. "The smart water system can reduce water costs by up to 30%, freeing up budgets to improve other agricultural areas," some farmers said. Activities supported by the farmer network, such as exhibitions and community markets, also help to promote knowledge exchange and increase access to new markets. "The network helps us learn new techniques, such as using organic fertilizers that are suitable for our local soil, and also allows us to sell our products to a wider market, especially at the annual exhibition," said one farmer.

In addition, farmers from resource-poor rural areas stated that "Although we started with difficulty accessing information, these activities have given us new directions and hopes for community development." The results also indicated that farmer networks and community enterprise groups play an important role in expanding knowledge through activities such as training, exchanging experiences, and developing community products. Farmers who joined community enterprise groups stated that "Groups

help us share problems and find solutions together, giving us more opportunities to develop products and increase income than working alone."

### 4.2. Quantitative Research Results on the Role of Community Participation by Information Professionals in Developing a Learning Community for Longan Farming in Chiang Mai Province

According to the results presented in Table 2, the overall analysis of the data indicates that the roles associated with community engagement undertaken by information professionals significantly contribute to the successful development of learning communities among longan farmers in Chiang Mai Province, with a high success rate ($\overline{\chi}$=3.83). Among these roles, outreach activities involving information provision and collaborative roles involving information evaluation are identified as the most impactful in fostering the development of learning communities, each rated at a high level ($\overline{\chi}$=3.88). Subsequent roles include involvement through information analysis ($\overline{\chi}$=3.80) and consulting through information consultation ($\overline{\chi}$=3.77).

When evaluating the outreach aspect of information provision and service delivery, it was observed that the role of information professionals in offering services that align with the needs of users significantly contributes to the successful development of learning communities, with a high level of effectiveness ($\overline{\chi}$=3.94). Following this, the provision of information services that are well-suited to community needs, the regular dissemination of up-to-date information for learning and career advancement ($\overline{\chi}$=3.91), and the preparation of documents for community dissemination to facilitate mutual learning ($\overline{\chi}$=3.87) are also important roles (Table 3).

In the context of collaborative roles as information evaluators, it was found that the involvement of information professionals in partnering with farmers and community members to evaluate data analysis results significantly contributes to the successful development of learning communities, achieving a high level of effectiveness ($\overline{\chi}$=3.94). This is followed by collaboration with agricultural experts in evaluating data analysis results ($\overline{\chi}$=3.93), and cooperation in assessing the quality of information based on the accuracy and reliability of agricultural data sources ($\overline{\chi}$=3.87) (Table 4).

Furthermore, analysis of the consulting roles as performed by information consultants revealed that advising on the creation of knowledge content tailored to farmers is the most significant role for information professionals

**Table 2.** Community engagement roles of information professionals

| Community engagement roles of information professionals | $\overline{\chi}$ | S.D. |
|---|---|---|
| 1. Outreach (Information Providing) | 3.88 | 0.71 |
| 2. Collaboratives (Information Evaluations) | 3.88 | 0.78 |
| 3. Involves (Information Analysis) | 3.80 | 0.78 |
| 4. Consults (Information Consulting) | 3.77 | 0.72 |
| Total | 3.83 | 0.68 |

S.D., standard deviation.

**Table 3.** Outreach roles of information professionals

| Roles | $\overline{\chi}$ | S.D. |
|---|---|---|
| (1) Providing information services that meet needs | 3.94 | 0.78 |
| (2) Providing appropriate information services | 3.91 | 0.80 |
| (3) Preparing documents and disseminating them to the community for regular joint learning | 3.87 | 0.85 |
| (4) Disseminating up-to-date information for regular learning and career development | 3.91 | 0.83 |
| (5) Creating databases, websites, and information systems to collect, store, and disseminate information to the community | 3.78 | 0.82 |
| Total | 3.88 | 0.71 |

S.D., standard deviation.

 https://doi.org/10.1633/JISTaP.2025.13.3.2

**Table 4.** Collaborative roles of information professionals

| Roles | $\bar{\chi}$ | S.D. |
|---|---|---|
| (1) Collaborating in the planning of agricultural information quality assessments that are aligned with the specific context of the community | 3.82 | 0.84 |
| (2) Cooperating in the evaluation of information quality based on the accuracy and reliability of agricultural information sources | 3.87 | 0.85 |
| (3) Working with agricultural experts to assess the outcomes of data analysis | 3.93 | 0.83 |
| (4) Collaborating with farmers and community members in the evaluation of data analysis results | 3.94 | 0.87 |
| Total | 3.88 | 0.78 |

S.D., standard deviation.

**Table 5.** Consulting roles of information professionals

| Roles | $\bar{\chi}$ | S.D. |
|---|---|---|
| (1) Providing consultation on the planning and development of information resources, including needs assessments and the formulation of plans for the storage and retrieval of information essential for learning and development | 3.76 | 0.79 |
| (2) Offering consultation on the creation of knowledge content tailored to the needs of farmers | 3.84 | 0.92 |
| (3) Providing consultation on the exchange and sharing of knowledge that is relevant and beneficial for farmers | 3.79 | 0.88 |
| (4) Offering consultation on the use of information technology to support collaborative learning initiatives | 3.82 | 0.78 |
| (5) Providing consultation on the development and maintenance of online learning environments for communities, including websites, learning platforms, and online communities | 3.66 | 0.87 |
| Total | 3.77 | 0.72 |

S.D., standard deviation.

**Table 6.** Involvement roles of information professionals

| Roles | $\bar{\chi}$ | S.D. |
|---|---|---|
| (1) Engaging in data analysis to understand community needs and trends, in order to formulate suitable development guidelines and plans | 3.75 | 0.83 |
| (2) Presenting the community with guidelines for data analysis and comparisons that are both reliable and of high quality | 3.78 | 0.85 |
| (3) Collaborating with agricultural experts to analyze information that aligns with the specific needs of the community | 3.87 | 0.89 |
| (4) Contributing to the systematic and reliable presentation of data analysis results, highlighting both advantages and disadvantages | 3.81 | 0.88 |
| Total | 3.80 | 0.78 |

S.D., standard deviation.

in fostering the success of learning communities, with a high mean score ($\bar{\chi}$=3.84). This was followed by consulting on the use of information technology to support collaborative learning ($\bar{\chi}$=3.82) and advising on the exchange and sharing of knowledge suitable for farmers ($\bar{\chi}$=3.79), respectively (Table 5).

In examining the role of involvement as an information analyst in Table 6, it was determined that the most impactful contribution of information professionals to the development of successful learning communities was their collaboration with agricultural experts to analyze information that aligns with the specific needs of the community

($\overline{\chi}$=3.87). This was followed by their role in systematically and reliably presenting the results of information analysis, including advantages and disadvantages ($\overline{\chi}$=3.81), and by their provision of guidelines for analyzing and comparing quality and reliable information for the community ($\overline{\chi}$=3.78).

Based on analysis of the normality test results for the data from all four roles: 1) Outreach (Information Providing), 2) Collaboratives (Information Evaluation), 3) Involves (Information Analysis), and 4) Consults (Information Consulting), the Shapiro-Wilk test (Shapiro & Wilk, 1965), which is commonly used for testing normal distribution in cases with small to moderate sample sizes, was applied. The test results showed that the $p$-value<0.05 for all roles, indicating that the data does not follow a normal distribution. Therefore, a non-parametric analysis using the Friedman test was conducted to compare the importance of the four roles. The results revealed significant differences in the participants' views ($\chi^2$ (3)=Friedman value, $p$=0.024), indicating that the participants perceived some roles as more important than others in developing the learning community for longan farmers.

For the next step, the researcher conducted a post-hoc analysis using pairwise comparison with the Wilcoxon Signed-Rank test to compare all six pairs: 1) Outreach vs. Consult, 2) Outreach vs. Involve, 3) Outreach vs. Collaborative, 4) Consult vs. Involve, 5) Consult vs. Collaborative, and 6) Collaborative vs. Involve. The $p$-value was then adjusted using Bonferroni Correction to control for Type I Error, using the formula $p$-adjusted=$p$-original×number of comparisons (n). The $p$-values after adjustment and the original $p$-values from the pairwise comparisons are reported in Table 7.

The pairwise comparison analysis using the Wilcoxon Signed-Rank test, along with $p$-value adjustment using the Bonferroni Correction, revealed a statistically significant difference in Role Pair 1 ($p$-adjusted=0.006), whereas no significant differences were observed for Role Pairs 2 to 6 ($p$-adjusted>0.05). Adjusted $p$-values exceeding 1.0 were reported as 1.0, following standard statistical conventions to ensure clarity and accuracy in interpretation. This adjustment did not influence the conclusions derived from the analysis.

In summary, the study found that the Outreach role (Information Providing) and Collaborative role (Information Evaluation) had identical mean scores of 3.88, indicating that both roles were perceived as similarly important. However, the S.D. for the Collaborative role was higher (0.78) compared to Outreach (0.71), reflecting greater variability in perceptions of the Collaborative role. While some respondents recognized the long-term significance of collaboration, others may not currently perceive its importance, possibly due to the absence of effective collaborative efforts.

Conversely, the Outreach role received more consistent evaluations, reflecting a shared acknowledgment of its importance in rapidly disseminating actionable information. Such information can be immediately applied to address agricultural issues, such as maintaining longan trees or responding to natural disasters. These findings suggest that the Outreach role is the most critical in the context of developing learning communities for farmers in Chiang Mai, particularly when timely information is essential to address pressing issues. At the same time, the Collaborative role, which is crucial for long-term development, may not yet demonstrate its full potential due to the lack of effective collaboration between farmers and agricultural professionals in certain areas.

In conclusion, prioritizing the Outreach role represents an effective starting point for developing learning com-

**Table 7.** Pairwise comparisons

| No | Comparison pair | Original $p$-values | Adjusted $p$-values | Results |
|---|---|---|---|---|
| 1 | Outreach (Information Providing) vs. Consults (Information Consulting) | 0.001 | 0.006 | Significant |
| 2 | Outreach (Information Providing) vs. Involves (Information Analysis) | 0.028 | 0.168 | Not significant |
| 3 | Outreach (Information Providing) vs. Collaboratives (Information Evaluations) | 0.985 | 1.00 | Not significant |
| 4 | Consults (Information Consulting) vs. Involves (Information Analysis) | 0.574 | 1.00 | Not significant |
| 5 | Consults (Information Consulting) vs. Collaboratives (Information Evaluations) | 0.035 | 0.21 | Not significant |
| 6 | Involves (Information Analysis) vs. Collaboratives (Information Evaluations) | 0.051 | 0.306 | Not significant |

 https://doi.org/10.1633/JISTaP.2025.13.3.2

munities, such as through organizing training programs or promoting knowledge dissemination via accessible and rapid channels. Simultaneously, efforts to develop the Collaborative role are necessary to ensure long-term sustainability by fostering partnerships between farmers and agricultural professionals within the community.

Upon obtaining the research findings from both phases, namely the qualitative results on the current status of learning community development among longan farmers in Chiang Mai Province and the quantitative results regarding the role of community participation by information professionals in the development of a learning community for longan farming, the researcher subsequently integrated these findings to provide a comprehensive and in-depth understanding of the development of learning communities and the contributions of information professionals in supporting the agricultural community. This integration aimed to establish a comprehensive and context-specific operational guideline for the learning community, as illustrated in Fig. 1, which presents a holistic overview of the research findings. The study emphasized the role of information professionals in supporting the relationship-building mission, focusing on four key functions: (1) outreach or information provision, (2) collaboration or information evaluation, (3) involvement or information analysis, and (4) consultation or information advising (Fig. 2). Furthermore, the findings revealed that agricultural communities in various areas of Chiang Mai Province already have organized longan farmer groups, such as the ALCs, which operate under the supervision of the Minis-

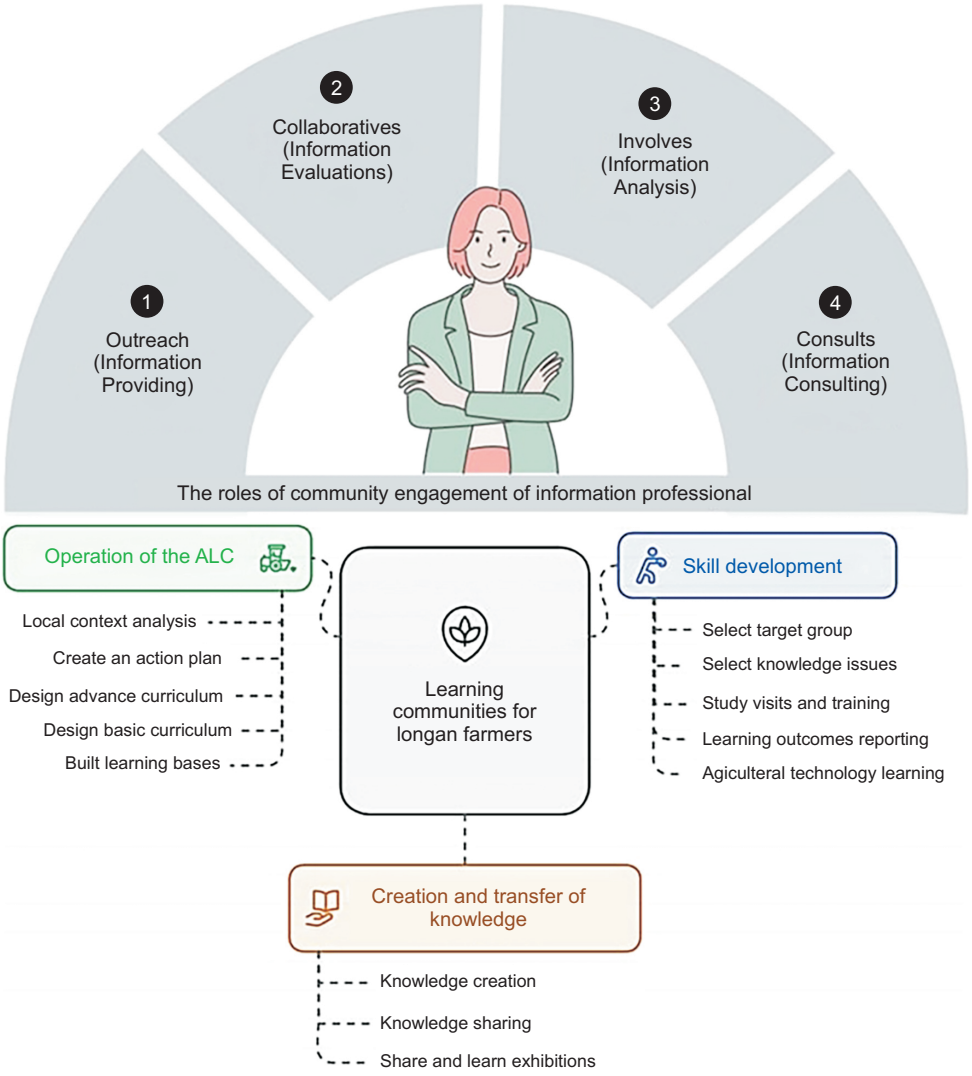



**Fig. 1.** Roles of community engagement of information professionals. ALC, Agricultural Learning Center.

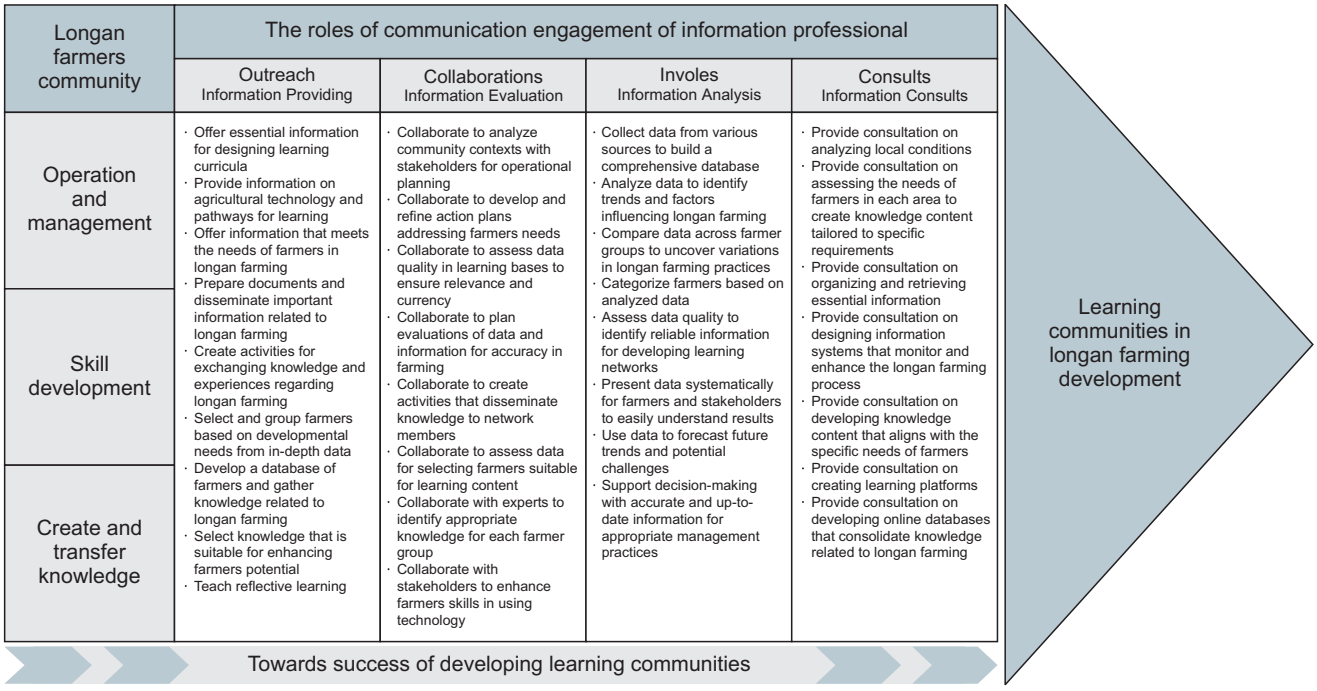



**Fig. 2.** Community engagement roles of information professionals in developing longan farmer learning communities.

try of Agriculture and Cooperatives. The researcher also examined the current operational processes of these centers, focusing on the development of farmers' skills, both among center leaders and network members, as well as the forms of knowledge creation and transfer.

The integration of quantitative and qualitative findings provided a deeper understanding of the roles of information professionals in developing learning communities for longan farming. Quantitative results indicated that outreach activities and collaborative roles were highly influential ($\overline{\chi}$=3.88), while information analysis and consultation also played important roles ($\overline{\chi}$=3.80 and $\overline{\chi}$=3.77, respectively). These findings were further elaborated by qualitative insights, where participants described how information professionals actively engaged in building trust through outreach, fostering cooperative learning environments, and tailoring information services to local needs.

For instance, while the quantitative results highlighted that outreach activities were highly valued ($\overline{\chi}$=3.88), qualitative data from interviews emphasized the importance of continuous engagement and trust-building efforts in sustaining long-term relationships within the community. Participants explained that the outreach role involved not only information dissemination but also maintaining a constant presence in the community to foster collaboration and ensure that the information provided met farmers' evolving needs. This qualitative insight complements the quantitative data, stressing the sustained effort required for successful outreach and the critical role it plays in enabling farmers to take immediate action based on timely and relevant information.

Similarly, the collaborative role was rated as highly influential ($\overline{\chi}$=3.88) in the quantitative results, and this was further enriched by qualitative findings. Interviewees explained that collaboration involved working closely with farmers to understand their unique needs, create customized solutions, and ensure the relevance of the information provided. This collaboration also aligned data collection with development goals, ensuring that the shared information was both accurate and actionable. The qualitative data revealed how successful collaborations were vital in addressing long-term agricultural challenges, with the support of information professionals directly improving farming practices.

The information analysis role, with a mean score of 3.80 in the quantitative results, was supported by qualitative data showing how information professionals systematically collected and analyzed data on variables such as weather patterns, soil conditions, and farming technology. This analysis enabled farmers to make informed decisions and adapt their practices based on data-driven insights. Participants spoke about how this role helped them pre-

dict and anticipate challenges, such as climate variations or pest outbreaks, that could affect crop yields.

Finally, the consultation role, with a mean score of 3.77, was supported by qualitative findings that described how information professionals guided farmers through specific agricultural challenges. Farmers emphasized the value of expert advice in making informed decisions on pest control, soil management, and overall crop health. The consultation provided by information professionals was considered essential not only for short-term problem-solving but also for long-term sustainable agricultural development.

The integration of these two strands revealed that while quantitative trends provided clear ratings of the roles, qualitative narratives enriched the understanding of how these roles were implemented and their contextual challenges. For example, although outreach activities were highly rated quantitatively, qualitative interviews illustrated the importance of continuous presence and relationship-building efforts in sustaining community trust. In this way, the two strands of data complemented and triangulated each other, offering a comprehensive and nuanced interpretation of the roles and contributions of information professionals in fostering learning communities for longan farming.

## 5. DISCUSSION

The research findings regarding the role of information professionals in supporting community engagement among farmers highlight the significance of having information experts in creating sustainable learning communities for farmers, especially in the context of longan farmers in Chiang Mai Province. Information professionals play a crucial role in connecting necessary agricultural information between farmers and appropriate sources. The dissemination of high-quality and relevant information can enhance the potential for collaborative learning among farmers (Idiegbeyan-Ose et al., 2019; Mia, 2020). Their role in evaluating information and transferring knowledge about data assessment to farmers enables them to make decisions and assess the value of using the information in agricultural practices independently, which promotes sustainable learning and enhances the community's ability to self-rely (Junaidi, 2022; Rufaidah & Junaidi, 2022).

The research also underscores the importance of collaboration between information professionals and farmers in exchanging information and jointly developing data evaluation skills. Working with farmers to identify key data for decision-making, such as crop selection, planting plans, or resource management, helps farmers develop the knowledge and skills necessary to adapt and enhance agricultural practices effectively (Wolfert et al., 2017). Additionally, information professionals act as consultants in data management and promote the use of appropriate tools and technologies to exchange information among farmers, focusing on developing sustainable and efficient agricultural outcomes (Lee et al., 2020).

A review of relevant literature and studies related to agricultural support in various regions worldwide demonstrates that information professionals in Southeast Asia, such as in Vietnam, have developed digital platforms, such as databases, that help farmers easily access information about climate conditions and agricultural technologies (Catelo et al., 2016; Tuan, 2024). Projects like the Smart Farmer Libraries in Vietnam are excellent examples of using technology to promote learning and agricultural development in rural areas. In Africa, the use of radio broadcasts and community radio to disseminate agricultural information in resource-limited areas is also an essential strategy (Buehren, 2023; Ng'eno & Mutula, 2022).

In the Americas, information professionals in the United States have supported farmers through Extension Services programs and the Ag Data Commons platform, which provides information on crop management and marketing, helping to support farmers' decision-making. They also connect farmers with researchers to promote agricultural technology development in Brazil and Mexico (U.S. Department of Agriculture, n.d.). In Europe, information professionals utilize advanced technologies such as Big Data and AI in agricultural data analysis, with the AGRIS database—a global agricultural research database—allowing farmers to access critical research information quickly. Furthermore, community libraries in Spain and Italy have organized activities that promote knowledge exchange on sustainable agriculture (FAO, 2022).

When considering and comparing the research and roles of information professionals in supporting agriculture across various regions globally, it is evident that, whether in Asia, Africa, Europe, or the Americas, the important role of information professionals in supporting the creation and development of learning communities for farmers is universally recognized. This is especially true in terms of providing necessary, high-quality information to support agricultural decision-making, which enables farmers to develop skills in assessing and using information to improve their agricultural methods. Information

professionals serve as consultants in data management and facilitate knowledge exchange among farmers for collaborative learning (Junaidi, 2022; Mia, 2020). Additionally, digital platforms that enable farmers to access up-to-date agricultural information, such as climate and agricultural technology databases, are being developed (Catelo et al., 2016).

However, our research focuses specifically on the context of longan farmers in Chiang Mai, where the role of information professionals is applied to managing data specifically related to longan farming and the selection of suitable agricultural technologies. Research from other areas highlights the role of information professionals in various global regions, including the use of radio broadcasts in Africa and advanced technologies such as AI and Big Data in Europe (FAO, 2022; Ng'eno & Mutula, 2022), demonstrating the diversity of approaches used in different regions. Thus, the role of information professionals in promoting community engagement among farmers not only contributes to agricultural development but also enhances the sustainability and adaptability of farming communities in all regions. This is a key element in fostering the sustainable development of agriculture in the future.

## 6. CONCLUSION AND RECOMMENDATIONS

The role of information professionals is pivotal in fostering an environment conducive to learning for longan farmers. These professionals act as intermediaries, connecting the farming community with essential information resources crucial for their agricultural activities. They play a significant role in the development of learning communities by providing and maintaining access to relevant information, which enhances the farmers' capacity for collective learning.

Information professionals also contribute by delivering knowledge and facilitating skill development, enabling farmers to acquire in-depth knowledge applicable to their farming practices. Collaboration between information professionals, farmers, and agricultural experts ensures the provision of high-quality, evaluated information that supports learning activities. In addition to their role in information evaluation, teaching farmers how to independently assess valuable information and organizing knowledge assessments empowers them to make informed decisions. This collaborative learning process fosters the development of sustainable learning communities.

Moreover, information professionals serve as information analysts, addressing farmers' needs by helping them make informed decisions related to agricultural operations, such as crop selection, optimal planting times, and resource allocation. A learning community approach, centered on relevant and community-specific information, positively influences the creation of effective development plans. Information professionals also act as consultants, assisting farmers in identifying critical knowledge and facilitating the management and exchange of data within farming communities. Their role in recommending and developing appropriate learning media further fosters a culture of continuous learning and adaptation.

With Thailand's robust Internet infrastructure, farmers are increasingly utilizing social media. Consequently, it is essential to enhance their capacity to leverage digital technology for practical benefits and innovation. Information professionals can play a crucial role in cultivating information literacy skills among farmers, enabling them to access reliable sources, critically assess the quality of information, and apply it ethically. Moreover, by equipping farmers with knowledge of digital technology's potential, they will be better positioned to navigate emerging challenges and adapt to technological advancements. This, in turn, will foster the rapid development of longan farming communities as farmers share and apply knowledge collaboratively, contributing to sustainable community development.

While this study focused on the context of longan farmers in Chiang Mai, Thailand, its findings provide a foundation for broader applications in global agricultural environments. The framework for developing learning communities and the roles of information professionals identified in this research can be adapted to various agricultural and socio-economic settings. For example, in resource-constrained regions such as Sub-Saharan Africa or South Asia, integrating digital tools with community engagement strategies can enhance agricultural productivity and sustainability.

Moreover, cultural adaptation is essential to ensure the success of such initiatives. In collectivist societies, for instance, leveraging traditional knowledge and local networks can foster trust and facilitate knowledge exchange. Similarly, regions with significant economic disparities may benefit from low-cost educational resources and partnerships with international organizations to bridge technological gaps.

This research underscores the importance of designing community engagement strategies that prioritize effective information dissemination and long-term sustainability. Establishing global networks of learning communities

 https://doi.org/10.1633/JISTaP.2025.13.3.2

could promote the exchange of best practices and drive innovation in agricultural education. Future research should explore cross-cultural adaptations to refine and expand the utility of this framework to diverse agricultural ecosystems worldwide.

However, to maximize the impact of information professionals in supporting longan farmers and fostering sustainable learning communities, it is essential to strengthen their role through targeted training, collaborative partnerships, and digital integration. Training programs should be designed to enhance their ability to evaluate and utilize agricultural information effectively. Establishing collaborative networks between information professionals, farmers, and agricultural scholars at various levels will promote knowledge sharing and interdisciplinary cooperation. Additionally, integrating traditional agricultural wisdom with modern technological advancements will ensure that farmers can access a holistic knowledge base. Developing an easily accessible agricultural information database will further empower farmers by providing reliable, localized information in user-friendly formats. By implementing these recommendations, the role of information professionals in agricultural learning communities can be significantly enhanced, leading to improved knowledge dissemination and long-term sustainability.

## 7. LIMITATIONS OF THE STUDY

The qualitative research section of this article has several limitations that should be considered when interpreting the results. The sample consisted of model longan farmers who are heads of ALCs for longan farming, with only five heads of ALCs in Chiang Mai. Additionally, five members of the ALCs were included, making a total of ten participants. The purposive sampling method was used to select participants with direct experience in longan farming. While some participants were unable to provide data (4 out of 10), the data obtained from these participants still offer valuable insights. These findings are relevant and can be applied to similar agricultural contexts, particularly in regions with comparable longan farming practices or operations.

## CONFLICTS OF INTEREST

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

## ACKNOWLEDGEMENTS

This research was supported by the Faculty of Liberal Arts, Maejo University, Chiang Mai, Thailand.

## REFERENCES

Akther, J. (2020). Influence of UNESCO in the development of lifelong learning. *Open Journal of Social Sciences*, 8, 103-112. https://www.scirp.org/pdf/jss_2020030911335598.pdf

Akuku, B., Waema, T. M., Oboko, R., & Brown, I. (2021). Knowledge management strategies adopted in agricultural research organizations in East Africa. *Information Development*, 37(4), 671-688. https://doi.org/10.1177/0266666920968165

Bonferroni, C. E. (1936). *[Statistical theory of classes and probability calculus]*. Seeber. Italy.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Buehren, N. (2023). *Gender & agriculture in Sub-Saharan Africa: Review of constraints and effective interventions*. https://openknowledge.worldbank.org/server/api/core/bitstreams/19966f50-f68b-42c9-9088-9098dfaca729/content

Caspe, M., & Lopez, M. E. (2018). Preparing the next generation of librarians for family and community engagement. *Journal of Education for Library and Information Science*, 59(4), 157-178. https://www.jstor.org/stable/90025869

Catelo, M. A. O., Francisco, H. A., & Darvin, B. A. C. (2016). *Economic instruments in environmental and natural resource management in Southeast Asia and China: Lessons and way forward*. https://ideas.repec.org/p/sag/seappr/2016326.html

Chiang Mai Agricultural Extension Office. (n.d.). *Agricultural Productivity Improvement Learning Center (APILC)*. https://chiangmai.doae.go.th/province/?page_id=4484

Clinical and Translational Science Awards Consortium, Community Engagement Key Function Committee, Task Force on the Principles of Community Engagement, United States, Agency for Toxic Substances and Disease Registry, & Centers for Disease Control and Prevention (U.S.). (2011). *Principles of community engagement (second edition)*. https://stacks.cdc.gov/view/cdc/11699

Cronbach, L. J. (1990). *Essentials of psychological testing* (5th ed.). Harper & Row.

Food and Agriculture Organization of the United Nations (FAO). (2017). *Information and communication technology (ICT) in agriculture: A report to the G20 Agricultural Deputies*. https://openknowledge.fao.org/items/fa02c775-

19fc-4a1f-b0a1-9369ac776854

Food and Agriculture Organization of the United Nations (FAO). (2022). *AGRIS: The International System for Agricultural Science and Technology*. https://openknowledge.fao.org/server/api/core/bitstreams/bb0e3f59-6e95-47db-9d3c-524197e8a869/content

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200), 675-701. https://doi.org/10.2307/2279372

Greer, R. C., Grover, R., & Fowler, S. G. (2013). *Introduction to the library and information professions* (2nd ed.). Libraries Unlimited.

Idiegbeyan-Ose, J., Aregbesola, A., Owolabi, S. E., & Eyiolorunshe, T. (2019). Relationship between motivation and job satisfaction of staff in Private University Libraries, Nigeria. *Academy of Strategic Management Journal*, 18(1), 1-13. https://www.abacademies.org/articles/relationship-between-motivation-and-job-satisfaction-of-staff-in-private-university-libraries-nigeria-7894.html

Jarolímek, J., Samek, J., Šimek, P., Stočes, M., Vaněk, J., & Pavlík, J. (2024). Information sources in agriculture. *Plant, Soil and Environment*, 70(11), 712-718. http://dx.doi.org/10.17221/361/2024-PSE

Joungtrakul, J., & Ferry, K. N. (2020). [Mixed methods research: Applying six mixed methods research strategies]. *Governance Journal*, 9(1), 27-53. Thai. https://so01.tci-thaijo.org/index.php/gjournal-ksu/article/download/242505/164555/

Junaidi, H. (2022). [Agricultural information literacy by urban agricultural communities in agricultural library laboratory based on social inclusion]. *Visi Pustaka*, 24(2), 161-170. Indonesia. https://doi.org/10.37014/visipustaka.v24i2.2858

Lee, B., Liu, J. Y., & Chang, H. H. (2020). The choice of marketing channel and farm profitability: Empirical evidence from small farmers. *Agribusiness*, 36(3), 402-421. https://doi.org/https://doi.org/10.1002/agr.21640

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140), 1-55.

Mancini, J. F. (2012). *The rise of the information professional: A career path for the digital economy*. https://computhink.in/wp-content/uploads/2017/01/Rise-of-the-Information-Professional-White-Paper.pdf

Mia, M. S. (2020). The role of community libraries in the alleviation of information poverty for sustainable development. *International Journal of Library and Information Science*, 12(2), 31-38. https://doi.org/10.5897/IJLIS2020.0942

Ministry of Agricultural and Cooperatives. (2023). *Five-year operational plan of the Ministry of Agriculture and Cooperatives (2023-2027)*. https://www.moac.go.th/about-strategic_plan_cost?utm_source=chatgpt.com

Ministry of Higher Education, Science, Research and Innovation. (2021). [*Action plan for driving Thailand's development with the BCG economic model 2021-2027*]. Thai. https://waa.inter.nstda.or.th/stks/pub/bcg/BCG-Actioion-Plan.pdf?utm_source=chatgpt.com

Ng'eno, E. J., & Mutula, S. M. (2022). Research data management in Kenya's Agricultural Research Institutes. In N. S. Mani & M. A. Cawley (Eds.), *Handbook of research on academic libraries as partners in data science ecosystems* (pp. 334-361). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-7998-9702-6.ch016

Nobnop, W., Thasanapremsin, S., Rabeabdee, P., Tawanpreeda, C., Saeyang, R., & thitiyotsak, A. (2021). [Learning community and community development]. *Journal of Local Management and Development Pibulsongkram Rajabhat University*, 1(2), 97-111. Thai. https://so06.tci-thaijo.org/index.php/jclmd_psru/article/view/248509

Office of the National Education Commission. (1977). [*National education plan 1977*]. Thai. https://online.fliphtml5.com/uqocz/scmy/#p=1

Rufaidah, V. W., & Junaidi, H. (2022). [Evaluation of the information literacy program during library visit day activities]. *Visi Pustaka*, 23(3), 197-206. Indonesia. https://doi.org/10.37014/visipustaka.v23i3.996

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), 591-611. https://doi.org/10.1093/biomet/52.3-4.591

Sindakis, S., & Showkat, G. (2024). The digital revolution in India: Bridging the gap in rural technology adoption. *Journal of Innovation and Entrepreneurship*, 13(1), 29. https://doi.org/10.1186/s13731-024-00380-w

Singh, V., Mehra, B., & Sikes, E. S. (2022). Agriculture-based community engagement in rural libraries. *Journal of Librarianship and Information Science*, 54(3), 404-414. https://doi.org/10.1177/09610006211015788

Siregar, N. M., Hasibuan, N., Syahbana, A., Kamaluddin, Rafiq, M., Nurhidayah, Marbun, Y., & Harahap, K. (2024). Empowerment of rural communities through information technology (IT) training for teenagers. *International Journal of Community Service Implementation*, 1(4), 117-124. https://afdifaljournal.com/journal/index.php/ijcsi/article/download/195/180

Trade Policy and Strategy Office. (2023). [*Deputy Minister of Commerce recommends accelerating the expansion of new fruit export markets and stimulating domestic consumption*]. Thai. https://tpso.go.th/news/2312-0000000016

Tuan, H. (2024). Agricultural practices and climate resilience: Case study in Vietnam. *International Journal of Climatic*

*Studies*, 3(1), 24-36. https://doi.org/10.47604/ijcs.2476

Tumwebaze, R. P., Walsh, J. N., & Lannon, J. (2025). Knowledge management in the agriculture sector: A systematic literature review. *Knowledge Management Research & Practice*, 23(2), 131-148. https://doi.org/10.1080/14778238.2024.2359419

UNESCO International Research and Training Centre for Rural Education. (2014). *Rural learning community for sustainable rural transformation*. https://inruled.bnu.edu.cn/docs/2021-09/20210916171829949633.pdf

U.S. Department of Agriculture. (n.d.). *Ag Data commons user guide*. https://www.nal.usda.gov/services/agdatacommons

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80-83. https://doi.org/10.2307/3001968

Winberry, J., & Bishop, B. W. (2021). Documenting social justice in library and information science research: A literature review. *Journal of Documentation*, 77(3), 743-754. https://doi.org/10.1108/jd-08-2020-0136

Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming – A review. *Agricultural Systems*, 153, 69-80. https://doi.org/10.1016/j.agsy.2017.01.023

World Bank. (n.d.). *Agriculture and rural development*. https://www.worldbank.org/en/programs/knowledge-for-change/brief/agriculture-and-rural-development

Yamane, T. (1973). *Statistics: An introductory analysis* (3d ed.). Harper & Row.

Yusuf, F. O., Owolabi, S. E., Fagbohun, M. O., Adebayo, O., & Iwu-James, J. (2019). Role of public libraries in promoting youth participation in agriculture in Nigeria: Information as a key driver. *Library Philosophy and Practice (e-journal)*, 2877. https://digitalcommons.unl.edu/libphilprac/2877

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

RESEARCH PAPER
J Inf Sci Theory Pract 13(3): 36-54, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.3.3

# Developing Essential Datasets for a Community Cultural Capital Information System

**Chatkamon Anontachai**
Department of Information Science, Khon Kaen University, Khon Kaen, Thailand
E-mail: chatkamon.a@kkumail.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

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

## ABSTRACT

This research analyzes and develops an essential dataset for creating a community information system (CIS) for managing cultural capital. The study has three primary objectives: (1) to identify and define key data categories required for a CIS focused on cultural capital management; (2) to assess the relevance and necessity of each data category through expert evaluation, particularly from individuals with practical experience in utilizing cultural capital in communities; and (3) to propose a dataset that can be adopted by local administrative organizations in Thailand to support cultural capital management and policy-making. The methodology employed in-depth interviews with six key informants: two cultural practitioners from local administrative organizations, two experts in dataset development, one expert in arts, culture, and local wisdom, and one user of cultural data for professional, educational, or research purposes. Data were analyzed using content analysis to identify standardized datasets, based on concepts of CIS, cultural capital information, and datasets, with verification through triangulation. Findings revealed that the essential data for developing the community cultural capital information system comprises two main datasets: community data, consisting of 117 data points with an overall high necessity rating, and cultural capital data, comprising 127 data points with an overall very high necessity rating. These findings establish the crucial data required for effective utilization of the system in managing cultural capital at the community level.

**Keywords:** cultural capitals, community information system, cultural capitals information, community information dataset, cultural capital dataset

**Received:** February 5, 2025 **Revised:** May 22, 2025
**Accepted:** June 12, 2025 **Published:** September 30, 2025

***Corresponding Author:** Kulthida Tuamsuk
https://orcid.org/0000-0003-0852-8945
**E-mail:** kultua@kku.ac.th

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 (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.



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

## 1. INTRODUCTION

Cultural capital is increasingly recognized as a fundamental resource for sustainable development. It encompasses intangible elements such as beliefs, thoughts, and creativity; tangible cultural products, including books, images, and world heritage; and institutional aspects such as norms and collective recognition. The accumulation of cultural capital shapes individual and collective "taste," reinforcing social stratification and identity (Bourdieu, 1986). Culture is also a convertible form of capital that influences perceptions and lifestyles and can be transformed into economic value (Griswold, 2004). In creative industries, cultural capital contributes to value-added production, resulting in culturally enriched goods and services (Department of Cultural Promotion, 2019). Proper management of these resources is essential for preserving cultural heritage, fostering social cohesion, and promoting local economic development (United Nations Educational, Scientific and Cultural Organization [UNESCO], 2011). One effective tool for this purpose is the community information system (CIS), which supports the collection, organization, and analysis of data to inform decisions and drive development initiatives.

Despite the recognized value of cultural capital, communities in Thailand continue to face challenges in managing these resources effectively. The absence of structured, comprehensive data systems has led to fragmented and underutilized information. Current data management practices often focus on isolated types of cultural capital, particularly tangible assets, while neglecting the broader cultural context and intangible elements. Moreover, the datasets frequently lack integration, consistency, and stakeholder input, limiting their usefulness for policy development and long-term cultural management. Moreover, in the digital age, the need for systematic and strategic data collection has become increasingly urgent. Without an effective CIS and clearly defined datasets, local administrative organizations (LAOs) are unable to fully harness the potential of cultural capital to promote sustainable development.

Theoretical foundations highlight the importance of cultural capital in shaping economic and social hierarchies (Bourdieu, 1986), and its potential to be converted into economic value (Griswold, 2004). CIS have been acknowledged for their capacity to collect, store, and disseminate data relevant to community development (Toubekis et al., 2011). These systems integrate datasets such as statistical and socioeconomic information, community needs, residents' opinions, and records of local activities and networks (Bunch, 1993; Department of Local Administration, 2019; Library of Congress, 2006; Lopez, 2015; Ozuomba et al., 2013; van der Waldt, 2019). Effective use of these datasets can enhance the representation of cultural heritage, promote conservation, and encourage community engagement (Hawkes, 2001). However, technical limitations such as incomplete or inaccurately stored data, redundancies, and inconsistencies continue to hinder the effectiveness of CIS (Auwattanamongkol, 2019; Little & Rubin, 2002).

In Thailand, most existing studies have focused on fragmented aspects of cultural capital, often prioritizing tangible assets without incorporating them into cohesive information systems (Boonto, 2025; Chansanam & Tuamsuk, 2022; Kwiecien et al., 2021; Nachom, 2018). Furthermore, previous research has largely failed to analyze the full range of necessary datasets or to involve community stakeholders in defining data parameters and content (Boonto, 2025; Nonthacumjane & Johansson, 2023). These gaps result in datasets that lack socio-cultural coherence and are inadequate for supporting strategic policy or planning.

This study aims to bridge the gap in cultural capital data management through the development of a comprehensive essential dataset for CIS. The specific objectives are: (1) to identify and define the key data categories required for the development of a CIS focused on managing cultural capital; (2) to assess the relevance and necessity of each data category through expert evaluation, particularly from individuals with practical experience in the use of cultural capital in communities; and (3) to propose a dataset that can be adopted by LAOs in Thailand to support cultural capital management and policy-making.

The study is expected to offer a clear dataset that can be used by LAOs to manage cultural resources more efficiently. It will provide local governments with a data-driven framework to guide the design of policies that effectively utilize cultural capital, and facilitate stronger engagement from community members in preserving and promoting cultural heritage. From an academic perspective, the research will contribute to closing existing gaps in the literature by presenting a comprehensive and integrated approach to cultural capital data management within the Thai context.

## 2. LITERATURE REVIEW

Grounded in theoretical frameworks and an extensive

review of related scholarship, this study integrates perspectives from community informatics (CI) and cultural capital informatics, alongside an examination of the following datasets.

### 2.1. Community Informatics

CI refers to the use of information and communication technology (ICT) to promote community development and participation. It emphasizes accessing and utilizing data to address local issues while creating opportunities for social, economic, and cultural development. CI enhances opportunities for communities to participate in decision making and local development through accessible information sources and digital tools for data management. Additionally, it connects community members, fosters transparency in data governance, and improves access to critical information for decision-making to achieve sustainable local development (Grisales Bohorquez et al., 2022; Gurstein, 2007, p. 11; Mehra et al., 2017).

CI is a field of study focusing on the relationship between ICT design and local communities, as well as implementing ICT projects within communities. It highlights using ICT to develop economical, social, and culturally appropriate and sustainable technological solutions. The field emphasizes the needs and local data practices involved in driving positive social change alongside community stakeholders (Davis et al., 2016; Ozuomba et al., 2013; Stillman & Linger, 2009).

Bunch (1993) categorized CI into two main types: (1) survival information: data related to essential aspects of living, such as health, housing, and political rights, and (2) citizen action information: information that facilitates participation in social, political, and economic processes within the community. The roles of CIS include strengthening social capital and expanding interactions within communities through the development and use of ICT applications that meet the needs of community members. These systems also foster networks that help sustain local resources in a sustainable manner (Gurstein, 2007, p. 74-77). CIS plays a crucial role in enhancing quality of life and providing access to essential services, such as healthcare, education, and economic opportunities (Eze et al., 2021; Lippeveld, 2017).

Community data or community profiles are analyses and reports based on data collected through community surveys. They provide an overview of the community's population, including various aspects such as community geographical area; population and household numbers; age and gender distribution; birth and death rates; income data and economic status; social/community needs; access to services; knowledge, attitudes, and beliefs; and understanding and attitudes of the population toward significant health issues impacting the community (OpenLearn Create, n.d.). Community profiles are seen as a comprehensive approach to gathering information about communities, particularly regarding community needs. They serve as a foundation for driving effective community change and development (van der Waldt, 2019).

From a review of literature related to the use of community data in the development of information systems, it was found that the types of data utilized vary based on the objectives of each system's development. These objectives significantly influence the types and characteristics of required data. Additionally, the unique characteristics of each community, including its specific needs and challenges, determine the appropriate data to use, ensuring

**Table 1.** Community data from literature reviews

| Sources | Bunch (1993) | Ozuomba et al. (2013) | Lopez (2015) | van der Waldt (2019) | Library of Congress (2006) | Department of Local Administration (n.d.) |
|---|---|---|---|---|---|---|
| Community data | -Statistical data<br>-Socio-economic data<br>-Local issues<br>-Residents' viewpoints | -Geographical location or area<br>-Key issues & environmental factors<br>-Community data<br>-Community needs | -Community data<br>-Online activities<br>-Social networks<br>-Shared content<br>-Tourist attractions | -Demographic statistics<br>-Community history<br>-Community structure<br>-Livelihoods<br>-Local institutions<br>-Community infrastructure<br>-Resources<br>-Environment<br>-Agriculture<br>-Conservation<br>-Water | -Individual<br>-Organization<br>-Program or service<br>-Event<br>-Other | -General data<br>-Infrastructure & public utilities<br>-Economics<br>-Society<br>-Public health<br>-Disaster prevention & mitigation<br>-Environments<br>-Finance and budgets |

 https://doi.org/10.1633/JISTaP.2025.13.3.3

alignment with the reality and requirements of community users. Community data has been applied in developing models for studying the context of community development, cultural heritage conservation projects, and as a guideline for data experts and technology developers aiming to design technologies for local communities (Huang et al., 2024). Furthermore, it serves as a medium for disseminating information from LAOs, enabling both public and private organizations to utilize the data effectively. Table 1 (Bunch, 1993; Department of Local Administration, n.d.; Library of Congress, 2006; Lopez, 2015; Ozuomba et al., 2013; van der Waldt, 2019) illustrates examples of community data from various concepts, research, and foundational information that have been employed in developing information systems.

From the analysis of community data based on concepts, research, and foundational information from agencies utilizing community data for the development of information systems, community data can be classified into five groups as follows: (1) statistical data–This category includes population statistics, such as population size, gender distribution, age, marital status, education, employment, health, and housing (Bunch, 1993; van der Waldt, 2019); (2) socio-economic data–This category encompasses general information (geographical factors, community history, community structure, livelihoods, local institutions, and individuals), infrastructure and public utilities (education, access to public services, medical and public health services, disaster prevention and mitigation, community facilities, volunteer organizations, and recreational facilities), social information (religion, ethnic groups), economic information (average household income), environmental information, and financial and fiscal data (Bunch, 1993; Department of Local Administration, n.d.; 2019; Library of Congress, 2006; Ozuomba et al., 2013; van der Waldt, 2019); (3) local issues–This data reflects problems or critical issues in the community, such as unemployment, poverty, lack of educational opportunities, access to healthcare services, infrastructure problems, and transportation. These factors impact the quality of life of residents in the area. Such data can be obtained from reports by government and private agencies, as well as from surveys of local residents' opinions (Bunch, 1993); (4) residents' viewpoints and community needs–This data includes community members' opinions about local situations and issues. It can be gathered through interviews, public meetings, opinion surveys, local media (e.g., letters to the editor or community newspapers), and discussions on social media platforms. It also includes other communication sources that reflect the views and needs of community members. This information is crucial for understanding the real needs of the community (Bunch, 1993; Ozuomba et al., 2013); and (5) activity and network information–This category involves online activities, social networks, shared content, and tourism sites (Library of Congress, 2006; Lopez, 2015).

## 2.2. Cultural Capital Information

Cultural capital is a concept used to study the foundation of communities built upon strength. The French sociologist Bourdieu (1986) defines cultural capital as resources accumulated within individuals, objects, and institutions, shaped and transmitted through educational systems. The result of this accumulation is "taste," which helps to create social distinctions and sustain social classes. Griswold (2004) expands on Bourdieu's concept, stating that culture can be viewed as a type of capital that can be accumulated and converted into economic capital. It influences perceptions, cultural tastes, and lifestyles. The Department of Cultural Promotion (2019) explains that cultural capital refers to the inheritance of cultural wisdom, which can be used to add value in creative industries, transforming it into goods or services with added value. The UNESCO Institute for Statistics (n.d.) categorizes cultural heritage into two types: (1) tangible cultural heritage, which includes movable cultural heritage, such as paintings, sculptures, coins, and manuscripts; immovable cultural heritage, such as monuments, archaeological sites, and other structures; underwater cultural heritage, such as shipwrecks, underwater ruins, and submerged cities; and (2) intangible cultural heritage, which includes oral traditions, performing arts, rituals, and other cultural expressions.

Cultural heritage information (CHI) refers to all information related to cultural heritage, including data on the heritage itself, its management, preservation, relocation, and the administration of institutions responsible for such heritage. CHI consists of resources with evaluated significance, such as ancient documents, photographs, and historical maps, which require proper storage and preservation to be passed on to future generations. CHI is classified into three main categories: (1) collection information, which involves data about heritage items such as museum objects and documents; (2) technical information, which includes data on technical activities such as conservation and exhibitions; and (3) administrative information, which encompasses data on the management of institutions responsible for heritage, such as policies and visitor

**Table 2.** Cultural capital information from literature reviews

| Cultural capital information | | | | | |
|---|---|---|---|---|---|
| Tangible cultural capital | Intangible cultural capital | | | | |
| Moveable/immoveable/ underwater cultural heritage | Oral traditions & expression | Performing arts | Social practices, rituals & festival events | Knowledge and practices, concerning nature and the universe | Traditional craftsmanship |
| Trust (2024) | Kwiecien et al. (2021) | Giannoulakis et al. (2018) | Chansanam & Tuamsuk (2022) | Nachom (2018) | Ye (2023) |
| 1. Object/work<br>2. Classification<br>3. Titles or names<br>4. Creation<br>5. Style/period/group/ movement<br>6. Measurements<br>7. Material and techniques<br>8. Inscription/marks<br>9. State<br>10. Edition<br>11. Facture<br>12. Orientation/ arrangement<br>13. Physical description<br>14. Condition/ examination history<br>15. Conservation/ treatment description<br>16. Subject matter<br>17. Context<br>18. Descriptive note<br>19. Critical responses<br>20. Related work<br>21. Current location<br>22. Copyright/ restriction<br>23. Ownership<br>24. Exhibition/loan history<br>25. Cataloguing history<br>26. Related visual document<br>27. Related textual reference<br>28. Person/corporate body authority<br>29. Place/location authority<br>30. Generic concept authority<br>31. Subject authority | 1. Identifier<br>2. Title<br>3. Creator<br>4. Contributor<br>5. Description<br>6. Keyword<br>7. Character<br>8. Moral<br>9. Ethnic group<br>10. Motif<br>11. Place<br>12. Relation<br>13. Country<br>14. Language<br>15. Medium<br>16. Source<br>17. Date<br>18. Rights | Dance<br>1. Title<br>2. History<br>3. Country/region of origin<br>4. Time of origin<br>5. Dance variation<br>6. Music motif<br>7. Lyrics<br>8. Dance annotation<br>Recording<br>9. Title<br>10. Description<br>11. Location in GPS format and human readable<br>12. Number of sensors<br>13. Sensor type & characteristic<br>14. Calibration<br>15. Software<br>3D Environment<br>16. Venue description<br>17. Object and their description<br>18. Description of the lighting<br>Dancer<br>19. Gender<br>20. Face expression<br>21. Traditional costume description | 1. Tradition name<br>2. Local tradition name<br>3. Organized month<br>4. Time according to the lunar calendar<br>5. Objective<br>6. Activity<br>7. Ritual<br>8. Literature<br>9. Belief<br>10. Place<br>11. Ceremony<br>12. Participant<br>13. Equipment<br>14. Building | 1. Title<br>2. Identifier<br>3. Creator<br>4. Description<br>5. Language<br>6. Ethics<br>7. Date time<br>8. Media<br>9. Keywords<br>10. Rights<br>11. Location<br>12. Categories<br>13. Thai Name<br>14. English name<br>15. Local name | 1. Apparel name<br>2. Keyword<br>3. Code character<br>4. Clothing category<br>5. Apparel owner<br>6. Operator<br>7. Nation to which the dress belongs<br>8. Clothing generation age<br>9. Clothing production area<br>10. Modeling<br>11. Size<br>12. Weight<br>13. Color<br>14. Pattern<br>15. Material<br>16. Process form<br>17. Accessories<br>18. Physical quality and evaluation<br>19. Wearing a group<br>20. Use occasion<br>21. Cultural connotation<br>22. Technological level<br>23. Value<br>24. Authority<br>25. Data collector<br>26. Data acquisition time<br>27. Physical storage location<br>28. Image storage location |
| 31 Elements | 18 Elements | 21 Elements | 14 Elements | 15 Elements | 28 Elements |
| 127 Elements | | | | | |

 https://doi.org/10.1633/JISTaP.2025.13.3.3

statistics. This information is crucial for preserving and transmitting cultural heritage from one generation to the next (Chaichuay, 2018).

Currently, institutions providing cultural information services, such as libraries, archives, and museums, emphasize the development of information services to improve the efficiency of access, sharing, and analysis of cultural data. These efforts aim to facilitate the effective use of such data for education, research, and conservation purposes. Efforts to enhance these services take various forms, particularly in the development and application of metadata. For example, Salse et al. (2022) studied metadata structures in museums and university collections across 23 institutions in Spain and Europe. They found that museums employed a variety of metadata standards, including Categories for the Description of Works of Art (CDWA), SPECTRUM, ICCD, MuseumDat, VRA Core, and LIDO. However, many museums still use custom metadata structures to meet the specific needs of their organizations. Similarly, a study by Pandey and Kumar (2023) examined metadata element sets used for digital art objects in Indian museums. Among the five museums studied, art objects were categorized into 13 categories, with each category employing different metadata standards to describe the data. It is evident that libraries, archives, and museums adopt diverse metadata standards and preserve various types of resources, making it impractical to use a single metadata standard for all resource types. However, having robust metadata is crucial for dissemination. Therefore, organizations must select metadata standards appropriate to the resources they manage and prioritize their effective implementation.

From a review of literature on cultural data management, there have been efforts to develop metadata capable of comprehensively and thoroughly describing cultural heritage according to its type. Based on the classification of cultural heritage by UNESCO, cultural heritage can be categorized into two types: tangible cultural heritage and intangible cultural heritage, as outlined in Table 2 (Chansanam & Tuamsuk, 2022; Giannoulakis et al., 2018; Kwiecien et al., 2021; Nachom, 2018; Trust, 2024; Ye, 2023).

### 2.3. Datasets

Datasets refer to collections of data organized into a structured format for practical use. These datasets often consist of interrelated data presented in an orderly manner, such as tables with rows and columns or files containing data in various formats, including text, numbers, images, or audio. Datasets can take many forms, such as tables, text files (e.g., CSV, JSON), or databases. A dataset typically includes multiple rows (records) and columns (attributes or fields) that store related information. In the fields of data science and data analysis, datasets serve as critical resources for researchers and analysts to process, summarize, predict, and create data analysis models (databricks, n.d.; Kelleher & Tierney, 2018).

Datasets in research and data analysis can be categorized into several types based on the characteristics of the records and the types of variables that make up the dataset (Auwattanamongkol, 2019, pp. 10-11). For example: (1) record data refers to a dataset consisting of individual records, each with attributes or characteristics (attributes) that have a defined number and type of data. Each record can be compared to a row in a database table, for instance, personal information, which may include variables such as ID number, name, address, etc. (2) A data matrix consists of records where all variables are numerical data (numeric). The dataset can be represented as a matrix, where each row represents one data instance or record. Each instance can be represented as a vector consisting of numerical values. The vector's size depends on the number of variables in the record: for example, the proportion of ingredients in a food mixture. (3) Transactional data consists of a list of events or items that appear in each record, such as customer purchase data for each transaction. The number of events or items on each record may vary. (4) Document data consists of individual documents (e.g., text). Each record contains variables indicating specific features of the document, such as the frequency of certain words appearing in the document. These variables can help determine the content of the document. This type of data often is high and most of the variable values are zero (sparse high dimensional data). (5) Binary data consists of records where each variable can only have a value of 1 or 0, indicating whether a specific event or status has occurred. This type of data is often converted from transactional data, with many variables, most of which are zero. (6) Graph data consists of data in a graph structure, where each record represents information in the form of nodes and edges, for example, chemical molecular structure data or protein-related data. These structures tend to be irregular and unpredictable. (7) Sequence data consists of a series of events or items occurring in sequence, such as gene sequences in chromosomes or web page visits within a certain time frame. This type of data reflects the occurrence of events in meaningful order. And (8) time series data consists of data where the variables change over time,

such as daily stock indices or daily temperatures throughout the year. This type of data is related to time periods and temporal changes (databricks, n.d.; Khan & Hanna, 2022).

The use of datasets in research is diverse and highly beneficial across many disciplines, such as medicine, marketing, education, finance, and social sciences. Utilizing datasets for analysis can help identify relationships, predict outcomes, and make data-driven decisions based on sound principles. Example include analyzing basic datasets to develop a system for linking family medicine data (Saensrisawat et al., 2024), or developing a dataset on patients with heart disease symptoms to train a model for predicting heart disease risk using various techniques, such as machine learning (Janosi et al., 1989). However, the use of datasets also comes with several limitations that may impact on the quality of analysis and data handling. These limitations may arise from incomplete or improperly stored data that does not meet the user's needs, such as incorrect, incomplete, redundant, or inconsistent data (Auwattanamongkol, 2019; Little & Rubin, 2002). Users of datasets may apply different techniques to prepare the data for more efficient and accurate analysis or model creation—for instance, using a data dictionary to explain the details of each field in the dataset, such as field names, data types, or descriptions. A data dictionary is a set of names, definitions, and characteristics of data elements used or collected in a database, information system, or research project. It explains the meaning and purpose of data elements in the context of the project while providing guidelines for interpreting and displaying the data. Additionally, the data dictionary includes metadata to define the scope, nature, and rules for using data elements, which helps reduce inconsistencies within the project, sets common standards, ensures consistency in data collection and use among research team members, facilitates data analysis, and establishes standards for data management. This makes it easier for users or programmers to understand and manage the dataset effectively (UC Merced Library, n.d.).

The management of cultural data in Thailand, despite efforts from various agencies to create databases and collect information, still faces challenges. A case study by Boonto (2025) on the development of digital cultural heritage management in local museums in Chiang Mai aimed to examine the lessons learned from managing and digitizing cultural heritage data using the Navanurak and Museum Pool platforms. The study also aimed to explore the development of digital cultural heritage data management for local museums in Thailand. The findings revealed that the studied area has a strong community network and is effectively able to connect museums with cultural tourism in the community. However, the data stored in the system primarily focuses on museum objects. The planning for data management from the initial stages of museum establishment lacked a systematic approach to storing cultural heritage data, such as the absence of an object register. This has resulted in challenges in the management and digitization processes, as there is a lack of necessary data in both the overall context and the specific context of each object. Furthermore, the absence of a connection between the object data and local community data has led to information that does not fully reflect the realities of the area or carry cultural significance in an authentic way (Boonto, 2025).

Studies on the role of LAOs in cultural capital information management reveal that LAOs have legal responsibilities for promoting and preserving local cultural heritage. Cultural operations typically fall under the purview of educational personnel who may lack specialized expertise in cultural management, resulting in unsystematic cultural information management. Consequently, most cultural data remains scattered and outdated. The cultural capital in the studied areas can be categorized into tangible and intangible assets, as well as policy-related information and administrative data. The research indicates that while LAOs have formal mandates for cultural preservation, the implementation of systematic cultural information management remains challenging due to the absence of systematic practices for data management. Despite these limitations, local cultural information management practices show promise in fostering educational opportunities, enhancing community engagement, and contributing to sustainable cultural conservation efforts (Anontachai et al., 2025).

Another study by Nonthacumjane and Johansson (2023) on shaping local information in Thailand identifies key issues in the digitalization of local information within the Provincial University Library Network (PULINET). These include contradictions in the conceptualization of local information among group members and the definitions established by the group. Additionally, there is a misalignment between the conceptual view of local information as a dynamic resource and the actual digitalization activities, which focus on formal boundaries and prominent objects, thus undermining the diversity and contextual richness of local information. The study also highlights the exclusion of marginalized groups and

 https://doi.org/10.1633/JISTaP.2025.13.3.3

expressions that are not deemed socially acceptable. Furthermore, these contradictions are not leveraged as opportunities for learning and development but are instead concealed and resisted, preventing the full potential of local information digitalization from being realized (Nonthacumjane & Johansson, 2023).

Building upon the aforementioned research, the review of relevant studies underscores key challenges in the management of cultural heritage and local information in Thailand. These challenges include the absence of systematic data storage, the fragmentation of data across different sources, and the lack of comprehensive data management frameworks that encompass both tangible and intangible cultural heritage. Furthermore, there is a notable deficiency in the diversity and contextual richness of local information. Developing datasets for a community cultural capital information system is essential for effectively linking museum data with community information. By constructing datasets that integrate cultural heritage and local administrative data while reflecting cultural diversity, these efforts can address the identified challenges and foster systematic data management, thereby facilitating sustainable cultural development and preservation at the local level.

In this research, the authors studied and analyzed standards for storing various types of data, including those used for managing community and cultural heritage data. This includes the CDWA standard for tangible cultural heritage data and research that developed metadata for intangible cultural heritage data, to analyze the necessary information for managing cultural heritage. They then developed a dataset using a data dictionary as a tool to manage and explain the data, such as field names, data types, data definitions, and related rules. This helps clarify the meaning of the data, reduces confusion, and ensures that data usage is organized and understood consistently.

## 2.4. Research Conceptual Framework

From an analysis of previous studies and literature, the researchers have established a research conceptual framework based on relevant concepts, theories, and related studies, including CI, cultural capital information, and datasets, as illustrated in Fig. 1 (Library of Congress, 2006; UNESCO, 2011).

According to Fig. 1 (Library of Congress, 2006; UNESCO, 2011), CI is a concept that focuses on utilizing ICT to strengthen communities in areas of development, learning, and network connectivity. Cultural capital information, on the other hand, refers to data related to cultural capital, which reflects the knowledge, wisdom, beliefs, and identity of the community. When integrated with systematically collected datasets, this information can be analyzed to provide a clearer and more comprehensive understanding of the community's overall picture and context. In developing the research conceptual framework, the researchers link the relationships found in the literature, beginning with the overarching concept of CI, proceeding in-depth into cultural capital information as a key data component that should be documented, employing datasets as tools for data collection, analysis, and transformation into new knowledge. These elements are interconnected in a systematic sequence, progressing from concept to practice and leading to the design of a clear and structured research conceptual framework.

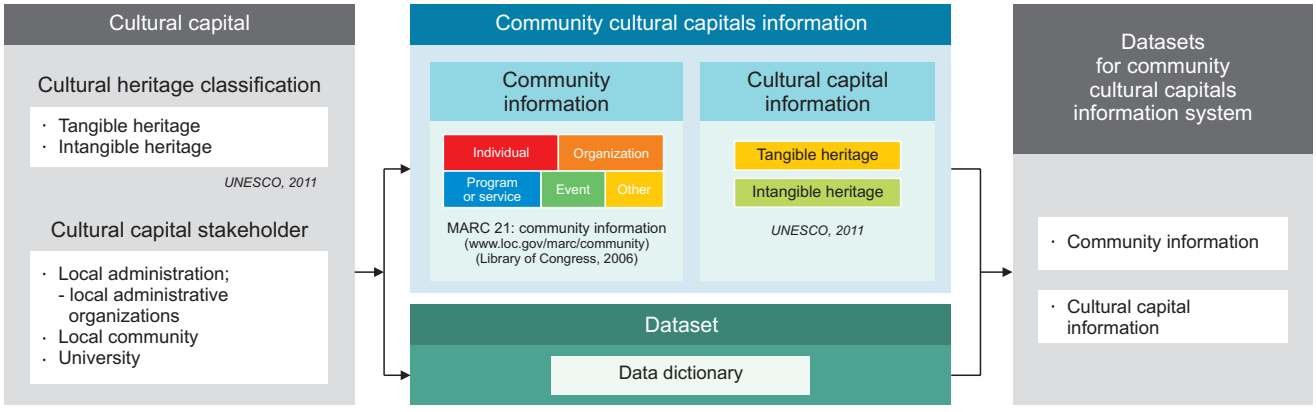



**Fig. 1.** Research conceptual framework. MARC, machine readable cataloging; UNESCO, United Nations Educational, Scientific and Cultural Organization.

## 3. RESEARCH METHOD

This study is part of a research project titled "Community Information System for Managing Cultural Capitals for Sustainable Development." The specific objectives are: (1) to identify and define the key data categories required for the development of a CIS focused on managing cultural capital; (2) to assess the relevance and necessity of each data category through expert evaluation, particularly from individuals with practical experience in the use of cultural capital in communities; and (3) to propose a dataset that can be adopted by LAOs in Thailand to support cultural capital management and policy-making. The research process comprises three main stages, as outlined below.

### 3.1. Formulating of the Research Topic

The researchers identified a topic that is both pertinent and significant to the study's objectives, emphasizing the analysis of key datasets essential for the development of a community cultural capital information system. The objective is to acquire datasets that are comprehensive, clear, and conducive to the efficient management of cultural capitals, thereby fostering sustainable development at both local and national levels. Based on the objectives of the study, the research questions can be identified as shown in Table 3.

### 3.2. Research Design

In designing the research process, the researchers established a framework and methodology that aligns with the objectives of the study, taking into account relevant elements such as theories, concepts, requirements, and related experiences. This approach ensures that the study is comprehensive and capable of addressing the research questions effectively. Additionally, the researchers defined a structured and standardized research framework and plan to ensure that the research process leads to accurate, reliable, and valid results. The details of the research design and process are summarized in Table 4.

### 3.3. Designing the Structure of Datasets to Identify and Define Key Data Categories

To fulfill the objective of identifying and defining the key data categories required for the development of a CIS focused on managing cultural capital, the researchers conducted a systematic process of dataset structuring. This

**Table 3.** Research objectives and research questions

| Research objectives | Research questions (RQ) |
|---|---|
| (1) To identify and define the key data categories required for the development of a CIS focused on managing cultural capital | RQ1: What "community data" is necessary to be collected for the benefit of utilizing community data?<br>RQ2: What "cultural capital data" is necessary to be collected for the benefit of utilizing cultural capital data? |
| (2) To assess the relevance and necessity of each data category | RQ3: What is the level of necessity for using community data?<br>RQ4: What is the level of necessity for using cultural capital data? |
| (3) To propose a dataset for the development of a CIS focused on managing cultural capital | |

CIS, community information system.

**Table 4.** Summary of research design and process

| Research objectives | Methodology/approach | Data sources/informants |
|---|---|---|
| 1. To identify and define the key data categories required for the development of a CIS focused on managing cultural capital | -Content analysis | Policy documents, academic research, and existing cultural metadata |
| 2. To assess the relevance and necessity of each data category | -In-depth interviews (with interview guidelines)<br>-Evaluation form (rating scale)<br>-Dataset modification | Six experts who are directly involved in the use of community cultural capital data |
| 3. To propose a dataset for the development of a CIS focused on managing cultural capital | | |

CIS, community information system.

 https://doi.org/10.1633/JISTaP.2025.13.3.3

process aimed to ensure that all essential data related to both community and cultural capital were clearly categorized and appropriately defined. The steps are outlined as follows.

#### *3.3.1. Identification of Data Components*

Based on document analysis and field data, the researchers compiled and listed essential data elements relevant to cultural capital and community context. This resulted in two primary data categories: Community data with 117 elements, and cultural capital data with 127 elements.

#### *3.3.2. Analyzing and Grouping Data Components*

After identifying the data components, the researchers analyzed their importance based on their relevance to practical applications. The components were then grouped based on concepts derived from the literature and related research:

- Community data: Sub-categorized into basic information about local government organizations (LGOs) responsible for the area and fundamental community data.
- Cultural capital data: Divided into two types—tangible and intangible cultural capital data.

The names and definitions of data elements were refined to ensure clarity and maintain the original meaning of all relevant components.

#### *3.3.3. Adopting Metadata Standards*

Based on literature and related research, researchers selected metadata standards as a framework for describing the data, including:

- Community information standards: Using the machine readable cataloging (MARC) format, specifically MARC 21 for community information, designed to systematically link community services data. MARC 21 is used to document non bibliographic resources meeting community information needs. Its scope includes individuals, organizations, programs or services, events, and others. This was combined with essential community data and basic LGO data.
- International metadata standards for museums: For instance, CDWA was used to describe cultural documents and objects. The researchers also designed datasets related to traditions, culture, and local wisdom, adapting metadata from relevant research.

#### *3.3.4. Knowledge Integration*

The researchers studied and analyzed content from documents, manuals, and research related to arts, culture, and local wisdom. They also interviewed individuals involved in utilizing cultural capital data in the community. The interviews focused on identifying necessary cultural capital data for practical use, which served as a guideline for dataset development.

#### *3.3.5. Dataset Design*

Based on the processes above, the researchers designed the dataset into two groups: community data, and cultural capitals data.

### 3.4. Evaluation and Refinement of the Dataset

In the final stage, the researchers evaluated the information derived from the previous stages to determine which data is essential for future development or research. This evaluation process involved additional interviews with experts to arrive at more definitive conclusions.

#### *3.4.1. Expert Evaluation*

The evaluation and refinement of the cultural capital dataset for sustainable development were conducted by experts selected through purposive sampling. These experts were individuals directly involved in using community cultural capital data and were divided into four groups: (1) Cultural practitioners in LGOs; (2) Data experts familiar with dataset structuring and management; (3) Experts in arts, culture, and local wisdom; and (4) Cultural data users for work, education, or research purposes. The number of experts participating in the study ranged from 6 to 12, which is consistent with the recommendations of Guest et al. (2006), who suggested that data saturation typically occurs within 6 to 12 interviews. Interviews continued until no new significant themes emerged from the data, a phenomenon known as thematic saturation. This ensures that the data collected is rich and comprehensive enough to effectively address the research questions. The selection of the number of experts also aligns with the approach suggested by Malterud et al. (2016), who recommended that the sample size be determined based on "information power," which refers to the number of interviews required to provide sufficient data to answer the research questions. Information on the key informants is presented in Table 5.

#### *3.4.2. Research Tools*

The research utilized interview guidelines with questions aligned to the study's objectives, focusing on essential

**Table 5.** Information on the key informants

| Group of key informants | Qualifications | Work experience | Number of people (N) |
|---|---|---|---|
| 1) Cultural practitioners in local government organizations in Roi Et | Executive or cultural practitioner in local government organizations in the study area, Roi Et province | At least 3 years of experience | 2 |
| 2) Data experts | Individuals with work experience or research related to community data or cultural arts and local wisdom | At least 3 years of experience | 2 |
| 3) Experts in cultural arts and local wisdom | Recognized experts in cultural arts and local wisdom with work experience or research related to the field | At least 3 years of experience | 1 |
| 4) Users of cultural data for work, education, or research | Individuals using cultural data for work, education, or research | At least 3 years of experience | 1 |

**Table 6.** Data classification for cultural capital

| Data record | Code group | Concomitant relationships | Key terms |
|---|---|---|---|
| "When designing a cultural capital exhibition, it's important to consider the necessary data based on the specific themes or objectives of the display. For example, if the goal is to showcase the production process of cultural capital, relevant data related to that process should be used. Or, if the focus is on the origins of cultural capital, historical data will be essential. Additionally, the characteristics of the objects being displayed must be taken into account, such as their size, shape, weight, source, and preservation methods. Information on the history, significance, and usage of cultural capital within the community, along with its connection to places and traditional technologies, should also be included. All of this data will help in designing an exhibition that effectively tells a story and determines the optimal layout of the display space." | 1. Context of cultural capital<br>2. Characteristics of objects<br>3. Cultural significance<br>4. Exhibition design | is part of | Cultural capital |

community and cultural capital data required for practical applications, and the level of necessity for community and cultural capital data. Data collection involved in-depth interviews conducted by the researchers from July to August 2024 over a two-month period.

### *3.4.3. Data Analysis*

The data analysis in this study employed content analysis methods, which involve categorizing, analyzing, and describing the data based on key themes that align with the research objectives. The data obtained from the interviews were recorded in MS Word. Subsequently, the researchers coded the data by assigning codes to sentences or paragraphs that were considered significant. This coding process facilitates the efficient extraction of information relevant to the research questions. This includes the coding of conceptual variables, grouping of terminology, analysis of relationships, and identification of key themes for further study. For example, an interview question regards the cultural capital data necessary for collecting information for the benefit of using community data (QA1: "What 'cultural capital data' is necessary to be collected for the benefit of using community data?").

The data coding process begins prior to data collection by creating initial code based on the established theoretical framework. Subsequently, these codes are revised and updated to more appropriate ones, with some irrelevant codes discarded. These codes are assigned meanings and recorded in the field notes. Once the data or groups of terms are categorized, the next step is to identify relationships and assign codes according to the examples in the data classification table. This process ensures that the analysis can be conducted effectively and accurately. An example is illustrated in Table 6.

After the researchers collected and refined the preliminary data, these data were summarized and analyzed in alignment with the research objectives using two methods of analysis: descriptive analysis and content analysis. Descriptive analysis was employed to provide an overview of the fundamental characteristics of the data, while content analysis was used for in-depth examination by categorizing and interpreting the data to extract key findings and

 https://doi.org/10.1633/JISTaP.2025.13.3.3

conclusions. The results of the data analysis were presented in tabular form to systematically display the key information and research findings. The analysis was explained using descriptive methods to ensure clarity in the presentation of the data and commentary. The combination of both methods will enhance the comprehensiveness and accuracy of the analysis, facilitating effective communication of the research findings in a clear and accessible manner.

The researchers further examined and synthesized content from documents, textbooks, manuals, and research related to cultural arts and local wisdom. This helped design the structure of the community dataset and the cultural capital data that need to be collected for practical use. The researchers then asked the experts to assess the necessity level of the community and cultural capital data once again, in order to prioritize or measure the importance of these data. This process contributed to obtaining more diverse and comprehensive results.

To evaluate the necessity level of community and cultural capital data, a 5-point rating scale was employed, ranging from most necessary, very necessary, moderately necessary, slightly necessary, to least necessary. The calculation approach involved using the average scores from the ratings provided by each expert group. Software such as Excel was used to compute the average and determine the frequency for statistical analysis.

## 4. RESEARCH RESULTS

### 4.1. Community Datasets Required for Developing CIS

The analysis of the community data required for collection, as identified by six experts, reveals the essential data to be collected in common, as shown in Table 7.

Based on Table 7, the community data required for collection and utilization consists of the following categories.

1) Spatial information, including geographic data, economic data, social data (such as community groups and social organizations), and cultural data (local traditions and customs, religious activities and beliefs, and local wisdom holders).

2) Community organization information, comprising (1) core function organizations, which are the primary government agencies directly responsible for the area, such as LAOs (subdistrict administrative organizations, subdistrict municipalities, etc.), and (2) area function organizations related to the community, which may include

**Table 7.** Community data

| Community data | Case | | | | | | Total |
|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | |
| 1. Information on community organizations, categorized into at least two types:<br>1) Core function organization: The main governmental organization responsible for the area (such as subdistrict administrative organization, subdistrict municipality, or others). Only one core function organization should be identified<br>2) Area function organizations: Other organizations related to the community, including public, private, civil society, and community-based organizations. More than one related organization can be included | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ | 5 |
| 2. Spatial information, including:<br>1) Geographic data<br>2) Economic data<br>3) Social data (social groups and organizations)<br>4) Cultural data (local traditions and customs, religious activities and beliefs, and local wisdom holders) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | 6 |
| 3. Information on community capital and potential, such as individuals, groups, agencies, organizations, local resources, and physical and social infrastructure | ✗ | ✗ | ✓ | ✗ | ✓ | ✓ | 3 |
| 4. Data for establishment and development purposes (e.g., for the creation of community museums), such as building location, surrounding environment, site accessibility, boundary definitions, budget information, and community needs | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ | 3 |

Case A-F refers to the identification codes of the experts who provided the information during the interviews. These codes are used to reference the experts involved in each case. Check mark (✓) indicates that an answer related to the specific issue was found from the interview with the respective expert. Cross mark (✗) indicates that no answer related to the specific issue was found from the interview with the respective expert.

public agencies, private organizations, civil society groups, and community-based organizations.

3) Community capital and potential information, including data on individuals, groups, agencies, organizations, local resources, and both physical and social infrastructure.

4) Information for development and construction purposes, such as data on building locations, surrounding environments, site accessibility, boundary definitions, budget allocation, and the specific needs of the community (e.g., for establishing community museums).

The analysis of cultural capital data required for collection and utilization, based on the perspectives of six experts, indicates that the necessary data to be collected are consistent across all experts, as shown in Table 8.

From Table 8, it was found that the cultural capital data required for collection and utilization consists of the following:

1) Tangible cultural capital, such as historical sites, monuments, and artifacts.

2) Intangible cultural capital, including traditions, local craftsmanship, and folk arts.

3) Information on individuals involved in cultural heritage, such as key cultural figures, experts, and successors.

4) Detailed in-depth information for reference purposes, including historical records from government sources, community legends, and oral histories from local wisdom keepers.

5) Natural heritage, such as landscapes, natural tourist attractions, and native plants and animals.

6) Information on cultural management and support, such as conservation, restoration, promotion, dissemination, financial assistance, and policy support.

7) Community valuation of cultural capital in various dimensions, which may include belief systems related to the cultural capital and its relationship with the community.

8) Potential benefits for extension and application, including commercial, professional, educational, or other uses.

Results of the development and evaluation on the necessity level of the 15 community datasets revealed that four datasets were rated as most necessary for the CIS for managing cultural capital, as follows: Individual data (e.g., community leaders/local scholars) ($\overline{\chi}$=4.69); Organizational data ($\overline{\chi}$=4.63); Area context data–community history ($\overline{\chi}$=4.50); and Community map data ($\overline{\chi}$=4.50). The remaining 11 datasets were rated as very necessary, including: Community location data ($\overline{\chi}$=4.48); Local project data ($\overline{\chi}$=4.42); Local personnel data ($\overline{\chi}$=4.36); LAO data ($\overline{\chi}$=4.31); Area context data–population size ($\overline{\chi}$=4.30); Activity data ($\overline{\chi}$=4.22); Program or service data ($\overline{\chi}$=4.20); Area context data–area size and administrative boundaries

**Table 8.** Cultural capital data

| Cultural capital data | Case | | | | | | Total |
|---|---|---|---|---|---|---|---|
| | A | B | C | D | E | F | |
| 1. Tangible cultural capital, such as historical sites and artifacts | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | 6 |
| 2. Intangible cultural capital, such as traditions, craftsmanship, and local arts | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | 6 |
| 3. Potential benefits for extension in commerce, careers, education, or other areas | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ | 2 |
| 4. Natural heritage, such as landscapes, natural tourist attractions, and native flora and fauna | ✗ | ✓ | ✗ | ✗ | ✓ | ✓ | 3 |
| 5. Information on individuals involved in cultural heritage, such as cultural experts, key persons and successors | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ | 5 |
| 6. Information on cultural management and support, such as conservation, restoration, promotion, dissemination, financial support, and policy | ✗ | ✓ | ✗ | ✗ | ✓ | ✓ | 3 |
| 7. In-depth data collection details (for reference purposes), such as historical data from official records, community legends, and oral histories from local wisdom keepers (key persons) | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ | 5 |
| 8. Community valuation of cultural capital in various dimensions, including belief systems and the relationship between cultural capital and the community | ✗ | ✗ | ✗ | ✓ | ✓ | ✓ | 3 |

Case A-F refers to the identification codes of the experts who provided the information during the interviews. These codes are used to reference the experts involved in each case. Check mark (✓) indicates that an answer related to the specific issue was found from the interview with the respective expert. Cross mark (✗) indicates that no answer related to the specific issue was found from the interview with the respective expert.

 https://doi.org/10.1633/JISTaP.2025.13.3.3

($\overline{\chi}$=4.61); Community slogan data ($\overline{\chi}$=3.92); and Budget management data ($\overline{\chi}$=3.67), respectively. In total, the 15 community datasets are comprised of 117 data elements. The required community datasets are presented in Table 9.

### 4.2. Cultural Capital Datasets Required for Developing CIS

Results of the development and evaluation of the necessity level of the cultural capital datasets revealed that

**Table 9.** Community datasets required for developing a community cultural capitals information system

| No. | Community datasets | Number of data elements | Means | Necessity level | Standard deviation |
|---|---|---|---|---|---|
| 1 | Local Administrative Organization (LAO) data: LAO code; LAO name; LAO type; house number; street name; village name; subdistrict name; district name; province name; postal code; location latitude; location longitude; Northern bordering; Southern bordering; Eastern bordering; Western bordering; responsibility area; historical background; geographical features; climatic conditions; LAO images; LAO website; contact information | 23 | 4.31 | Very necessary | 0.63 |
| 2 | Budget management data: Budget code; budget name; budget type/category (categorized according to relevant laws) | 3 | 3.67 | Very necessary | 1.54 |
| 3 | Local project data: Project code; budget name; project name; plan code; plan volume; plan page; ordinance/regulation code; ordinance/regulation volume; ordinance/regulation page; department code; LAO code; LAO name; project type | 13 | 4.42 | Very necessary | 0.76 |
| 4 | Local personnel data: Citizen ID number; first name; last name; title; gender; date of birth; education level; personnel type; position number; position name; term end date; personnel level; department code; department name; department type; LAO code; LAO name | 17 | 4.36 | Very necessary | 0.84 |
| 5 | Community location data: Community code; village number; village/community name; subdistrict name; district name; province name; latitude; longitude | 8 | 4.48 | Very necessary | 0.55 |
| 6 | Area context data: Community history: community history code; community history | 2 | 4.50 | Most necessary | 1.01 |
| 7 | Area context data–Area size and administrative boundaries: area code; Northern boundary; Southern boundary; Eastern boundary; Western boundary; topographical features; administrative divisions (or community/village names within the administrative area); area size (sq. km.); transportation and accessibility (e.g., public transport routes) | 9 | 3.96 | Very necessary | 0.92 |
| 8 | Area context data–Population: Area code; male population; female population; number of households; primary, secondary, and supplementary occupations | 5 | 4.30 | Very necessary | 0.73 |
| 9 | Community slogan data: Slogan code; community slogan | 2 | 3.92 | Very necessary | 1.13 |
| 10 | Community map data: Map code; community map | 2 | 4.50 | Most necessary | 1.01 |
| 11 | Individual data (e.g., community leaders or local scholars): Individual code; name; address; phone number; expertise; notes | 6 | 4.69 | Most necessary | 0.55 |
| 12 | Organization data: Organization code; organization name; organization address; phone number; fax number; email address; location; operating hours; specific information/facilities/services provided | 9 | 4.63 | Most necessary | 0.55 |
| 13 | Program or service data: Service code; service name; service provider; location; specific information/facilities/services provided | 5 | 4.20 | Very necessary | 0.69 |
| 14 | Activity data: Activity code; activity name; activity venue; date; time; responsible person; notes on funding source; budget; activity details (1.07) | 9 | 4.22 | Very necessary | 0.88 |
| 15 | Other data may involve community-related operations and consists of 4 data elements: Data code, data name, data address, notes | 4 | 3.58 | Very necessary | 1.07 |
| | Total number of data elements | 117 | | | |

among the two main datasets, intangible cultural datasets were rated at the most necessary level ($\overline{\chi}$=4.4.64), while the tangible cultural dataset was rated at a very necessary level ($\overline{\chi}$=4.37) for the CIS for managing cultural capital. For the intangible cultural datasets, they are divided into five subsets. It is revealed that all five subsets were rated

**Table 10**. Cultural capital datasets required for developing a community cultural capitals information system

| No. | Cultural capital datasets | Number of data elements | Means | Necessity level | Standard deviation |
|---|---|---|---|---|---|
| 1 | Tangible cultural heritage data (in referenced to CDWA Standards): Type and quantity of work; classification, name; creation; form of the work, dimensions; materials used in creation; marks or inscriptions on materials; relationships among multiple created works; edition; construction methods or specific techniques used; description of display or presentation methods; description of the work's appearance; physical condition assessment; conservation or restoration; artistic subject matter; object context; descriptive text or explanation of the work; critical commentary on the work; related artworks or architecture; current location and geographic location of the artwork or architecture; individuals or groups holding usage, display, or reproduction rights; provenance or ownership history from creation to present; exhibition or loan details; documentation of creation and description modifications; related documentation; references to textual sources related to the artwork or architecture; information about artists, architects, or other relevant individuals; information about geographical locations significant to the artwork, architecture, or creator; general concepts necessary for cataloging or describing the work (e.g., object type, materials, activities, styles); information about characters, animals, themes, stories, literature, myths, religions, or historical or fictional events mentioned by name | 31 | 4.37 | Very necessary | 0.58 |
| 2 | Intangible cultural heritage data | 93 | 4.64 | Most necessary | 0.49 |
| 2.1 | Narrative data (e.g., folk legends - oral tradition & expression) consists of 17 data elements: Identifier; name; creator; contributors; description; keywords; characters; morals; ethnic group; location; relationships; country; language; media; sources; date; rights | 17 | 4.53 | Most necessary | 0.53 |
| 2.2 | Performing arts data consists of 20 data elements: Name; history; country/region of origin; time of origin; dance styles; musical concepts; lyrics; dance annotations; description; GPS coordinates; performance duration; software; location; objects and descriptions; explanations; gender; facial expressions; traditional costume description; number of performers+performer attributes | 20 | 4.61 | Most necessary | 0.45 |
| 2.3 | Social practices, rituals, and festivals data consists of 13 data elements: Tradition/local name; month of celebration; frequency/scale of event; lunar calendar; purpose; activities; rituals; literature; beliefs; location; participants; equipment; buildings | 13 | 4.73 | Most necessary | 0.45 |
| 2.4 | Knowledge and practices related to nature and the universe data consists of 15 data elements: Title of the work; work code; creator; description; language; ethnicity; date and time; supplementary media; keywords; copyright owner; habitat/location; categories of cultural and intellectual knowledge; Thai name; English name; local name | 15 | 4.61 | Most necessary | 0.55 |
| 2.5 | Traditional craftsmanship data consists of 28 data elements: Name; keywords; craft code; category; owner; executor; affiliated nation; generational age; production area; modeling; dimensions; weight; color; design; materials; process forms (production process); ornaments; physical qualities and evaluation; user groups; usage occasions; cultural implications; technology level; values; rights; data collectors; time of data collection; physical storage location; image storage location | 28 | 4.74 | Most necessary | 0.47 |
| Total number of cultural capital data elements | | 124 | | | |

CDWA, Categories for the Description of Works of Art.

 https://doi.org/10.1633/JISTaP.2025.13.3.3

at the most necessary level, respectively, as follows: Traditional craftsmanship data ($\overline{\chi}$=4.74); Social practices, rituals, and festivals data ($\overline{\chi}$=4.73); Performing arts data ($\overline{\chi}$=4.61); Knowledge and practices related to nature and the universe data ($\overline{\chi}$=4.61); and Narrative data ($\overline{\chi}$=4.53). In total, the two cultural capital datasets comprise 124 data elements. The required community datasets are presented in Table 10.

## 5. DISCUSSION

The findings of this research on developing a community cultural capital information system underscore the critical importance of structured data management in nurturing local cultural assets, aligning closely with the principles of CI as discussed in the literature. As noted by Gurstein (2007), CI emphasizes the role of ICT in fostering community participation and development. Our study's qualitative approach, engaging local cultural practitioners and experts, embodies this principle by grounding the developed datasets in the lived realities and needs of the community.

The delineation of community data and cultural capital data reflects the theoretical framework established in the literature, which identifies various dimensions of cultural capital, including tangible and intangible aspects (Bourdieu, 1986; Griswold, 2004). In this study, the categorization into intangible cultural capitals information, tangible cultural capitals information, and administrative cultural capitals information mirrors the need for a multifaceted understanding of cultural resources, as identified in earlier research. This classification allows for nuanced data analysis that can directly inform policy making and community initiatives aimed at sustainable development, resonating with the findings of UNESCO (2011) regarding the preservation of cultural heritages' importance.

Moreover, the study supports the assertion that the effective management of cultural capital contributes to social cohesion and local economic development, which has been a recurrent theme in the literature (Department of Cultural Promotion, 2019). The engagement of community members in the evaluation process not only enhances the relevance of the cultural capital datasets but also builds a participatory knowledge base, echoing the sentiments expressed by Auwattanamongkol (2019) on the significance of involving stakeholders in data governance and utilization.

The researchers designed a dataset for cultural capital aimed at sustainable development by applying international metadata standards for museums (CDWA) to describe cultural documents and objects related to tangible cultural capital. This was combined with the adaptation of metadata in related research for designing datasets on traditions, culture, and local wisdom, including (1) a *Metadata construction scheme of a traditional clothing digital collection* (Ye, 2023); (2) *Development of Thai culture ontology-based metadata: Heet Sib Song* (Chansanam & Tuamsuk, 2022); (3) a *Metadata schema for folktales in the Mekong River Basin* (Kwiecien et al., 2021); (4) *Metadata for intangible cultural heritage - the case of folk dances* (Giannoulakis et al., 2018); and (5) *Local culture and wisdom metadata* (Nachom, 2018). Our application of these metadata standards, utilized to structure the datasets, resonates with the literature's recommendation for clear data categorization to enhance accessibility and usability (Toubekis et al., 2011). By integrating these standards, the research ensures that essential cultural and community data can be effectively shared and operationalized across various platforms, thereby enhancing transparency and fostering collaborative approaches to cultural capital management.

While the primary focus of this study is on Thailand, the proposed dataset framework and CIS model offer valuable insights that may inform similar efforts in other contexts. By enabling LAOs to systematically manage and utilize cultural capital, the study provides a scalable approach that could inspire regional and national strategies for cultural heritage preservation and policy development. However, it is important to recognize that cultural capital is inherently shaped by unique historical, social, and cultural contexts. As such, while the structure and methodology of the dataset may be adaptable, its direct application in other countries would require careful contextualization to reflect local cultural norms, values, and institutional frameworks. This highlights the importance of flexibility and cultural sensitivity when transferring or scaling such systems beyond their original setting.

## 6. CONCLUSION

This study presents a foundational framework for the development of essential datasets to support a CIS in Thailand. The research underscores the value of integrating community data with cultural capital data to enable more effective and context-sensitive management of cultural assets. By structuring these datasets, community data, and cultural capital data, the study provides a practical tool that can be employed by local government

agencies, cultural practitioners, community members, and policymakers to inform decision-making, resource allocation, and local development planning.

Key findings highlight the importance of a dual-data approach: Community data offers insight into demographic, social, and economic characteristics, while cultural capital data documents heritage, traditions, and other local cultural assets. The integration of these data types enhances the ability of stakeholders to design development projects that are aligned with local values and conditions, supporting sustainable growth and cultural conservation. Furthermore, the inclusion of metadata standards facilitates data interoperability across platforms, fostering inter-community collaboration and the dissemination of best practices.

The research also affirms the central role of cultural capital in advancing sustainable development and social cohesion, particularly in the context of increasing globalization. By adopting a data-driven approach to cultural resource management, communities can strengthen local identity, increase participation in cultural initiatives, and build resilience against external pressures. This study contributes to the broader discourse on CI and cultural capital by bridging theoretical frameworks with practical implementation strategies. It offers empirical evidence supporting the use of structured data in strengthening local governance and cultural sustainability.

To maximize the impact of this research, it is recommended that LAOs and cultural institutions adopt and apply the proposed dataset structure. Capacity-building initiatives, such as training programs for community stakeholders, should be developed to ensure effective data collection, interpretation, and application. In addition, continuous monitoring and evaluation of the CIS should be undertaken to ensure its relevance and responsiveness to evolving community dynamics.

Future research should examine the long-term outcomes of implementing CIS frameworks, with particular attention to their influence on community development, policy integration, and cultural vitality. Studies could also explore the role of digital technologies—such as mobile applications, geospatial mapping, and participatory platforms—in enhancing the accessibility and utility of cultural data. Lastly, while this study is contextually grounded in Thailand, future comparative research could investigate the adaptability of the framework in other cultural settings, taking into account local variations in social structures, governance systems, and cultural values.

The limitations of this study include the limited scope of community data collection, where certain groups may not have participated in providing information, resulting in data that does not fully reflect the actual needs of the community. The cultural diversity of each community poses challenges in developing a data structure that comprehensively covers all aspects. Furthermore, stakeholder participation may be incomplete, affecting the practical application of the data as expected. Additionally, technological limitations, such as a lack of modern technology or insufficient training in certain areas, may hinder the appropriate use of data management systems within the context of each community.

## 7. ETHICAL CONSIDERATIONS

The project has been reviewed and approved by the Khon Kaen University Ethics Committee for Human Research based on the Belmont Report and Good Clinical Practice in Social and Behavioral Research (Date of approval: 21 September 2024; Reference No.: HE653236).

## CONFLICTS OF INTEREST

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

## REFERENCES

Anontachai, C., Tuamsuk, K., & Chansanam, W. (2025). Cultural capital information management and sustainability in the context of Thailand local government organizations. *Journal of Posthumanism*, 5(3), 117-137. https://doi.org/10.63332/joph.v5i3.717

Auwattanamongkol, S. (2019). *Data mining* (2nd ed.). National Institute of Development Administration Publishing.

Boonto, P. (2025). Development of digital cultural heritage management: Lessons and guidelines from local museums in Chiang Mai Province. *Journal of Arts and Thai Studies*, 47(1), E4065 (4061-4021). https://so08.tci-thaijo.org/index.php/artssu/article/view/4065

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241-258). Greenwood. https://home.iitk.ac.in/~amman/soc748/bourdieu_forms_of_capital.pdf

Bunch, A. (1993). *The basics of community information work* (2nd ed.). Library Association.

Chaichuay, V. (2018). Cultural heritage information: Concept and research issues. *Journal of Information Science Research and Practice*, 35(2), 130-153. https://so03.tci-thaijo.

org/index.php/jiskku/article/view/108375

Chansanam, W., & Tuamsuk, K. (2022). Development of Thai culture ontology-based metadata: Heet Sib Song (Twelve Months Festival). *Humanities, Arts and Social Sciences Studies*, 22(3), 725-737. https://doi.org/10.14456/hasss.2022.62

databricks. (n.d.). *What is a dataset?* https://www.databricks.com/glossary/what-is-dataset

Davis, S., Palmer, L., & Etienne, J. (2016). The geography of digital literacy: Mapping communications technology training programs in Austin, Texas. In B. Passarelli, J. Straubhaar, & A. Cuevas-Cerveró (Eds.), *Handbook of research on comparative approaches to the digital age revolution in Europe and the Americas* (pp. 371-384). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-4666-8740-0.ch022

Department of Cultural Promotion. (2019). *Culture, value to value*. https://www.culture.go.th/culture_th/ewt_news.php?nid=3972&filename=index

Department of Local Administration. (2019). [*User manual for the local administration organizations system*]. Thai. https://www.nongkratumkorat.go.th/pdf/16464852061.pdf

Department of Local Administration. (n.d.). *Public health and environment*. https://info.dla.go.th/onepage/info07.jsp

Eze, P. U., Ezenkwu, C. P., & Etteh, C. C. (2021). Community informatics for sustainable management of pandemics in developing countries: A case study of COVID-19 in Nigeria. *Ethics, Medicine and Public Health*, 16, 100632. https://doi.org/10.1016/j.jemep.2021.100632

Giannoulakis, S., Tsapatsoulis, N., & Grammalidis, N. (2018). Metadata for intangible cultural heritage - the case of folk dances. *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 5, 634-645. https://doi.org/10.5220/0006760906340645

Grisales Bohorquez, C., Ruan, L., & Williams, K. (2022). How librarians and firefighters built a special library in Champaign, Illinois, USA: A community informatics story. *Digital Transformation and Society*, 2(1), 42-59. https://doi.org/10.1108/dts-08-2022-0035

Griswold, W. (2004). *Cultures and societies in a changing world* (2nd ed.). Pine Forge Press.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough?: An experiment with data saturation and variability. *Field Methods*, 18(1), 59-82. https://doi.org/10.1177/1525822X05279903

Gurstein, M. (2007). *What is community informatics (and why does it matter)?* Polimetrica. https://doi.org/10.48550/arXiv.0712.3220

Hawkes, J. (2001). *The fourth pillar of sustainability: Culture's essential role in public planning*. Common Ground Publishing. https://www.culturaldevelopment.net.au/downloads/FourthPillarSummary.pdf

Huang, Y., Tu, R., Tuerxun, W., Jia, X., Zhang, X., & Chen, X. (2024). A community information model and wind environment parametric simulation system for old urban area microclimate optimization: A case study of Dongshi Town, China. *Buildings*, 14(3), 832. https://doi.org/10.3390/buildings14030832

Janosi, A., Steinbrunn, W., Pfisterer, M., & Detrano, R. (1989). *Heart disease*. https://archive.ics.uci.edu/ml/datasets/heart+disease

Kelleher, J. D., & Tierney, B. (2018). *Data science*. MIT Press.

Khan, M., & Hanna, A. (2022). *The subjects and stages of ai dataset development: A framework for dataset accountability*. https://dx.doi.org/10.2139/ssrn.4217148

Kwiecien, K., Chansanam, W., Supnithi, T., Chitiyaphol, J., & Tuamsuk, K. (2021). Metadata schema for folktales in the Mekong River Basin. *Informatics*, 8(4), 82. https://doi.org/10.3390/informatics8040082

Library of Congress. (2006). *MARC 21 community information*. https://www.loc.gov/marc/community/ciintro.html

Lippeveld, T. (2017). Routine health facility and community information systems: Creating an information use culture. *Global Health: Science and Practice*, 5(3), 338. https://doi.org/10.9745/GHSP-D-17-00319

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. John Wiley & Sons.

Lopez, C. (2015). *Modeling sustainability of participatory information systems for urban communities: A mixed-method approach* (doctoral dissertation). University of Pittsburgh, Pittsburgh, PA, USA.

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research*, 26(13), 1753-1760. https://doi.org/10.1177/1049732315617444

Mehra, B., Singh, V., Hollenbach, N., & Partee, R. P., II. (2017). Rural librarians as change agents in the twenty-first century: Applying community informatics in the Southern and Central Appalachian region to further ICT literacy training. In B. Real (Ed.), *Rural and small public libraries: Challenges and opportunities* (Vol. 43, pp. 123-153). Emerald Publishing Limited.

Nachom, C. (2018). Local culture and wisdom metadata. *PULINET Journal*, 5(1), 304-315. https://pulinet.oas.psu.ac.th/index.php/journal/article/viewFile/306/307

Nonthacumjane, P., & Johansson, V. (2023). Shaping local information in Thailand: Hidden contradictions in the digitisation activities of the Provincial University

Library Network (PULINET). *Journal of Librarianship and Information Science*, 55(1), 246-258. https://doi.org/10.1177/09610006221143101

OpenLearn Create. (n.d.). *Health management, ethics and research module: 11. Developing your community profile*. https://www.open.edu/openlearncreate/mod/oucontent/view.php?id=229

Ozuomba, S., Chukwudebe, G., Opara, F., & Ndinechi, M. (2013). Preliminary context analysis of community informatics social network web application. *Nigerian Journal of Technology*, 32(2), 266-272. https://www.ajol.info/index.php/njt/article/view/123594

Pandey, R., & Kumar, V. (2023). A survey of the metadata element sets used for digital art objects in the online collections of the museums of India. *DESIDOC Journal of Library & Information Technology*, 43(4), 226-233. https://www.semanticscholar.org/paper/A-Survey-of-the-Metadata-Element-Sets-Used-for-Art-Pandey-Kumar/e954443a80a87a307d7eb50401b24059af958885

Saensrisawat, A., Chalotham, N., Dechchit, P., Wongtan, T., Bunma, P., & Kittikarnnakorn, N. (2024). [Analysis of standard data sets to develop a family medicine data linkage system]. *Thai Journal of Pharmacy Practice*, 16(4), 941-953. Thai. https://he01.tci-thaijo.org/index.php/TJPP/article/view/263740

Salse, M., Guallar-Delgado, J., Jornet-Benito, N., Mateo Bretos, M. P., & Silvestre-Canut, J. O. (2022). GLAM metadata in museums and university collections: A state-of-the-art (Spain and other European countries). *Global Knowledge, Memory and Communication*, 73(4-5), 477-495. https://doi.org/10.1108/gkmc-06-2022-0133

Stillman, L., & Linger, H. (2009). Community informatics and information systems: Can they be better connected? *The Information Society*, 25(4), 255-264. https://doi.org/10.1080/01972240903028706

Toubekis, G., Cao, Y., Klamma, R., Spaniol, M., Jarke, M., & Jansen, M. (2011). ACIS: A collaborative hypermedia geospatial community information system of Afghan monuments and sites. In: E. Jerem, F. Redő, & V. Szeverényi (Eds.), *Proceedings of the 36th International Conference* (pp. CD-ROM 550-558). Computer Applications and Quantitative Methods in Archaeology (CAA).

Trust, J. P. G. (2024). *Categories for the description of works of art*. https://www.getty.edu/publications/categories-description-works-art/

UC Merced Library. (n.d.). *What is a data dictionary?* https://library.ucmerced.edu/data-dictionaries

UNESCO Institute for Statistics. (n.d.). *Cultural heritage*. https://uis.unesco.org/en/glossary-term/cultural-heritage

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2011). *2003 Convention for the safeguarding of the intangible cultural heritage*. https://ich.unesco.org/doc/src/15164-EN.pdf

van der Waldt, G. (2019). Community profiling as instrument to enhance project planning in local government. *African Journal of Public Affairs*, 11(3), 1-21. https://journals.co.za/doi/abs/10.10520/EJC-19603347f5

Ye, Q. (2023). Metadata construction scheme of a traditional clothing digital collection. *The Electronic Library*, 41(4), 367-386. https://doi.org/10.1108/el-01-2023-0004

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

RESEARCH PAPER
J Inf Sci Theory Pract 13(3): 55-71, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.3.4

# The Effect of Institutional Pressures on the Intention to Adopt Environmental Disposal of Electronic Waste: Attitude as the Mediating Variable

**Ghaith Abdulridha Mubdir**
Faculty of Economics and Management, Universiti Kebangsaan Malaysia (UKM), Bangi, Malaysia
Faculty of Administration and Economics, University of Thi-Qar, Thi-Qar, Iraq
E-mail: P104676@siswa.ukm.edu.my

**Sharizal Hashim***
Faculty of Economics and Management, Universiti Kebangsaan Malaysia (UKM), Bangi, Malaysia
E-mail: hsharizal@ukm.edu.my

**Abdul Aladheem Jabbar Alziady**
Faculty of Administration and Economics, University of Thi-Qar, Thi-Qar, Iraq
E-mail: dr.abdaladheem@yahoo.com

**Sabah Hussein Enayah**
Faculty of Science, University of Thi-Qar, Thi-Qar, Iraq
E-mail: sabah792007@yahoo.com

## ABSTRACT

The rising volume of electronic waste generated by public organisations in the Thi-Qar province of Iraq adversely affects the environment due to inadequate recycling and disposal methods, thereby intensifying pressures on enterprises. This research seeks to identify the factors influencing organisational managers' adoption of environmentally sustainable e-waste disposal methods. A research model was established to connect three components: coercive pressure, normative influence, and mimicry. It investigates the impact of adopting environmentally responsible e-waste disposal and analyses the mediating role of attitude. The model underwent validation through data collected from a field survey involving 302 managers of public organisations in the Thi-Qar province. A questionnaire was created to gather data, consisting of five primary variables represented by twenty-one items. The seven-dimensional scalar form was employed for measurement and underwent reliability and validity assessments. Partial least squares were utilised to analyse the survey responses. The findings demonstrated that coercive, mimetic, and normative pressures significantly influenced attitudes and the adoption of environmentally responsible e-waste disposal practices. Attitudes regarding the environmental disposal of e-waste influence the adoption of eco-friendly disposal methods and mediate the effects of coercive, mimetic, and normative pressures on this adoption. This study advances the green information technology literature by identifying critical institutional factors influencing environmental behaviour and provides practical recommendations for policymakers and organisations to improve e-waste management strategies.

**Keywords:** institutional pressures, environmental disposal of e-waste, coercive pressures, mimetic pressures, normative pressures

**Received:** February 12, 2025 **Revised:** April 19, 2025
**Accepted:** May 21, 2025 **Published:** September 30, 2025

***Corresponding Author:** Sharizal Hashim
https://orcid.org/0000-0003-1287-2490
**E-mail:** hsharizal@ukm.edu.my

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 (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

© 2025 [Ghaith Abdulridha Mubdir, Sharizal Hashim, Abdul Aladheem Jabbar Alziady, Sabah Hussein Enayah] This article is distributed under the Creative Commons Attribution License (CC BY), allowing unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. It is published by the Korea Institute of Science and Technology Information (KISTI).

## 1. INTRODUCTION

The advancement of technology has resulted in the extensive use of electrical and electronic devices, leading to significant e-waste composed of metals, plastics, rubber, glass, and ceramics (Mat Nawi et al., 2024). Public organisations in Thi-Qar province, Iraq, generate significant amounts of e-waste annually, which is disposed of in landfills, resulting in the leakage of harmful chemicals (Alziady & Enayah, 2019). E-waste comprises hazardous materials such as lead, mercury, arsenic, cadmium, selenium, and chromium, complicating its disposal. Harmful metals are associated with malignancies, miscarriages, neurological damage, blood disorders, and dioxins that adversely affect the central nervous system (Regel-Rosocka, 2018). E-waste management in Iraq presents challenges due to environmental, technical, economic, and health factors during treatment and recycling processes, resulting in significant exposure to lead, mercury, cadmium, and arsenic. Public organisations in Thi-Qar generate significant amounts of e-waste due to increasing domestic wealth and declining product prices post-2003 (Ikhlayel, 2018). A communication and media commission indicated that these organisations acquired 695,000 computers in 2018, along with numerous electronic devices, including TVs, phones, DVD players, printers, satellite receivers, remote controls, mobile chargers, copy machines, lighting equipment, microwaves, ovens, air conditioners, and others on an annual basis (Mat Nawi et al., 2024). Cities produce a significant amount of electrical and electronic waste, one of the fastest-growing sources of waste. The disposal of abandoned devices poses significant environmental and health risks due to their complex composition of materials (Kumar et al., 2017; Sharma et al., 2024).

Consequently, hazardous e-waste processing that poses risks to human health and the environment has raised concerns among federal governments, regulators, consumers, competitors, communities, environmental advocacy groups, business associations, researchers, and local authorities (Regel-Rosocka, 2018). Authorities have identified harmful compounds infiltrating the environment, resulting in city residents being exposed to elevated levels of lead, mercury, cadmium, and arsenic (Waghmode et al., 2021). The inadequate environmental disposal of e-waste by public organisations in Thi-Qar and the increasing volumes of e-waste constitute a significant concern. The increasing recognition of environmental degradation in recent years highlights the importance of e-waste disposal as a critical factor in attaining community satisfaction (Alziady, 2018). Recognizing the incentives that motivate organisational managers to manage e-waste environmentally is essential, particularly in regions such as Thi-Qar, where e-waste is increasing and environmental pressures are intensifying (Alziady & Enayah, 2019). Reducing e-waste accumulation enables an organisation to utilise resources more efficiently and recover secondary raw materials via recycling and other recovery methods, thereby enhancing its environmental performance throughout the product life cycle.

Public organisations have increasingly utilised electronic devices in the past decade, with newer and more powerful models replacing older versions. Consequently, numerous gadgets have a limited lifespan, making a significant quantity obsolete. E-waste generated from obsolete electronics is disposed of in landfills, posing risks to environmental integrity and human health (Ikhlayel, 2018). E-waste comprises hazardous substances that can harm the environment if not recycled properly. E-waste also comprises valuable metals such as iron, copper, aluminium, gold, and silver. To maximise resource efficiency, e-waste contains valuable raw materials (Borthakur & Govind, 2018). Administrators of public organisations must prioritise the environmental disposal of electronic waste. The issue of e-waste has heightened awareness. Methods for disposing of e-waste include reusing, remanufacturing, recycling, incineration, and landfilling (Echegaray & Hansstein, 2017). This represents an innovative approach for organisations to adhere to environmental regulations. Public sector managers may leverage environmental knowledge to reconcile economic, social, and environmental performance.

The significant increase in e-waste in Thi-Qar prompts an examination of public organisation managers' perspectives on environmental disposal and their willingness to implement appropriate e-waste management practices and disposal techniques. E-waste poses significant environmental pollution risks and potential health hazards. Consequently, managers' perceptions regarding the hazardous nature of corporate e-waste may influence the relationship between institutional pressures and the inclination to dispose of e-waste in an environmentally responsible manner. A recent study from Iraq has highlighted a concerning situation (Mat Nawi et al., 2024). The quantity of discarded e-waste remains uncertain, necessitating further investigation into the environmental consequences of institutional pressures on managers' intentions to dispose of e-waste in an environmentally responsible manner and the mediating influence of attitude. This study aims to analyse

 https://doi.org/10.1633/JISTaP.2025.13.3.4

and understand the significant challenges associated with e-waste. The research structure includes the introduction, research question framework, objectives, theoretical background, methods, findings, discussion, and conclusions.

Post-2003, there was a global increase in economic development and public sector purchasing power, which led to a rise in the utilisation of computers, laptops, televisions, DVDs, printers, copy machines, and mobile phones. The volume of e-waste has significantly increased. Improper e-waste disposal, including municipal solid waste and open burning, has resulted in environmental pollution; however, informal recycling remains prevalent in cities. Consequently, numerous e-waste components are disposed of in open landfills and dumps. As a result, administrators of public organisations in Thi-Qar have acknowledged the gravity of e-waste and have indicated a commitment to its environmentally responsible disposal. Research indicates that institutional constraints, including coercive pressure, normative influence, and imitation, impact the adoption and dissemination of environmental practices among organisational managers (Lin & Ho, 2016). This present research investigates the impact of institutional factors on intention and attitude as mediating variables. The following questions are examined.

Do institutional influences impact public organisation managers' willingness to dispose of e-waste in an environmentally responsible manner? Do the same institutional powers affect the adoption of e-waste? What is the role of attitude in mediating institutional pressures and the inclination of managers to dispose of e-waste in an environmentally responsible manner?

## 2. LITERATURE REVIEW

E-waste represents a multifaceted category of waste, characterised by its material heterogeneity and diverse product lifecycles. The challenge in identifying the material type and product lifecycle indicates that e-waste and its associated commodity chains are complex. After their useful life, electronics become e-waste. The United States Environmental Protection Agency (EPA) identifies e-waste as used electronics, recognizing that recyclable products reduce improper disposal in unmanaged landfills. E-waste includes any electrical and electronic equipment dumped by the owner as garbage without purpose of reuse. Ananno et al. (2021) define e-waste as all non-recyclable electrical and electronic gadgets. Zhang and Xu (2016)'s study found that electronic gadgets include both harmful and valuable components, making recycling ecologically and economically beneficial (Zhang & Xu, 2016). E-waste disposal must reduce or eliminate hazardous materials such as lead, cadmium, mercury, hexavalent chromium, polybrominated biphenyls, polybrominated diphenyl ethers, arsenic, and polyvinyl chloride from electronic devices. Chemicals and metals disposed of in landfills or incinerated impair the environment and human health by contaminating water, air, soil, dust, and food (Ikhlayel, 2018). These materials carry health and environmental concerns, including water and soil contamination (Sharma et al., 2024). Research shows that e-waste also includes gold, palladium, silver, indium, and rare earth metals, making recycling more beneficial.

E-waste is one of the fastest-growing issues since typical product lifespans have shortened, leading to the dumping of outmoded technology into the waste stream (Sharma et al., 2024). E-waste is a major issue with potentially serious repercussions. According to Zhang et al. (2011), widespread information technology (IT) use may cause e-waste-related environmental catastrophes. About 500 million PCs died between 1994 and 2003. Five hundred million personal computers contain about 2,872,000 tonnes of plastics, 718,000 tonnes of lead, 1,363 tonnes of cadmium, and 287 tonnes of mercury, posing serious environmental and public health risks if not disposed of properly. Around 60 million tonnes (Mt) of e-waste are produced globally, with some ending up in landfills, creating a significant risk of hazardous chemical leakage. Televisions, computers, cell phones, printers, scanners, and faxes weighed 2.37 Mt in 2009, according to the EPA. Recycling was applied to just 25% of these gadgets, with the rest going to landfills.

Regel-Rosocka (2018) reported that The European Union created 9 Mt of e-waste in 2005 and may produce 12 Mt by 2020. In 2016, the United Nations anticipated 44.7 Mt of e-waste, rising to 52.2 Mt by 2021 (Ananno et al., 2021). Poor direction and management cause incorrect behaviours, public health risks, and environmental hazards, especially in countries without processing facilities. Nishant et al. (2017) indicated that e-waste is projected to become increasingly problematic, with an annual growth rate of approximately 4%, making it the fastest-growing waste stream. Annually, around 40 million metric tonnes of electronic waste are generated. In 2014, e-waste generation totalled 41.8 Mt, predominantly from Asia, North America, and Europe (Dias et al., 2018). The total volume of e-waste was projected to attain 50 Mt by 2018, with an annual increase of 3% to 5% (Cucchiella et al., 2015). Europe exhibited the highest per capita waste generation, fol-

lowed by Oceania (Dias et al., 2018). E-waste constitutes 1-3% of the total global municipal waste annually (1,636 Mt); in developed nations, this proportion is roughly 5%. According to the United States Environmental Protection Agency (USEPA), global annual e-waste generation is increasing at 5-10% per year, whereas the recovery rate remains approximately 10% (Ananno et al., 2021).

While developed nations generate a greater volume of e-waste annually than developing nations, the adverse effects of e-waste are more complex in the latter context. The cost of e-waste recycling in developed countries is significantly higher, ten times that of developing countries. The export of e-waste from developed to developing countries is a significant debate and controversy, with developed nations asserting that the rationale is business, a highly inequitable stance. Furthermore, due to the insufficient experience and equipment for e-waste disposal in developing countries, authorities often resort to incineration or landfill dumping, exacerbating the issue. IT manufacturers are responding to pressures from various stakeholders to create environmentally friendly products by utilising raw materials that are not harmful to the environment (Echegaray & Hansstein, 2017). In terms of e-waste disposal, several countries have achieved notable advancements in this area; however, in several Middle Eastern countries, the implementation is still limited.

Despite the significance of e-waste, many individuals prioritise other issues as more pressing. Consequently, society continues to confront a significant challenge posed by e-waste, prompting governments to implement strategies to address the issue. When a computer malfunctions, the owner typically either disposes of the device or retains it at home; both scenarios pose risks to public health. Public organisations face pressure from customers, competitors, regulators, and society to adopt green information technology (GIT) practices (Coffey et al., 2013). Public organisations have increased their investments in IT in recent years, which serves as the primary source of e-waste. The increasing e-waste generated by such organisations has rapidly reached alarming levels; however, organisational managers have historically overlooked its detrimental environmental effects. The Thi-Qar province in Iraq lacks e-waste statistics; however, numerous computers, photocopiers, printers, and unused faxes are abandoned on shelves and streets. In Thi-Qar, most e-waste management operations are conducted by garbage collectors, encompassing collection, segregation, dismantling, and disposal activities. Many methods employed for e-waste treatment pose risks without the appropriate technologies and equipment. Additionally, inadequate waste disposal practices employed by various Thi-Qar municipalities typically involve the open burning of plastic waste, exposure to hazardous solvents, the dumping of acids, and extensive general dumping.

Therefore, a survey questionnaire was created to examine the institutional pressures influencing public organisation managers' behaviours concerning adopting environmentally sustainable e-waste disposal practices. The research included 302 organisations from diverse sectors: higher education, construction, banking, health services, manufacturing, oil and gas, agriculture, education, information and communication, insurance, water supply, sewerage management, and transportation. Emails were dispatched to the managers of the targeted organisations, containing requests for participation and accompanying instructions. Participants were selected based on prior research demonstrating that managers recognize the motivations for adopting these practices (Krell et al., 2016). Akman and Mishra (2015) observed that diverse sociodemographic characteristics among societal groups can result in differing behavioural patterns.

## 3. RESEARCH MODEL AND HYPOTHESES DEVELOPMENT

Institutional theory explains how social and cultural pressures impact organizations, leading them to adopt similar practices. It identifies three types of pressures: coercive, which arise from dependencies on other organizations; normative, which stem from professional standards; and mimetic, which occur in uncertain situations where firms imitate successful peers. These pressures result in a shared set of values and practices among organizations in the same field, ultimately influencing managerial decisions and creating a uniform organizational landscape.

Institutional pressures significantly influence the adoption of e-waste management systems, as evidenced in the literature. This discussion primarily focuses on external pressures that drive firms to adopt voluntary environmental strategies, exceeding the performance levels mandated by environmental legislation (Daddi et al., 2016). Developed an institutional theory to explain how firms in Iraq adopt environmentally responsible e-waste disposal practices. The research posited that coercive forces have predominantly influenced environmental management practices, leading firms across various industries to adopt comparable approaches. Consequently, public organisations faced growing pressure from customers and

 https://doi.org/10.1633/JISTaP.2025.13.3.4

shareholders, leading to proposed legislative changes to enhance their environmental performance. Scholars, journalists, academics, and governments have studied e-waste's environmental impact. Executives influence e-waste disposal rates, promoting pro-environmental behaviour as companies seek for sustainability. Implementing ecologically acceptable electronic waste disposal depends on practical, theoretical, and legal considerations (Chen et al., 2011).

To mitigate the negative consequences of IT, academics have established the concept of coordination to address these concerns throughout the e-waste lifecycle. Environmental disposal of e-waste has become a hot concern. Most explanations for ecologically responsible e-waste disposal use the institutional approach. This method stresses how companies gain legitimacy by following societal norms and industry practises (DiMaggio & Powell, 1983). During institutionalisation, organisations incorporate e-waste disposal values and practises, resulting in a convergence of practises and reactions (Oliver, 1991).

This study developed a research model that emphasises coercive pressure, normative influence, and imitation as institutional constraints affecting e-waste environmental disposal, following an examination of the institutional perspective. The research demonstrates how external pressures compel organisations to manage e-waste in an environmentally responsible manner. From an institutional perspective, the pressures identified as independent variables in the research model influence the rapid organisational attitude that impacts behavioural outcomes (Krell et al., 2016). A research model has been developed to elucidate the key variables influencing the adoption of environmental disposal practices for e-waste, grounded in a thorough literature review. The study assessed three institutional forces to demonstrate a comprehensive approach. The model demonstrates that institutional pressures serve as independent variables influencing managers' intentions regarding e-waste disposal. Attitudes served as mediators in independent-dependent relationships. Fig. 1 illustrates the research model and its associated constructs. The subsequent section examines the three pressures.

### 3.1. Coercive Pressures

Numerous government agencies impact the e-waste disposal practices of public organisations. Legislation enables authorities to establish and implement coercive restrictions. Numerous studies have indicated that enacted laws and regulations impact companies (Alziady, 2018;

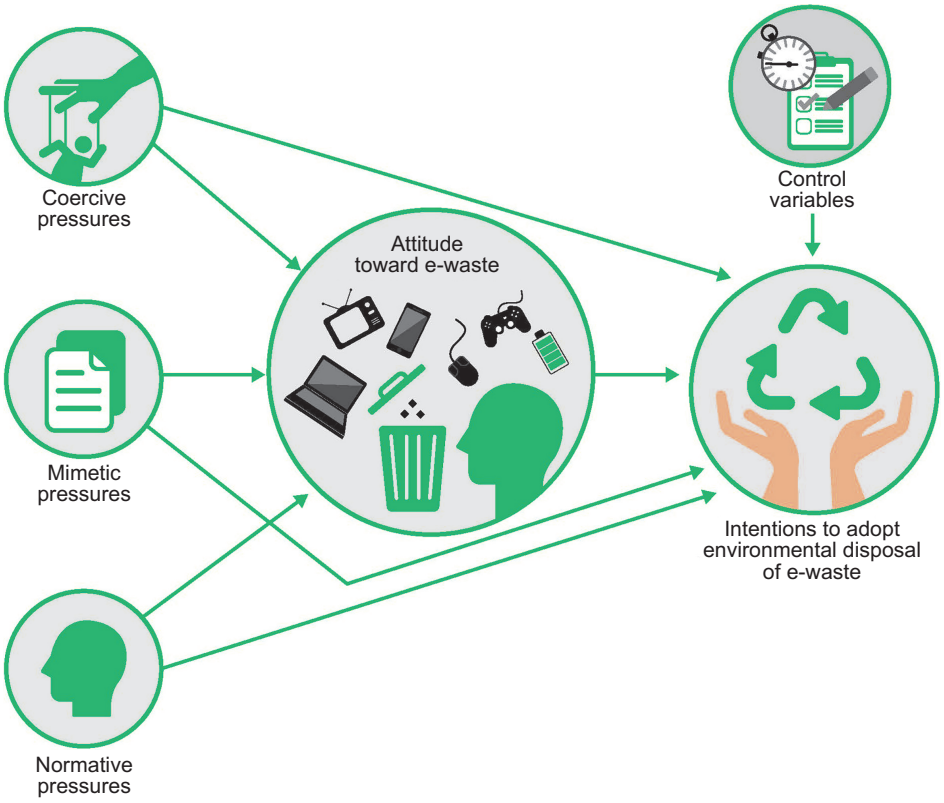



**Fig. 1.** Research model.

Alziady & Enayah, 2019). Public organisations impose penalties on e-waste disposal enterprises for noncompliance, either directly or indirectly. Institutional theory posits that coercive pressures originate from institutions establishing explicit norms that corporations must adhere to. The concept posits that institutions must possess sufficient strength to administer rewards or penalties for disobedience (DiMaggio & Powell, 1983). Coercive pressures influence environmental management practices, leading to similar procedures across industries, as noted by Jennings and Zandbergen (1995).

Institutions allocate resources to manage the disposal of e-waste. Furthermore, federal and municipal authorities require public organisations to manage e-waste disposal, thereby limiting the actions of companies. According to DiMaggio and Powell (1983), resource dependency is a driving factor behind coercive forces. Pfeffer and Salancik (1978) contended that influential actors controlling limited resources may require dependent organisations to implement structures that align with their interests. The survival of resource-dependent groups often necessitates adherence to specific protocols. Political influence and the requirements for legitimacy exert pressure. Recent federal and local government regulations mandate that public organisations adhere to environmental protection standards. Addressing requests is advantageous for businesses and mitigates the risk of penalties.

Resource-dominant organisations, government agencies, and other entities may compel corporations to implement effective e-waste disposal policies if they perceive that the activities of public organisations are detrimental to society. Research indicates that regulations are essential for fostering environmental practices (Daddi et al., 2016; Delmas & Toffel, 2004; Lin & Ho, 2016). Research indicates that coercive pressures are essential for adopting GIT (Daddi et al., 2016; Vejvar et al., 2018). The following hypotheses are presented:

H1: Coercive pressures are positively related to the manager's attitude towards the environmental disposal of e-waste.

H2: Coercive pressures are positively related to the manager's intentions to adopt the environmental disposal of e-waste.

### 3.2. Mimetic Pressures

The success of other adopters or the popularity of a practice incentivises public bodies to adopt it. Public organisations reflect behaviours similar to those of comparable companies. Delmas and Toffel (2004) asserted that networked organisations replicate the behaviours of other firms. Mimetic isomorphism refers to the tendency of companies to emulate industry leaders who benefit from being first movers. The description addresses the environmental disposal of e-waste. An organisation may engage in mimetic pressure by observing that other companies have successfully disposed of e-waste while generating profits. Some companies adopt a prevalent practice without due consideration, leading to its devaluation (Cappellaro, 2014). Mimetic pressure arises from uncertainty regarding the appropriate methods to address a problem, perform a task, or attain an objective (Krell et al., 2016). Organisations may mimic others in response to these pressures during periods of uncertainty. Organisations are believed to emulate successful structurally similar entities (Coffey et al., 2013). Institutions participate in the institutional trend to avoid appearing unique compared to other organisations (DiMaggio & Powell, 1983), while still pursuing social legitimacy. Krell et al. (2016) also suggested that organisations may seek solutions from other entities facing similar challenges in the absence of information to address difficulties. Chen et al. (2011) and Coffey et al. (2013) demonstrate significant mimetic pressures. We propose the following hypotheses:

H3: Mimetic pressures are positively related to the manager's attitude towards the environmental disposal of e-waste.

H4: Mimetic pressures are positively related to the manager's intentions to adopt the environmental disposal of e-waste.

### 3.3. Normative Pressures

Professionalisation represents an organisational response to normative pressures (Deng & Ji, 2015). Global decision-makers are addressing the adverse implications of IT and seeking to mitigate these effects. Normative pressures arising from supplier-customer interactions facilitate the identification of innovations, advantages, and expenses. The rise in professionalisation, driven by concerns regarding education and training, generates normative pressures (Vejvar et al., 2018). Companies will engage in environmentally responsible disposal of e-waste when such practices become standard due to normative pressures arising from ongoing interactions with suppliers, consumers, or trade organisations (Deng & Ji, 2015). Normative pressures within institutions originate from professional or industry associations (Krell et al., 2016).

 https://doi.org/10.1633/JISTaP.2025.13.3.4

In contrast to coercive pressures, normative signals have increased; however, organisations that implement them lack the ability to directly enforce compliance or impose penalties for disobedience (Chen et al., 2011). Decision-makers align with industrial and professional institutions, leading enterprises to adhere to regulations voluntarily. Consequently, decision-makers believe that adherence to professional and industry standards will benefit the organisation. Public organizations are likely to manage e-waste responsibly when decision-makers perceive it as enhancing consumer appeal and showcasing their commitment to environmental sustainability. According to Kuo and Dick (2010), normative forces significantly influence the adoption of GIT practices. The discussion proposes the subsequent hypotheses:

H5: Normative pressures are positively related to the manager's attitude towards the environmental disposal of e-waste.

H6: Normative pressures are positively related to the manager's intentions towards adopting environmental disposal of e-waste.

### 3.4. Attitude Towards Adopting Environmental Disposal of E-Waste

Numerous scholars and researchers assert that attitudes towards the environmental disposal of e-waste are crucial for acceptability and have incorporated these perspectives into corporate operations. The study of attitude, which influences individuals' perceptions, emotions, and behaviours regarding e-waste, has received significant attention. The proliferation of new technologies in the twenty-first century necessitates disposing of electronic waste responsibly. Simultaneously, the perspectives of public organisation managers regarding e-waste disposal should be taken into account during implementing practices aimed at environmental protection. Allport (1935) defined attitude as "a mental and neurological state of readiness that influences the individual's behaviour towards all related objects and situations." The individual's attitude towards the activity influences their preference for it (Ajzen et al., 2018). Schwartz (1992) characterised attitude as a system of ideas regarding an object or action that may result in its realisation. The attitude towards conduct reflects an individual's perception of the behaviour in question (Wang et al., 2016). This research analyses managerial perspectives regarding the detrimental effects of electronic waste and its environmental disposal practices.

The extent to which individuals perceive themselves as integral to the natural environment influences their environmental perspectives. Managers will view a behaviour positively if they perceive it yields beneficial outcomes. The conduct will yield adverse effects. The behaviour aligns with the disposition. An initial assessment of attitudes can forecast behavioural performance and elucidate the reasons behind specific behaviours (Ajzen, 1985). This study characterises attitude as the preferences and aversions of organisational managers, noting that environmental attitudes generally affect decisions regarding e-waste disposal. Consequently, the analysis indicates:

H7: A manager's attitude towards the environmental disposal of e-waste is positively related to the intentions towards adopting the environmental disposal of e-waste.

H8: A manager's attitude towards the environmental disposal of e-waste mediates the relationship between institutional pressures and intentions towards adopting the environmental disposal of e-waste.

### 3.5. Relationship Between Institutional Pressures and Intentions to Adopt Environmental Disposal of E-Waste

Improving e-waste collection, treatment, and recycling at the end of a device's life could enhance resource efficiency and support the circular economy in Iraq. Presumably, adopting environmental disposal of e-waste practices is a realistic approach for organisations to manage current environmental problems from accumulated e-waste. Significantly, companies can acquire a competitive advantage by distinguishing themselves from other organisations by adopting the practice while improving their organisational image and economic performance. Furthermore, adopting the environmental disposal of e-waste practice is an example of institutions' organisational behaviour request, as the motive to adopt is acceptance instead of maximising organisational efficiency.

Environmental disposal of e-waste is currently a developing field that lacks theoretical and empirical research in Iraq. Public organisations are under increasing institutional pressures to reduce the negative impact of IT on the environment; hence, adopting the environmental disposal of e-waste is vital. Moreover, more managers are experimenting with adopting the environmental disposal of e-waste practice; thus, studies should investigate the intentions towards adopting the practice as key factors. The intention towards adopting the environmental disposal of e-waste practices has additional motivational factors beyond standard IT adoption (Molla et al., 2014). There-

fore, the motivational factors include economic benefits, regulatory requirements, stakeholder obligations, and ethical reasons that must be considered when exploring and analysing factors that influence the adoption of environmental disposal of e-waste practices (Thomson & van Belle, 2015).

Although research on the environmental disposal of e-waste has recently increased within the IT discipline, there is still considerable overlap among the influencing factors, making them difficult to distinguish. Nevertheless, managers may choose to adopt environmentally responsible e-waste disposal practices either in response to external pressures or to gain strategic benefits—an understandable decision given the complexity and ambiguity surrounding the issue. Based on earlier studies, this study conceptualised the intention to adopt the environmental disposal of e-waste practices as the dependent variable. The economic performance of public organizations is understandable. Recently, managers of these organizations have started to realise that ignoring their environmental performance is absurd. Diffusion and penetration of e-waste practices in organisations reduces power consumption and carbon emissions, improves operating system performance, and increases interaction and collaboration (Deng & Ji, 2015). Hence, the emerging role of e-waste management and benefits highlight the importance of understanding that continuously practicing environmental disposal of e-waste practices could develop and enhance GIT practices.

## 4. METHODS

Public entities that are environmentally detrimental are developing green technologies. The focus here was on leaders within the public sector. E-waste is a recent development, leading managers to be more aware, skilled, and informed in selecting and evaluating environmental disposal methods. The target respondents were executives familiar with the company and responsible for e-waste disposal. Prior studies yielded quantitative assessments. Mat Nawi et al. (2024) and Krell et al. (2016) provided six items related to coercive pressure, five items about mimetic pressure, and four items associated with normative pressure. Wang et al. (2016) posed three enquiries regarding attitudes and intentions.

A cover letter proposed the study and solicited questionnaires from managers. Anonymous data analysis holds potential for advancing scientific research. The study conducted in Thi-Qar province involved 490 managers and yielded 440 responses. The analysis excluded 11 incomplete responses from a total of 313, representing 71% of the dataset. A total of 302 surveys were deemed analysable. Reminder letters and phone calls elevated the net study response to 67.6%, indicating a positive outcome. The accuracy of the survey questionnaire was verified through translations into Arabic and subsequent back-translation into English. Measurements were verified by three experts, and a pilot study assessed Cronbach's alpha. Cronbach's alpha was greater than 0.70, signifying reliability.

All research constructs employed a seven-point Likert scale ranging from very little (1) to very much (7). Appendix 1 presents the construct and measurement items of the most commonly utilised and effective survey-type research scaling instruments. The research incorporated data from public organisations in Thi-Qar, which are notable contributors to e-waste generation. More than 80% of the 302 respondents, comprising 60 women and 242 men, held college degrees. Employees in higher education, health services, construction, oil and gas, manufacturing, water and sewage, banking, agriculture, information and communication technology, insurance, and transportation sectors have responded. Descriptions of responders are presented in Table 1.

The partial least squares (PLS) is advantageous for investigation due to its robust predictive capabilities and minimal requirements regarding sample size and residual distribution (Chin, 1998). Hair et al. (2010) demonstrated that Smart PLS exhibits greater efficiency. The PLS also assessed the hypothesised correlations among the research variables. PLS is a second-generation multivariate technique that evaluates the psychometric properties of scales designed to measure a variable and estimates the parameters of structural models, including the strength and direction of correlations among model variables (Fornell & Larcker, 1981). The study employed a two-step approach to model testing, as outlined by Ananno et al. (2021). The study initially evaluated the measurement qualities of reflective latent concepts through confirmatory factor analysis (CFA). The study employed structural equation modelling to test the hypotheses.

## 5. RESULTS

The study operationalised all components using existing metrics based on literature studies. Mat Nawi et al. (2024) and Liang et al. (2007) developed six-item coercive pressures, five-item mimetic pressures, and four-item normative pressures. Wang et al. (2016) assessed attitudes

 https://doi.org/10.1633/JISTaP.2025.13.3.4

**Table 1.** Descriptive information of respondents

| Category | Label | Proportion (%) | No. of respondents |
|---|---|---|---|
| Gender | Woman | 19.90 | 60 |
| | Man | 80.10 | 242 |
| Education level | Bachelor's degree | 79.9 | 240 |
| | Master's degree | 20.1 | 59 |
| Sector representation | Higher education | 26.00 | 79 |
| | Education | 17.00 | 51 |
| | Health services | 13.00 | 39 |
| | Building & construction | 10.00 | 30 |
| | Oil & gas | 8.00 | 24 |
| | Manufacturing | 6.00 | 18 |
| | Water & sewage | 6.00 | 18 |
| | Banking | 5.00 | 15 |
| | Agriculture | 4.00 | 12 |
| | Information & communication | 3.00 | 9 |
| | Insurance activities | 1.00 | 3 |
| | Transportation | 1.00 | 3 |

toward e-waste disposal methods using a three-item construct, while also measuring the dependent variable and institutional pressures influencing the adoption of e-waste practices, with all scale items rephrased in English. For Arabic responses, back-translation confirmed the translation. The study identified the measurement model to ensure the survey instrument's validity and robustness per MacKenzie et al. (2011). Assessing the scale's psychometric characteristics, convergent, discriminant, and nomological validity required data. Experts proposed questions for each construct to assess content validity. Seven academics from Thi-Qar University and the University of Sumeru examined the survey's suitability, while IT managers analysed question understandability. Feedback validated the instrument's suitability and clarity.

### 5.1. Measurement

We assessed the measuring scale's reliability, discriminant validity, and convergent validity. Examining item loadings on latent variables examined reflecting item and construct validity and reliability (Hulland, 1999). Using principal component analysis and the maximum likelihood approach, CFA assessed scale reliability and validity. In regions with strong a priori theory, Bagozzi et al. (1991) preferred CFA over exploratory factor analysis. This study tested ideas with pre-validated scales. Higher loadings indicate larger construct-item variation than mistakes in PLS-based CFA (Hulland, 1999).

The study identified five factors explaining 77% of the total variance, with eight values greater than 1.0. Each construct was confirmed to be unidimensional and distinct, with items loading significantly on their respective factors. The analysis followed established guidelines for reliability and validity, utilizing composite reliability instead of Cronbach's alpha for better accuracy. All constructs showed high reliability, with loading values exceeding 0.70 and average variance extracted values above 0.50, ensuring strong discriminant validity. This rigorous assessment confirms the effectiveness of the measurement scales used in the study (Table 2).

The text details the validation of a measurement model, confirming its discriminant and convergent validity. It highlights significant factor loadings and t-statistics, indicating that the measures meet the necessary criteria. The analysis also addressed common method bias, using the Harman one-factor test, which revealed no dominant factor and suggested that common method bias was not an issue. Moreover, the intercorrelation matrix showed that factors were not highly correlated, supporting the findings. Overall, the results demonstrate that the constructs are distinct and adequately measured (Table 3).

**Table 2.** Factor analysis matrix

| Items | Mimetic | Coercive | Normative | Attitude | Intention | Communalities |
|---|---|---|---|---|---|---|
| MI1 | 0.79 | 0.15 | 0.13 | 0.14 | 0.15 | 0.78 |
| MI2 | 0.80 | 0.09 | 0.115 | 0.22 | 0.14 | 0.77 |
| MI3 | 0.78 | 0.11 | 0.14 | 0.13 | 0.16 | 0.75 |
| MI4 | 0.81 | 0.13 | 0.20 | 0.10 | 0.14 | 0.80 |
| MI5 | 0.86 | 0.1i | 0.10 | 0.18 | 0.17 | 0.70 |
| CI1 | 0.15 | 0.81 | 0.14 | 0.15 | 0.14 | 0.75 |
| CI2 | 0.16 | 0.84 | 0.16 | 0.16 | 0.08 | 0.74 |
| CI3 | 0.08 | 0.83 | 0.17 | 0.14 | 0.15 | 0.75 |
| CI4 | 0.12 | 0.83 | 0.16 | 0.10 | 0.13 | 0.71 |
| CI5 | 0.10 | 0.84 | 0.11 | 0.12 | 0.11 | 0.74 |
| C16 | 0.12 | 0.79 | 0.12 | 0.10 | 0.09 | 0.76 |
| NI1 | 0.09 | 0.11 | 0.81 | 0.11 | 0.15 | 0.76 |
| NI2 | 0.17 | 0.14 | 0.84 | 0.17 | 0.09 | 0.77 |
| NI3 | 0.16 | 0.08 | 0.81 | 0.16 | 0.13 | 0.80 |
| NI4 | 0.11 | 0.07 | 0.81 | 0.22 | 0.18 | 0.81 |
| AT1 | 0.08 | 0.09 | 0.11 | 0.83 | 0.16 | 0.79 |
| AT2 | 0.07 | 0.12 | 0.12 | 0.82 | 0.11 | 0.84 |
| AT3 | 0.09 | 0.10 | 0.16 | 0.90 | 0.09 | 0.78 |
| IN1 | 0.11 | 0.07 | 0.12 | 0.18 | 0.89 | 0.84 |
| IN2 | 0.13 | 0.14 | 0.11 | 0.11 | 0.91 | 0.81 |
| IN3 | 0.10 | 0.07 | 0.10 | 0.22 | 0.90 | 0.80 |
| Eigenvalue | 4.43 | 4.08 | 3.19 | 2.30 | 2.22 | 16.22 |
| Explained variance (%) | 0.21 | 0.19 | 0.15 | 0.11 | 0.11 | 0.77 |

CI, coercive pressure; MI, mimetic pressure; NI, normative pressure; AT, attitude; IN, intention.

**Table 3.** Assessment of reliability, convergent, and discriminant validity of research constructs

| Construct | Coercive | Mimetic | Normative | Attitude | Intention | CA | CR |
|---|---|---|---|---|---|---|---|
| Coercive | 0.80 | | | | | 0.77 | 0.78 |
| Mimetic | 0.37 | 0.82 | | | | 0.78 | 0.80 |
| Normative | 0.34 | 0.31 | 0.82 | | | 0.75 | 0.76 |
| Attitude | 0.36 | 0.29 | 0.35 | 0.85 | | 0.74 | 0.75 |
| Intention | 0.47 | 0.49 | 0.50 | 0.55 | 0.89 | 0.78 | 0.79 |

Diagonal elements represent the square roots of AVE, CA, and composite reliability.
CA, Cronbach's alpha; CR, composite reliability.

## 5.2. Structural Model

A path analysis was conducted using ordinary least squares hierarchical multiple regressions to test the proposed hypotheses. This method evaluates how explanatory variables influence outcomes by assessing individual paths. The analysis confirmed good internal consistency with reliability scores between 0.75 and 0.8. An intercorrelation matrix showed no multicollinearity issues, as cor-

 https://doi.org/10.1633/JISTaP.2025.13.3.4

relations remained below 0.50 and variance inflation factor values were under 3.3. Additionally, a PLS-structural equation modeling analysis was performed using SmartPLS 3.0 (Ringle et al., 2015) to visualize the structural model, which illustrated the explained variance of the endogenous variables and standardized path coefficients. A bootstrap analysis with 500 resamples was also conducted to determine the significance of the estimates.

Fig. 2 illustrates that institutional forces (coercive, normative, and mimetic) positively affect managers' attitudes towards e-waste environmental disposal. The results also demonstrated that coercive, normative, and imitation pressures affected e-waste environmental disposal intentions. The path coefficient from attitude to intention to dispose of e-waste environmentally was substantial, supporting all predictions (H1) suggested that coercive forces strengthen e-waste environmental disposal attitudes. The predicted hypothesis (H1) was supported (β=0.27, T=2.31, $p$<0.05). The path from mimetic pressures to e-waste disposal attitude (β=0.28, T=2.16, $p$<0.05) supported (H3). The study found that normative pressures favourably impact attitudes towards e-waste disposal (H5) (β=0.31, T=2.41, $p$<0.05) and (H2) (β=0.32, T=2.39, $p$<0.05) (H7) said attitude favourably affects e-waste environmental disposal intentions. The hypothesis was well-supported (β=0.47, T=3.11, $p$<0.05).

As mentioned in (H4), mimetic pressures increase e-waste environmental disposal intentions. Results supported the hypothesis (β=0.30, T=2.11, $p$<0.05) (H6) concluded that normative pressure favourably influences e-waste environmental disposal intentions. Results supported the hypothesis (β=0.26, T=2.42, $p$<0.05). The numbers in Table 4 confirm the path model in Fig. 2. Table 2 presents the factor loadings and communalities for all measurement items, confirming the validity and reliability of the constructs. Table 3 summarizes the results of reliability analysis, convergent validity, and discriminant validity for all research constructs. Attitude towards environmental disposal of e-waste mediates or supplements institutional pressures and the intention to do so, since the indirect and direct effects exhibited comparable signs, confirming H8. Comparing the calculated correlations (i.e., total effects) with the real correlations between the independent variables and the dependent variables confirmed the model's "goodness of fit." The study required that the absolute difference between replicated (total effects) and original correlations not exceed 0.10. Fig. 2 shows structural model route coefficients and explained variances. Fig. 2 shows

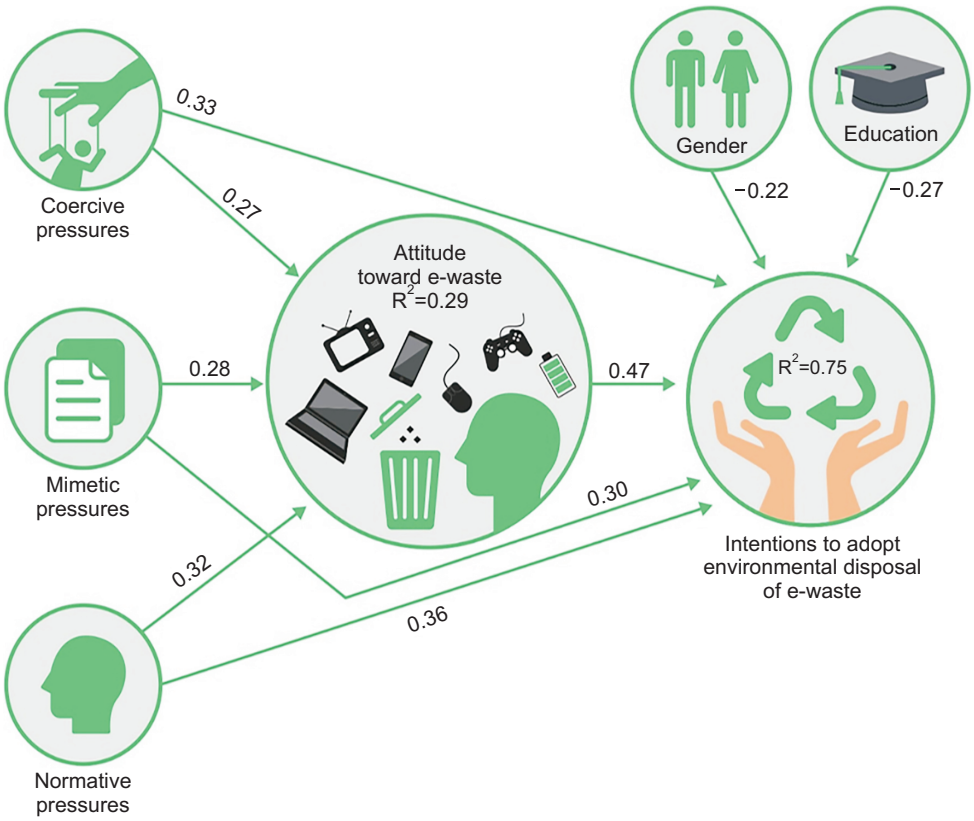



**Fig. 2.** Research model after test.

**Table 4.** Path model statistics

| Path coefficients | Direct effect | T | Signature $p$<0.05 | Indirect effect | T | Signature $p$<0.05 | Mediation |
|---|---|---|---|---|---|---|---|
| CI----IN | 0.33 | 2.39 | Yes | 0.12 | 2.03 | Yes | Partial |
| MI----IN | 0.30 | 2.11 | Yes | 0.14 | 2.11 | Yes | Partial |
| NI----IN | 0.36 | 2.43 | Yes | 0.15 | 2.99 | Yes | Partial |
| Gender-----IN | -0.22 | 3.02 | Yes | | | | |
| Size--------IN | 0.08 | 1.23 | No | | | | |
| Education---IN | 0.27 | 3.02 | Yes | | | | |
| Sector-----IN | 0.06 | 1.34 | N0 | | | | |
| Age-------IN | 0.09 | 0.98 | No | | | | |

CI, coercive pressure; MI, mimetic pressure; NI, normative pressure; AT, attitude; IN, intention.

the structural model PLS findings, including standardised path coefficients, significance, and variance explained (R2). Control factors, such as sector, age, and firm, did not significantly affect intention ($p$-value>0.05), while gender and education did ($p$-value<0.05).

## 6. DISCUSSION

Public corporations are pressured by rivals, consumers, regulators, and community organisations to adopt more environmentally friendly practices that benefit internal and external stakeholders. E-waste disposal must be considered in organisational strategy. Most Iraqi organisations now dispose of e-waste environmentally. Firms must embrace the practice to fit in and function well. Thus, this study seeks to empirically evaluate institutional pressures on e-waste environmental disposal willingness and attitude as a mediator. The study asks: (1) Do institutional forces (coercive, mimetic, and normative) impact public organisation managers' e-waste practices? (2) How does attitude mediate the relationships among variables? The current work developed and tested a theoretical model to address the problems.

The following section details this study's primary discoveries, theoretical contributions, and practical consequences. The results showed that the three forces favourably affected respondents' e-waste environmental disposal. Several interviewees said their companies had had recent successes and a local reputation for e-waste environmental disposal. However, government, supplier, and consumer pressures pressure corporations to prioritise the issue. Coercive pressures have a positive impact on e-waste disposal. Based on the high study sample agreement on pertinent topics, rules are a greater source of coercive pressures than suppliers and customers. Governments were also important in businesses' choice to dispose of e-waste environmentally. Governments may show their support by boosting the image of e-waste disposal companies. Governments can also reduce information and search costs by providing technical assistance to potential adopters, demonstrating their effectiveness in promoting green behaviours across organisations, especially when they have an organisation-wide impact (Chen et al., 2011). The results verified Molla et al. (2014), who said that government policies that promote GIT practices across the IT life cycle were a significant factor. Chen et al. (2011) found that rules affected pollution control and sustainable development more than mimetic pressure. According to the study, governments in developing countries need to establish clear environmental guidelines for the adoption and disposal of e-waste, as electronic and electrical waste are among the fastest-growing waste categories globally. Therefore, governments worldwide have created legislation to handle device post-consumption (Favot & Grassetti, 2017). In this study, the results are due to the federal government and local administrators' efforts to impose regulations to control public organisations' environmental activities and diffuse pro-environmental practices, especially since most public organisation managers try to balance economic and environmental goals, which should be balanced but are often forgotten by companies. Thus, external forces shape e-waste disposal attitudes.

Although Chen et al. (2011) excluded normative pressure, this study found that normative pressures influence attitudes towards e-waste environmental disposal. This suggests that company decision-makers have realised the benefits of adopting the practice. Equally, corporations aim to build links with other firms, and professional

groups have used GIT to profit. Due to the ubiquity and maturity of e-waste procedures, the two studies found different results. Most responders want to recycle electronics, justifying their e-waste disposal habits. Mimetic stresses also improve e-waste environmental disposal. Krell et al. (2016) found that copying similar institutions enhances a firm's learning experience. Likewise, it was revealed that decision-makers often prefer adopting other companies' IT strategies instead of depending solely on internal decisions. A decision-maker could ignore internal technology evaluation conclusions to mimic another company's selections. In addition, rivals used e-waste environmental disposal to gain significant benefits and satisfy suppliers and consumers. The results match those of Molla et al. (2014), who revealed that select enterprises used GIT investment as a competitive advantage that others copied. The number of other businesses adopting GIT practices was unlikely to affect enterprises. However, if they saw favourable results in other organisations, they were inclined to embrace them (Wang et al., 2016). Since all this is still new, most public organisations are refining their operating methods. However, numerous companies have realised that environmental disposal of e-waste allows a significant change in IT capabilities to increase operational performance and address environmental challenges. Imitating successful businesses is widespread.

Additionally, corporations understood that e-waste management is vital to market share. According to an extensive study, businesses' effective and efficient e-waste disposal depends on external market dynamics. Thus, institutional forces influence e-waste environmental disposal. This solved the study's initial equations. Research suggests that e-waste environmental disposal should boost performance and market competitiveness rather than merely complying with environmental protection standards. Thus, environmental e-waste disposal becomes a strategic alternative subject to institutional pressures (Sharma et al., 2024; Wang et al., 2016). The study found three institutional elements affecting GIT: coercive pressure, normative influence, and imitation. The variables also impact e-waste environmental disposal intentions. In contrast, coercive constraints greatly influenced attitude, with imitation and normative pressures impacting managers' intention to dispose of e-waste environmentally. Since other companies in the sector view e-waste disposal as a legitimate business strategy to win market share, tolerating the conduct has become a daily habit. The company will be more open to IT innovations that cut costs or boost income, regardless of greenness.

The study found that consumers, suppliers, government, and professional bodies no longer significantly impact small businesses' e-waste practices. When early adopters achieved positive results from green practices, market pressures became a mimetic incentive (Molla et al., 2014). Attitude towards e-waste environmental disposal also affects adoption. Public organisation changes drastically alter industry structure and modify small businesses' competitive tactics to gain new advantages. Organisations must practise GIT to combat environmental deterioration. In addition, public organisations that embrace environmental e-waste disposal practices develop an environmental culture and pro-environment attitude after recognizing the issue. Finally, organisation managers' awareness leads to a belief in its effectiveness and ease of implementation. Thus, manager knowledge enhances the desire to continue e-waste environmental disposal (Mat Nawi et al., 2024). Thus, the findings shed light on institutional influences on e-waste attitude and environmental disposal. The practical contribution is that first, the data can help managers adapt the adoption process, spread, and build trust in e-waste environmental disposal practices in enterprises. Managers adopted and maintained the practices to acquire acceptance and economic rewards. Additional study with a broader sample of managers is needed to confirm the applicability of the findings to other public companies.

This study's results on attitude's mediation function between institutional pressures and behavioural intentions advance theory and practice. The findings show that management attitudes are a key psychological mechanism for internalising external institutional influences and promoting environmentally responsible e-waste disposal. The mediation effect has serious ramifications. The findings improve theoretical knowledge of e-waste adoption and continuation intentions. The model explained 75% of e-waste environmental disposal intentions and 29% of attitudes. Interestingly, a considerable fraction of the variance remains unexplained, highlighting the need for more research incorporating unmeasured factors.

## 7. CONCLUSION

This study investigated the influence of institutional pressures on public sector managers' intentions to adopt environmentally responsible e-waste disposal practices in Iraq's Thi-Qar province, with particular attention to the mediating role of managerial attitudes. Through a rigorous empirical examination of 302 public organization managers, our research yields several important theoreti-

cal and practical contributions while also identifying valuable avenues for future scholarship. The analysis revealed three key findings that advance our understanding of institutional drivers in environmental management. First, coercive pressures emerged as the most potent predictor of both managerial attitudes ($\beta$=0.27, $p$<0.05) and adoption intentions ($\beta$=0.32, $p$<0.05), underscoring the critical role of regulatory frameworks in developing nation contexts. Second, the study empirically validated attitude as a significant mediator (H8 supported), bridging the gap between external institutional pressures and internal behavioral intentions. Third, while mimetic and normative pressures demonstrated measurable effects, their comparatively weaker influence suggests that isomorphic pressures operate differently in Iraq's unique institutional environment than in developed economies.

Our work makes three substantive contributions to the literature. Theoretically, we extend institutional theory by demonstrating its applicability to e-waste management in conflict-affected regions, while simultaneously enriching the Theory of Planned Behavior by contextualizing attitude formation within institutional frameworks. Methodologically, we advance measurement approaches for developing contexts through our rigorous translation and validation process, which included expert panels, pilot testing, and CFA. Practically, our findings offer evidence-based guidance for policymakers seeking to improve e-waste management in similar institutional environments.

Several limitations warrant consideration. The cross-sectional design precludes causal inferences, suggesting the need for longitudinal studies to examine how institutional pressures and attitudes evolve. Our focus on public sector organizations in one Iraqi province limits generalizability, inviting future research across sectors and geographical contexts. Additionally, while we examined three institutional pressures, other factors such as organizational culture or resource constraints may moderate these relationships. These limitations notwithstanding, our findings have important implications. For theory, they demonstrate how institutional forces shape pro-environmental behavior in developing economies, challenging assumptions derived primarily from Western contexts. For practice, they suggest that policy interventions should prioritize regulatory enforcement while simultaneously cultivating positive attitudes through training and awareness programs. Future research should explore how digital technologies might amplify institutional effects and how cost-benefit considerations influence adoption decisions. By illuminating the institutional determinants of e-waste management in Iraq, this study provides a foundation for both scholarly advancement and practical improvement in environmental governance. As electronic waste continues to grow as a global challenge, understanding these institutional-behavioral linkages becomes increasingly vital for sustainable development.

## CONFLICTS OF INTEREST

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

## REFERENCES

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *Action control: From cognition to behavior* (pp. 11-39). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-69746-3_2

Ajzen, I., Fishbein, M., Lohmann, S., & Albarracín, D. (2018). The influence of attitudes on behavior. In D. Albarracin, & B. T. Johnson (Eds.), *The handbook of attitudes, volume 1: Basic principles* (pp. 59-367). Routledge. https://doi.org/10.4324/9781315178103

Akman, I., & Mishra, A. (2015). Sector diversity in green information technology practices: Technology acceptance model perspective. *Computers in Human Behavior*, 49, 477-486. https://doi.org/10.1016/j.chb.2015.03.009

Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *Handbook of social psychology* (pp. 798-844). Clark University Press. https://www.scribd.com/document/333120405/Allport-G-W-1935-Attitudes-in-Handbook-of-Social-Psychology-C-Murchison-798-844

Alziady, A. A. D. J. (2018). Studying the effect of institutional pressures on the attitudes of the managers of small enterprises in Thi-Qar province toward intention continuance environmental disposal of e-waste. *Journal of Economics and Administrative Sciences*, 24(107), 24. https://doi.org/10.33095/jeas.v24i107.1287

Alziady, A. A. D. J., & Enayah, S. H. (2019). Studying the effect of institutional pressures on the intentions to continue green information technology usage. *Asian Journal of Sustainability and Social Responsibility*, 4(1), 4. https://doi.org/10.1186/s41180-018-0023-1

Ananno, A. A., Masud, M. H., Dabnichki, P., Mahjabeen, M., & Chowdhury, S. A. (2021). Survey and analysis of consumers' behaviour for electronic waste management in Bangladesh. *Journal of Environmental Management*, 282, 111943. https://doi.org/10.1016/j.jenvman.2021.111943

Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing construct validity in organizational research. *Administrative Science Quarterly*, 36(3), 421-458. https://doi.org/10.2307/2393203

Borthakur, A., & Govind, M. (2018). Public understandings of e-waste and its disposal in urban India: From a review towards a conceptual framework. *Journal of Cleaner Production*, 172, 1053-1066. https://doi.org/10.1016/j.jclepro.2017.10.218

Cappellaro, G. (2014). Book review: Institutional theory and organizational change. *Organization Studies*, 35(12), 1893-1896. https://doi.org/10.1177/0170840614534386

Chen, A. J., Watson, R. T., Boudreau, M. C., & Karahanna, E. (2011). An institutional perspective on the adoption of green IS & IT. *Australasian Journal of Information Systems*, 17(1), 5-27. https://doi.org/10.3127/ajis.v17i1.572

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. In G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 295-336). Psychology Press. https://www.taylorfrancis.com/chapters/edit/10.4324/9781410604385-10/partial-least-squares-approach-structural-equation-modeling-wynne-chin

Coffey, P., Tate, M., & Toland, J. (2013). Small business in a small country: Attitudes to "Green" IT. *Information Systems Frontiers*, 15(5), 761-778. https://doi.org/10.1007/s10796-013-9410-4

Cucchiella, F., D'Adamo, I., Lenny Koh, S. C., & Rosa, P. (2015). Recycling of WEEEs: An economic assessment of present and future e-waste streams. *Renewable and Sustainable Energy Reviews*, 51, 263-272. https://doi.org/10.1016/j.rser.2015.06.010

Daddi, T., Testa, F., Frey, M., & Iraldo, F. (2016). Exploring the link between institutional pressures and environmental management systems effectiveness: An empirical study. *Journal of Environmental Management*, 183, 647-656. https://doi.org/10.1016/j.jenvman.2016.09.025

Delmas, M., & Toffel, M. W. (2004). Institutional pressure and environmental management practices. In S. Sharma & M. Starik (Eds.), *Stakeholders, the environment and society* (pp. 230-245). Edward Elgar Publishing. https://www.hbs.edu/faculty/Pages/item.aspx?num=22075

Deng, Q., & Ji, S. (2015). Organizational green IT adoption: Concept and evidence. *Sustainability*, 7(12), 16737-16755. https://www.mdpi.com/2071-1050/7/12/15843

Dias, P., Machado, A., Huda, N., & Bernardes, A. M. (2018). Waste electric and electronic equipment (WEEE) management: A study on the Brazilian recycling routes. *Journal of Cleaner Production*, 174, 7-16. https://doi.org/10.1016/j.jclepro.2017.10.219

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147-160. https://doi.org/10.2307/2095101

Echegaray, F., & Hansstein, F. V. (2017). Assessing the intention-behavior gap in electronic waste recycling: The case of Brazil. *Journal of Cleaner Production*, 142, 180-190. https://doi.org/10.1016/j.jclepro.2016.05.064

Favot, M., & Grassetti, L. (2017). E-waste collection in Italy: Results from an exploratory analysis. *Waste Management*, 67, 222-231. https://doi.org/10.1016/j.wasman.2017.05.026

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. https://doi.org/10.2307/3151312

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis : A global perspective* (7th edition ed.). Pearson.

Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strategic Management Journal*, 20(2), 195-204. http://www.jstor.org/stable/3094025

Ikhlayel, M. (2018). An integrated approach to establish e-waste management systems for developing countries. *Journal of Cleaner Production*, 170, 119-130. https://doi.org/10.1016/j.jclepro.2017.09.137

Jennings, P. D., & Zandbergen, P. A. (1995). Ecologically sustainable organizations: An institutional approach. *The Academy of Management Review*, 20(4), 1015-1052. https://doi.org/10.2307/258964

Krell, K., Matook, S., & Rohde, F. (2016). The impact of legitimacy-based motives on IS adoption success: An institutional theory perspective. *Information & Management*, 53(6), 683-697. https://doi.org/10.1016/j.im.2016.02.006

Kumar, A., Holuszko, M., & Espinosa, D. C. R. (2017). E-waste: An overview on generation, collection, legislation and recycling practices. *Resources, Conservation and Recycling*, 122, 32-42. https://doi.org/10.1016/j.resconrec.2017.01.018

Kuo, B. N., & Dick, G. N. (2010). *Organizational green IT: It seems the bottom line rules*. Paper presented at AMCIS 2010 Proceedings: Sixteenth Americas Conference on Information Systems, Lima, Peru. http://aisel.aisnet.org/amcis2010/99

Liang, H., Saraf, N., Hu, Q., & Xue, Y. (2007). Assimilation of enterprise systems: The effect of institutional pressures and the mediating role of top management. *MIS Quarterly*, 31(1), 59-87. https://doi.org/10.2307/25148781

Lin, L. H., & Ho, Y. L. (2016). Institutional pressures and environmental performance in the global automotive industry:

The mediating role of organizational ambidexterity. *Long Range Planning*, 49(6), 764-775. https://doi.org/10.1016/j.lrp.2015.12.010

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly*, 35(2), 293-334. https://doi.org/10.2307/23044045

Mat Nawi, M. N., Fauzi, M. A., Ting, I. W. K., Wider, W., & Amaka, G. B. (2024). Green information technology and green information systems: science mapping of present and future trends. *Kybernetes*, 54(6), 3136-3155. https://doi.org/10.1108/k-10-2023-2139

Molla, A., Abareshi, A., & Cooper, V. (2014). Green IT beliefs and pro-environmental IT practices among IT professionals. *Information Technology & People*, 27(2), 129-154. https://doi.org/10.1108/itp-10-2012-0109

Nishant, R., Teo, T. S. H., & Goh, M. (2017). Do shareholders value green information technology announcements? *Journal of the Association for Information Systems*, 18(8). https://doi.org/10.17705/1jais.00466

Oliver, C. (1991). Strategic responses to institutional processes. *The Academy of Management Review*, 16(1), 145-179. https://doi.org/10.2307/258610

Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations : A resource dependence perspective*. Harper & Row.

Regel-Rosocka, M. (2018). Electronic wastes. *Physical Sciences Reviews*, 3(5), 20180020. https://doi.org/10.1515/psr-2018-0020

Ringle, C. M., Wende, S., & Becker, J. M. (2015). *SmartPLS 3*. https://www.smartpls.com

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1-65). Academic Press. https://doi.org/10.1016/S0065-2601(08)60281-6

Sharma, M., Kaushal, D., Joshi, S., Kumar, A., & Luthra, S. (2024). Electronic waste disposal behavioral intention of millennials: A moderating role of electronic word of mouth (eWOM) and perceived usage of online collection portal. *Journal of Cleaner Production*, 447, 141121. https://doi.org/10.1016/j.jclepro.2024.141121

Thomson, S., & van Belle J. P. (2015). Antecedents of green IT adoption in South African higher education institutions. *Electronic Journal Information Systems Evaluation*, 18(2), 172-186. https://academic-publishing.org/index.php/ejise/article/view/180

Vejvar, M., Lai, K. H., Lo, C. K. Y., & Fürst, E. W. M. (2018). Strategic responses to institutional forces pressuring sustainability practice adoption: Case-based evidence from inland port operations. *Transportation Research Part D: Transport and Environment*, 61, 274-288. https://doi.org/10.1016/j.trd.2017.08.014

Waghmode, M. S., Gunjal, A. B., & Patil, N. N. (2021). Bioleaching of electronic waste. *Pollution*, 7(1), 141-152. https://doi.org/10.22059/poll.2020.308119.872

Wang, Z., Guo, D., & Wang, X. (2016). Determinants of residents' e-waste recycling behaviour intentions: Evidence from China. *Journal of Cleaner Production*, 137, 850-860. https://doi.org/10.1016/j.jclepro.2016.07.155

Zhang, H., Liu, L., & Li, T. (2011). Designing IT systems according to environmental settings: A strategic analysis framework. *The Journal of Strategic Information Systems*, 20(1), 80-95. https://doi.org/10.1016/j.jsis.2011.01.001

Zhang, X., & Xu, Z. (2016). A review of current progress of recycling technologies for metals from waste electrical and electronic equipment. *Journal of Cleaner Production*, 127, 19-36. https://doi.org/10.1016/j.jclepro.2016.04.004

**Appendix 1.** Please mark the number to indicate the extent to which you agree with the following statements

| | Item | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|
| 1 | The government requires from my enterprise to disposal e- waste from the environment | | | | | | | |
| 2 | Using environmental disposal e-waste is necessary for legal compliance | | | | | | | |
| 3 | Regulatory requirements to enforce penalties for enterprise which not doing environmental disposal of e-waste | | | | | | | |
| 4 | My enterprise well-being depends on suppliers that have adopted the environmental disposal of e-waste | | | | | | | |
| 5 | My enterprise cannot easily switch away from suppliers that have adopted the environmental disposal of e-waste | | | | | | | |
| 6 | My enterprise well-being depends on their customers that supported environmental disposal of e-waste | | | | | | | |
| 7 | Our main competitors who have adopted environmental disposal of e-waste greatly benefitted | | | | | | | |
| 8 | Our main competitors who have adopted environmental disposal of e-waste are favorably perceived by others in the same industry | | | | | | | |
| 9 | Our main competitors who adopted environmental disposal of e-waste which are favorably perceived by their customers | | | | | | | |
| 10 | My enterprise's competitors that have adopted the environmental disposal of e-waste are favorably perceived by their suppliers | | | | | | | |
| 11 | An environmental disposal of e-waste have been used by our competitors | | | | | | | |
| 12 | The proportion of my enterprise's customers have used environmental disposal of e-waste | | | | | | | |
| 13 | The proportion of my firm's suppliers used environmental disposal of e-waste | | | | | | | |
| 14 | The extent of my enterprise's decision to use environmental disposal of e-waste was affected by promotions by the government | | | | | | | |
| 15 | The extent of my company's decision the using the environmental disposal of e-waste was affected by promotions by industry, trade, or professional bodies | | | | | | | |
| 16 | Environmental disposal of e-waste makes me feel very satisfied | | | | | | | |
| 17 | Environmental disposal of e-waste makes contribution to the society | | | | | | | |
| 18 | Environmental disposal of e-waste is everyone's responsibility | | | | | | | |
| 19 | I plan to keep using environmental disposal of e-waste practice in my job in the future | | | | | | | |
| 20 | I intend to continue using environmental disposal of e-waste practice in my job in the future | | | | | | | |
| 21 | I expect my use of environmental disposal of e-waste practice to continue in my job in the future | | | | | | | |

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

RESEARCH PAPER
J Inf Sci Theory Pract 13(3): 72-84, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.3.5

# Emoji Votes: Predicting Kathmandu's 2022 Mayoral Election via Facebook Sentiment

**Kapil Gurung***
College of Global AI Convergence, Seoul Christian University, Seoul, Korea
E-mail: kapilgurung@scu.ac.kr

**Shreejana Gnawali**
Global College, Kyungsung University, Busan, Korea
E-mail: ganawaliks@ks.ac.kr

## ABSTRACT

As political engagement is increasingly migrating to social media platforms, understanding the emotional dynamics embedded in online reactions has become critical to interpreting voter behavior. This study explores how Facebook's reactions function as affective signals of voter sentiment and predictive indicators of electoral outcomes, using the 2022 Kathmandu metropolitan mayoral election. Drawing on emotion theory and affective information behavior, the research analyzes 322 Facebook posts related to the campaign from three leading candidates: Balen Shah, Keshav Sthapit, and Sirjana Singh, focusing on emoji-based reactions, shares, and comments. Sentiment scores were computed using a weighted classification of emojis, and predictive potential was assessed through rank-order analysis and correlation with actual vote shares. The result reveals a positive correlation between pre-election sentiment indicators and electoral outcomes. Balen Shah, who led in both emotional engagement and final vote count, demonstrated how affective online support can signal electoral strength. Time-truncation and unweighted scoring further validated the robustness check. Additionally, text analytics highlight Shah's resonance with majority voters and his alignment with calls for systemic change, contrasting with mixed sentiment toward the latter two. This study contributes to the growing literature on digital political engagement by demonstrating that emoji reactions can serve as reliable proxies for public sentiment in emerging democracies. The findings suggest practical implications for political strategists, campaign managers, and communication agencies seeking to understand and respond to digital sentiment in real time.

**Keywords:** social media platform, sentiment analysis, emoji reaction, election campaign, voter behavior, affective information behavior

**Received:** March 30, 2025 **Revised:** June 11, 2025
**Accepted:** June 27, 2025 **Published:** September 30, 2025

***Corresponding Author:** Kapil Gurung
https://orcid.org/0000-0002-1534-977X
**E-mail:** kapilgurung@scu.ac.kr

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 (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.



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

## 1. INTRODUCTION

The rise of social media has significantly reshaped the landscape of political communication, enabling new forms of engagement between political actors and the public. Platforms such as Facebook, X (formerly Twitter), and TikTok have become instrumental in political discourse. They not only disseminate information but also shape political campaigns (Nahon & Hemsley, 2013; Tumasjan et al., 2010), civic participation (Boulianne, 2020), political protest (Theocharis et al., 2015), and voter behavior through affective digital cues shaping voter attitudes. Rather than functioning merely as content distribution systems, these platforms create dynamic affective environments where users signal support, dissent, or ambivalence in real-time.

Among social media platforms, Facebook stands out due to its widespread adoption and rich set of emotional affordances. It's seven emoji reactions: like, love, care, ha-ha, wow, sad, and angry (Fig. 1), offer researchers a granular and interpretable set of affective signals (Tian et al., 2017). These reactions enable users to communicate the intensity and character of their emotional stance toward political messages, providing a valuable lens through which to observe affective information behavior (AIB). The latest studies integrating emoji reactions with natural language processing and computer vision techniques have shown that these multimodal expressions can serve as reliable proxies for voter sentiment (Khan et al., 2025; Wisniewski et al., 2020). Compared to traditional surveys, social media data offers real-time, large-scale insights into public sentiment, with the potential to decode how individuals emotionally process and respond to political content (Muraoka et al., 2021; Saifeddine & Abdellatif Chakor, 2024).

Despite growing interest in affective digital engagement, much of the existing literature remains concentrated in Western contexts (Chambers et al., 2015) or larger democracies, such as India, leaving regions like Nepal relatively underexplored. Nepal's increasing digital adoption offers a compelling context to explore how AIB unfolds in emerging democracies. Following its transition from a monarchy to federal republic, Nepal has witnessed rapid growth in political discourse on digital platforms, particularly Facebook (Bhattarai, 2023). Recent studies have documented the use of social media for youth engagement (Wagle, 2024) and political mobilization (Kharel, 2024), but few have analyzed how Facebook's emotional reactions serve as behavioral indicators of political alignment.

This study seeks to address that gap by reframing Facebook emoji reactions as affective signals within the broader framework of AIB. Using the 2022 Kathmandu mayoral election as a case, we examine how the digital expression of emotions correlates with real-world electoral outcomes. This research seeks to: (1) examine affective patterns expressed through Facebook emoji reactions expressed toward political candidates; (2) evaluate the predictive power of sentiment polarity and user engagement to actual vote shares; (3) explore how nonpartisan emotional engagement challenges traditional party-based political alignment; and (4) demonstrate how emoji-based expressions predict electoral outcomes in Nepal's evolving political landscape.

## 2. LITERATURE REVIEW

### 2.1. Social Media Platforms and Digital Political Engagement

Social media platforms have reshaped political engagement by directly enabling emotionally charged interactions between candidates and voters. Platforms such as Facebook, X, and TikTok are no longer passive information outlets. They function as an active space where political sentiment is expressed, shaped, and amplified in real time (Boulianne, 2020; Persily & Tucker, 2020).

Emotional reactions have emerged as a key metric of user sentiment on social media platforms. Tian et al. (2017) showed that Facebook emoji reactions offer interpretable affective signals, which are often more reliable than text-based sentiment. Bond et al. (2012) found that emotionally resonant content on Facebook could influence real-world voter turnout, underscoring the behavioral power of digital emotion. Several studies link social media sentiment to electoral outcomes. Tumasjan et al. (2010) and

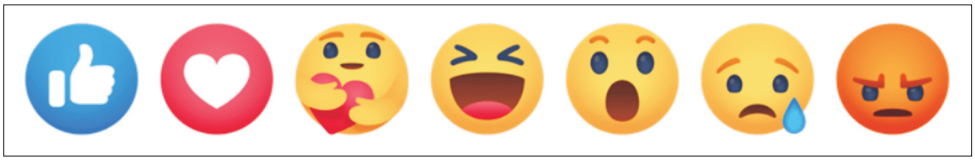



**Fig. 1.** Facebook's seven emojis.

Haryanto et al. (2019) demonstrated how X and Facebook sentiment mirrored election results in Germany and Indonesia, respectively. Muraoka et al. (2021) found that love and anger reactions to party posts correlated with public support across 79 democracies. These findings suggest that emoji reactions are more than mere expressions, as they serve as behavioral signals with predictive potential.

In the context of South Asia, Agarwal and Bansal (2020) examined Facebook posts during India's 2017 state election and found that sentiment polarity correlated with candidate performance. In Nepal, Bhattarai (2023) noted Facebook's growing role in political socialization, particularly among youth, while Upreti et al. (2025) found that emotional engagement on Facebook correlated with electoral success in the 2022 parliamentary election. Their study revealed patterns in voter sentiment, messaging strategies, and emotional tone. However, challenges including misinformation and echo chambers can skew sentiment (Guess et al., 2020), and in Nepal, limited digital access may affect representativeness (Kharel, 2024).

Despite this emerging body of work, comprehensive analyses of Facebook-based affective signaling in Nepal remain limited. In particular, the role of emoji reactions and emotional cues as behavioral indicators of political alignment remains underexplored, highlighting the need for further research into AIB in the country's digital political sphere.

### 2.2. Emotion Theory and Affective Information Behavior

Understanding how emotions shape information processing is crucial for interpreting digital political behavior. For classifying emotional responses, emotion theory provides the psychological foundation, while AIB explains how these emotional states shape the way individuals interact with information in real-world and digital environments.

Ekman (1993)'s theory of six universal emotions (happiness, sadness, fear, anger, surprise, and disgust) offers a widely accepted framework for decoding affective expression. Similarly, Plutchik (1965)'s wheel of emotions illustrates how primary emotions blend into complex affective states that influence perception, communication, and decision-making. In digital spaces like Facebook, emoji reactions such as love, care, anger, or sadness serve as proxies for these emotional expressions, allowing users to communicate affective responses to political content on a scale.

To further explore this emotional analysis with information science, this study draws from the field of AIB, which explores how emotional states influence information seeking, processing, and use. Kuhlthau (1991)'s information search process (ISP) model was among the first to integrate affect, identifying phases of uncertainty, optimism, confusion, and relief during search behavior. Nahl (2007) later expanded this with a tripartite model incorporating affective, cognitive, and sensorimotor dimensions. Savolainen (2015) emphasized the everyday, contextual nature of affective engagement, where emotions guide not only what information individuals attend to, but also how they evaluate and act upon it.

In this study, Facebook emoji reactions are conceptualized not just as affective signals but as manifestations of collective emotional information processing. Users reacting to political posts on social media are engaging in a form of public affective behavior—an expression of sentiment that reflects not only individual preferences but also broader political association. By aggregating and analyzing millions of such reactions, we argue that these aggregated affective signals function as behavioral indicators, offering predictive insight into electoral outcomes. This integration of emotion theory with AIB thus enables a more holistic understanding of how affect drives both digital engagement and real-world political decision-making.

## 3. RESEARCH QUESTIONS AND HYPOTHESES

This study explores the extent to which Facebook reactions serve as a real-time indicator of voter sentiment and electoral outcomes based on the 2022 Kathmandu metropolitan election in Nepal. Due to Nepal's growing social media penetration and its evolving democratic landscape (Kharel, 2024), this context presents a unique opportunity to examine how digital engagement reflects political preferences. Thus, the central research question guiding this study is: To what extent do Facebook emoji reactions reflect AIB during mayoral political campaigns, and how might they serve as predictors of collective political sentiment?

To address this question, we propose the following hypotheses:

H1: Facebook emoji reactions reflect users' emotional responses toward political candidates during electoral campaigns.

H2: Pre-election emotional expressions on Facebook can predict electoral outcomes.

H3: Emoji reactions signal voter sentiments towards candidates, independent of party affiliation.

 https://doi.org/10.1633/JISTaP.2025.13.3.5

These three hypotheses are tested through the analysis of emoji reactions, post shares, and comments on Facebook posts related to campaigning from April 26 to May 10, 2022, focusing on three prominent candidates: Balen Shah, independent; Keshav Sthapit, Communist Party of Nepal-United Marxist-Leninist (CPN-UML); and Sirjana Singh, Nepali Congress (NC). By focusing on Nepal, this research extends the global discourse on social media's electoral influence, which is often centered on Western contexts, and contributes to understanding digital politics in emerging democracies. Insights from the study could inform campaign strategies, enhance electoral predictions, and shape policies for digital governance, offering lessons for South Asia and beyond.

## 4. METHODOLOGY

This study employs a multi-step methodology designed to analyze Facebook engagement as a behavioral signal of voter sentiment and electoral outcomes. Our approach integrates sentiment classification, engagement normalization, predictive analysis, and robustness check. This section is organized into three subsections: (1) Data collection and sample, (2) Sentiment categorization and scoring, and (3) Predictive signal assessment.

### 4.1. Data Collection and Sample

We collected data from the official Facebook pages of three major mayoral candidates in the 2022 Kathmandu metropolitan election: Balen Shah (independent), Keshav Sthapit (CPN-UML), and Sirjana Singh (NC). These mayoral candidates were selected due to their prominence and public engagement throughout the campaign. Given the capital city's central political role and strong digital infrastructure, the election attracted a highly engaged online audience.

Table 1 summarizes the demographic and politician background of the three main candidates in the 2022 Kathmandu mayoral election. Balen Shah, a 35-year-old independent candidate and structural engineer with a background in rap music, appealed strongly to younger, digitally active voters through his outsider image and reform-oriented message. In contrast, Keshav Sthapit, a 69-year-old seasoned politician from the CPN-UML with prior experience as mayor (1997-2006), drew support from older, traditional party members. Sirjana Singh, aged 63 and representing the NC, came from a well-established political lineage and was supported primarily by unions and party-affiliated constituents.

Earlier, a national newspaper conducted an opinion poll in the days before the election (Setopati Team, 2020). The poll results showed that Shah, the independent candidate, was in the lead, followed by Sthapit from the CPN-UML party in second place, and Singh from the NC party in third place. The poll also revealed minimal support for other candidates, reflecting the preferences of voters who had not decided yet (Fig. 2).

To mimic real-time predictive analysis, we retrieved Facebook posts, reactions, comments, and shares from April 26 to May 10, 2022, excluding the final three days to

**Table 1.** Candidates' demographics

| Name | Balen Shah | Keshav Sthapit | Sirjana Thapa |
|---|---|---|---|
| Age | 35 years (born in 1990) | 69 years (born in 1956) | 63 years (born in 1962) |
| Education | Master's degree | Bachelor's degree | Master's degree |
| Profession | Rapper, structural engineer, politician | Politician | Politician |
| Party affiliation | Independent | Communist Party of Nepal-United Marxist Leninist | Nepali Congress |
| Political experience | - | Past mayor (1997-2006) | Former youth leader/chairperson (women's union) |
| Party founded | - | 1991 | 1950 |
| Total membership | - | 855,000 (2021) | 870,106 (2021) |
| Election date | 13 May 2022 | | |
| Constituency | Kathmandu Metropolitan | | |
| Votes cast/total | 61,767/159,906 | 38,117/159,906 | 38,341/159,906 |

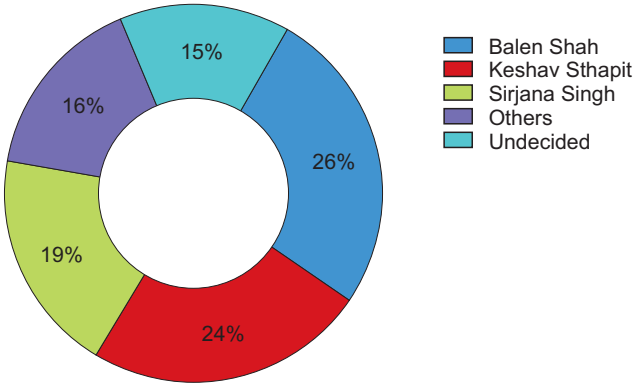



**Fig. 2.** Opinion poll for mayoral candidates in percentages.

avoid post hoc bias. Preprocessing included the removal of spam, duplicates, and non-emotive content to ensure a high-quality dataset for sentiment evaluation.

### 4.2. Sentiment Categorization and Scoring

The initial task involved categorizing Facebook emojis into five impressions: very positive, positive, neutral, negative, and very negative expressions (Table 2), drawing on the emotional models identified by Ekman (1993) and Plutchik (1965), which primarily delineate three types of emotions. We assigned a weighted sentiment score to each reaction type based on its implied emotional value into positive (e.g., love, care, and like), negative (e.g., sad and angry), and neutral (e.g., ha-ha and wow) reactions. These emojis were deemed universally applicable to any text, audio, video post, or share on the platform. To further ensure reliability, we calculated the weights using a pilot study with human annotators, achieving an inter-rater agreement.

Users typically employ the 'like' emoji when expressing agreement and positive sentiment on a post. The 'care' reaction is used to convey empathy, kindness, and understanding, and it is also inclined towards positive feelings. Expressing greater affection, users utilize the 'love' emoji, considered twice as strong as the like and care emoji, to ensure an accurate calculation of positive polarity on Facebook.

On the other hand, 'sad' and 'angry' reactions were used to express disapproval or frustration. Given its heightened intensity, the 'angry' emoji was assigned double the weight of 'sad.' Emojis such as 'ha-ha' and 'wow' were treated as context-dependent: 'ha-ha,' often used sarcastically in political discourse, was classified as negative, while wow was considered neutral to mildly positive, particularly when found in combination with other favorable reactions. Utilizing these data, the overall positive and negative sentiments were calculated accordingly as follows:

**Table 2.** Description of Facebook emoji

| Emojis | Description | Impression |
|---|---|---|
| Love | Users show strong support for the Facebook post | Very positive |
| Care | Users show support for the Facebook post | Positive |
| Like | Users show pleasure towards the Facebook post | Positive |
| Ha-ha | Users show sarcasm, but the feeling is neutral | Neutral |
| Wow | Users show surprise, but the feeling is neutral | Neutral |
| Sad | Users were upset by the Facebook post | Negative |
| Angry | Users were completely upset by the Facebook post | Very negative |

$$\begin{aligned}\text{Total positive sentiment} &= \text{like}+\text{care}+\text{wow}+2\times\text{love}\\ \text{Total negative sentiment} &= \text{ha-ha}+\text{sad}+2\times\text{angry}\end{aligned} \quad (1)$$

Table 3 illustrates the sum of all seven emojis, including followers, posts, comments, and shares. Balen Shah had 30 posts, Keshav Sthapit had 168 posts, and Sirjana Singh had 124 posts on their respective Facebook walls. The cumulative number of posts amounted to 322, and all the emotions and reactions, including followers associated with each post and candidates, were meticulously analyzed.

Using the first formula and data from Table 3, we calculate the total positive and negative sentiments in Table 4 below.

Based on the data of total positive and negative sentiment, each post's net sentiment polarity was calculated as:

$$\text{Sentiment polarity} = \frac{\text{Total positive sentiment-total negative sentiment}}{\text{Total sentiment reactions}} \quad (2)$$

### 4.3. Predictive Signal Assessment

To evaluate the predictive potential of affective engagement of Facebook, we employed a combination of interpretable, signal-based approaches grounded in sentiment metrics and pre-election data. These included:

**Table 3.** Facebook posts and reactions

| Candidate (party) | Balen Shah (independent) | Keshav Sthapit (CPN-UML) | Sirjana Singh (NC) | Total |
|---|---|---|---|---|
| Followers | 1,400,000 | 209,000 | 13,000 | 1,622,000 |
| Posts | 30 | 168 | 124 | 322 |
| Facebook engagement | | | | |
| Comments | 81,041 | 6,629 | 9,127 | 96,797 |
| Shares | 39,875 | 2,150 | 5,185 | 47,210 |
| Likes | 884,000 | 122,288 | 96,127 | 1,102,415 |
| Emotion reactions | | | | |
| Love | 668,700 | 4,897 | 8,257 | 681,854 |
| Care | 84,465 | 399 | 444 | 85,248 |
| Ha-ha | 2,179 | 6,068 | 510 | 8,757 |
| Wow | 5,786 | 159 | 82 | 6,027 |
| Sad | 153 | 39 | 14 | 206 |
| Angry | 69 | 180 | 43 | 292 |

CPN-UML, Communist Party of Nepal-United Marxist-Leninist; NC, Nepali Congress.

**Table 4.** Total positive and negative sentiments

| Candidates | Total positive sentiments | Total negative sentiments |
|---|---|---|
| Balen Shah (independent) | 2,311,651 | 2,470 |
| Keshav Sthapit (CPN-UML) | 132,580 | 6,467 |
| Sirjana Singh (NC) | 113,167 | 610 |
| Total impact | 2,557,398 | 9,547 |

CPN-UML, Communist Party of Nepal-United Marxist-Leninist; NC, Nepali Congress.

(i) Rank-order comparison between digital sentiment metrics and vote shares.

(ii) Pearson correlation to quantify the strength of association.

Key indicators such as average sentiment polarity, engagement rate (e.g., reactions, shares, comments per 1,000 followers), and overall emotional reaction scores were computed for each candidate and compared against their final vote percentage. To further strengthen this analysis, we applied the logarithmic sentiment index developed by Oh and Kumar (2017), which expresses the ratio of positive to negative reactions:

$$\text{Sentiment index}=\text{Ln}\left[\frac{1+\text{Total positive}}{1+\text{Total negative}}\right] \quad (3)$$

This framework enabled a focused assessment of whether Facebook-based affective expressions collected before elections could meaningfully signal voter support. The strong alignment between sentiment indicators and electoral outcomes highlights the potential of social media data for early-stage electoral forecasting and behavioral analysis in digital political environments.

## 5. DATA ANALYSIS AND FINDINGS

### 5.1. Findings from Emotions and Engagement

Table 5 reveals great contrasts in both emotional intensity and engagement patterns across the three mayoral candidates. Balen Shah has an exceptionally high emotion score of 1,429,539, compared to just 17,450 and 16,021 for Sirjana Singh and Keshav Sthapit, respectively. This shows many folds of difference, suggesting that Shah not only generated more interactions, but also evoked far stronger emotional responses among Facebook users. His posts elicited highly affective reactions, indicating that followers were not merely passive consumers of content but deeply engaged participants.

In terms of engagement rate and reactions per 1,000 followers, Shah again dominated, with values of 980 out-

**Table 5.** Engagement rate and sentiment rate

| Candidates | Emotion score | Engagement rate | Reactions rate | Shares per post | Comments per post |
|---|---|---|---|---|---|
| Balen Shah | 1,429,539 | 980 | 980 | 1,329 | 2,701 |
| Keshav Sthapit | 16,021 | 243 | 243 | 12.8 | 39.5 |
| Sirjana Singh | 17,450 | 355 | 355 | 42.8 | 73.6 |

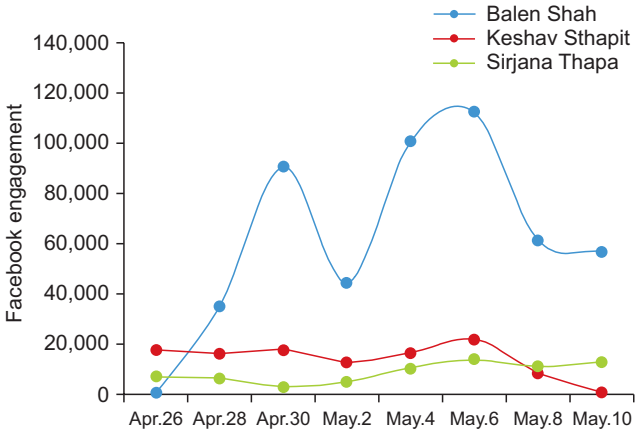



**Fig. 3.** Daily Facebook engagement trend.

pacing Singh (355) and Sthapit (243) by fair margins. This indicates that his emotional appeal translated into active engagement on a per-follower basis, further reinforcing the notion of authentic resonance between candidate and audience.

Notably, Shah also led in shares per post (1,329) and comments per post (2,701), both orders of magnitude higher than Singh (42.8 shares, 73.6 comments) and Sthapit (12.8 shares, 39.5 comments). High sharing rates suggest that Shah's content was not only engaging but amplified organically through network effects, reaching a broader audience and reinforcing campaign visibility. The volume of comments further reflects a high level of public discourse and two-way interaction surrounding his candidacy.

Together, these metrics reflect a candidate whose campaign successfully leveraged affective communication and participatory momentum on Facebook. The scale and depth of digital engagement observed for Shah support the broader finding that emotion-driven social media activity is a meaningful predictor of electoral support.

Similarly, Fig. 3 shows the average Facebook engagement (likes, shares, and comments) for each candidate, calculated over two-day intervals from April 26 to May 10, 2022. Balen Shah began his campaign on April 28 but quickly gained traction, consistently leading with significantly higher engagement, peaking at nearly 120,000 on May 4 before declining later. In contrast, Sthapit and Singh showed relatively low and stable engagement, with minor fluctuations. The stark contrast highlights Shah's strong digital presence and sustained public interest during the campaign period, aligning with his eventual electoral success.

**Table 6.** Sentiments vs. voting results

| Candidates | Net sentiment polarity | Sentiment index | Actual vote share |
|---|---|---|---|
| Balen Shah | 1.00 (1st) | 6.84 (1st) | 38.64% (1st) |
| Keshav Sthapit | 0.91 (3rd) | 3.02 (3rd) | 24.03% (2nd) |
| Sirjana Singh | 0.99 (2nd) | 5.22 (2nd) | 23.83% (3rd) |

## 5.2. Findings from Predictive Signals

To assess the predictive strength of Facebook reactions, we compare each candidate's pre-election sentiment indicators with their actual vote share.

A rank-order comparison was conducted to evaluate the alignment between pre-election sentiment metrics and actual electoral outcomes based on both average sentiment polarity and the logarithmic sentiment index. Balen Shah consistently ranked first across all three dimensions—average polarity (1.00), sentiment index (6.84), and vote share (38.64%), demonstrating a strong alignment between emotional engagement and electoral success. Although Singh and Sthapit's vote shares were nearly identical (23.83% and 24.03%, respectively), their positions were reversed compared to their sentiment ranks (Table 6). Despite this minor discrepancy, the overall ranking demonstrates strong alignment between affective digital engagement and electoral performance, reinforcing the predictive value of pre-election Facebook sentiment signals.

To further evaluate the relationship between online sentiment and electoral outcomes, a Pearson correlation was conducted using both the sentiment index and net sentiment polarity. The correlation between the log-transformed sentiment index and vote share was r=0.81, indi-

 https://doi.org/10.1633/JISTaP.2025.13.3.5

cating a strong positive association. This metric reflects the relative dominance of positive sentiment over negative. A moderate correlation was observed between net sentiment polarity and vote share (r=0.58), suggesting that the overall emotional tone of user reactions (particularly the balance between positive and negative sentiment) was meaningfully aligned with electoral performance. This was evident in Shah's case, where high user engagement and affective support corresponded closely with voting outcomes.

These findings suggest that Facebook sentiment signals (net polarity and sentiment index) offer a reliable behavioral proxy for electoral support when measured prior to elections. However, minor rank mismatches between sentiments and vote share, as seen between Singh and Sthapit, indicate that offline factors such as party affiliation, demographic outreach, or mobilization strategies may still influence voting outcomes beyond what is captured digitally.

### 5.3. Robustness Checks

To ensure the reliability of our findings, we conducted two robust validation tests. First, we conducted a time-truncation analysis by excluding the final three days of Facebook data (May 8-10) from the dataset. After truncating, 201 posts remained across all candidates. Keshav Sthapit had the most posts, followed by Sirjana Singh and Balen Shah, consistent with the post distribution in the full dataset. We then ran the sentiment analysis on the truncated dataset and compared the results with the actual vote share. While the correlation between sentiment analysis and vote share declined slightly (r=0.78), the ranking of candidates and the overall predictive direction remain unchanged. This suggests that the model's predictive strength is not solely dependent on last-minute surges in online engagement but instead reflects more stable affective patterns throughout the campaign (Jaidka et al., 2019).

Second, we tested the robustness of our sentiment scoring method by removing the emotion-based weighting assigned to specific emoji reactions. In this unweighted model, all reactions were treated equally, so each emoji contributed to one unit regardless of its affective intensity. While this simplified approach still yielded a strong correlation between sentiment polarity and vote share (r=0.62), the predictive precision declined slightly compared to the weighted model. This result reinforces the theoretical and empirical value of emotion-weighted scoring, which captures varying degrees of affective expression and improves the sensitivity of the sentiment-vote relationship (Giuntini et al., 2019; Tian et al., 2017).

These checks confirm that our results are not dependent on a specific time window and emoji weighting scheme, strengthening the claim that Facebook emoji reactions are robust behavioral signals of electoral sentiment (Hauthal et al., 2019).

### 5.4. Findings from Text Analytics

We conducted text analytics of Facebook comments directed toward mayoral candidates Balen Shah, Keshav Sthapit, and Sirjana Singh to uncover the most frequently mentioned messages, terms, and sentiments associated with each. Using techniques such as word clouds and sentiment analysis, we gained valuable insights into the narratives and conversations of the electorate. Word clouds offer a visually intuitive method for highlighting commonly used terms, making them crucial for identifying key public conversations.

We began by analyzing the word cloud for Balen, who received the most comments—76,141 in total, which is 85% more than the combined total for the other two candidates. Most of these comments were positive, with phrases like '*I support Balen*,' '*my vote for Balen*,' and '*time for change*' appearing frequently (Fig. 4). Additionally, many comments from users included terms like '*youth*' and similar phrases, such as '*vote for youth*,' '*youth for change*,' '*youths of Kathmandu*,' and '*Balen rock*.' Balen's posts primarily featured photos from campaign rallies and live videos. He was the first to announce his mayoral candidacy and start a Facebook campaign. Judging by the comments and rally videos, it appears that he had many young supporters. On the negative side, Shah was underestimated for his independent candidacy without the support of any political party, which is a significantly necessary factor in Nepal's partisan political landscape.

In the word clouds for Keshav Sthapit and Sirjana Singh, we observed a noticeable concentration of support for their rival, Shah, with comments like '*vote for Balen*' and '*this time it is Balen*' appearing occasionally. Despite this, there were numerous positive comments for both candidates. Keshav Sthapit's posts received comments such as '*good luck*,' '*congratulations*,' and '*be the best again*,' reflecting his previous tenure as mayor two decades earlier from 1997 to 2006, as he sought this position for a second time. Similarly, Sirjana Singh's posts attracted comments including '*best wishes*,' '*Jay (hail) Nepal*,' '*congratulations*,' and '*gentlewoman*.' However, despite support from various unions, including those representing teachers, students, traders, and drivers, both candidates also faced substantial




**Fig. 4.** Word clouds for Balen Shah (top left), Keshav Sthapit (top right), and Sirjana Singh (bottom).

criticism. Keshav Sthapit was criticized with comments like '*please talk less*' and '*you are not eligible*', while Singh encountered remarks such as '*no nepotism*,' and '*shame to Ganesh Man*.' Mrs. Singh comes from a prominent political background, with her husband serving as a parliamentarian and her father-in-law, Ganesh Man, a significant figure in Nepal's political history. She was often criticized for alleged candidacy bargaining and partisan politics. Additionally, users who expressed support were derogatorily referred to as '*jholye*' (porter) and '*bheda*' (sheep). The term '*jholye*' is used to describe people who blindly carry a politician's ideology or actions, while '*bheda*' refers to those who follow their leaders or the crowd without much knowledge.

Our analysis shows that in the days leading up to the election, there was a high volume of positive comments, particularly targeting Shah, and negative comments targeting Keshav Sthapit and Sirjana Singh, which may have contributed to the polarity for them. Unfortunately, Facebook posts about election rallies and meetings by the latter two major political candidates were either sparsely commented on or met with negative sentiments.

## 6. DISCUSSION

To our knowledge, this study marks the first empirical attempts to analyze Nepal's mayoral election through the lens of Facebook emoji reactions and behavioral signals of voter sentiment, interpreted with the AIB framework. Nepal's evolving digital landscape and unique political dynamics make this investigation timely and significant, contributing to broader understandings of digital political engagement in emerging democracies (Boulianne, 2020; Boulianne & Larsson, 2023). By examining how emoji reactions reflect emotional information processing, this study advanced AIB by showing how digital affective signals shape voter decision-making (Kuhlthau, 1991; Savolainen, 2015).

The findings support all three hypotheses. First, Facebook emoji reactions reflect users' emotional responses to political candidates, serving as proxies for AIB processes such as emotional evaluation and preference formation (Nahl, 2007). Sentiment distribution and polarity scores closely aligned with voting behavior, supporting studies that endorse social media sentiment as a reliable proxy for public opinion (Ceron et al., 2023; Sandoval-Almazan &

 https://doi.org/10.1633/JISTaP.2025.13.3.5

Valle-Cruz, 2020). Similar findings from India (Rita et al., 2023) reinforce this, although Jungherr et al. (2012) caution that social media users may not fully represent the electorate, suggesting emoji reactions be interpreted as behavioral signals rather than precise forecasts.

Second, the predictive value of engagement was evident, with candidates showing higher positive sentiment and interactions (reactions, shares, comments) performing better at the ballot box. Rank order analysis (Balen at number 1 on all) and Pearson's correlation (r=0.81 for sentiment index, r=0.58 for polarity) confirm that emoji-driven engagement predicts vote shares, echoing Bond et al. (2012) and Oh and Kumar (2017). However, contextual factors like limited digital access, misinformation, or offline mobilization may mediate online sentiments in different environments (Sandoval-Almazan & Valle-Cruz, 2020).

Third, emotional engagement extended beyond party lines, with independent candidate Balen Shah generating the most intense positive reactions due to his perceived integrity and issue alignment. This suggests that emoji reactions reflect AIB's assessment of candidates' authenticity, and this perceived authenticity, as interpreted through AIB, subsequently influence voter decisions beyond party loyalty (Savolainen, 2015). Shah's passionate, issue-focused communication (Gautam, 2022) aligns with research on personalized engagement (Enli & Skogerbø, 2013), with parallels to nonpartisan campaigns such as Emmanuel Macron's in France (Cole, 2022). However, party identity remains influential where echo chambers persist (Barberá, 2015).

Overall, this study highlights emoji reactions as AIB-driven indicators, offering diagnostic and predictive insights into voter sentiment in Nepal. Timer-truncation and unweighted sentiment analyses ensured robust results, providing a methodological foundation for future studies in emerging democracies.

## 7. IMPLICATIONS FOR THE STUDY

This study highlights the importance of AIB in understanding digital political engagement on social media platforms, contributing to both AIB and political communication. Drawing from emotion theory (Ekman, 1993; Plutchik, 1965) and the AIB framework (Kuhlthau, 1991; Nahl, 2007; Savolainen, 2015), the findings demonstrate that emotionally charged expressions, specifically Facebook emoji reactions, serve as reliable behavioral predictors of electoral outcomes. The positive correlation between pre-election sentiment polarity and vote share validates the role of collective emotional information processing in shaping voter decision-making, advancing AIB by showing how digital affective signals reflect and predict real-world behaviors.

This study's results extend Kuhlthau (1991)'s ISP model by illustrating how emotions like enthusiasm (expressed via love and like reactions) or dissatisfaction (angry and sad reactions) guide users' interactions with political information, influencing their voting decisions. Similarly, Nahl (2007)'s emphasis on affective states as drivers of information behavior is supported by the high engagement rates and emotional intensity observed for Balen Shah, whose campaign resonated with users' desires for change and authenticity. By demonstrating how emoji reactions serve as proxies for collective emotional processing, this study enriches the AIB framework, showing that these reactions are not merely expressions but predictive indicators of behavioral decision-making in democratic contexts (Fulton, 2009; Lopatovska & Arapakis, 2011).

The findings offer multifaceted practical implications for political strategists, campaign managers, and political communication agencies in digitally connected societies. First, social media platforms such as Facebook offer real-time feedback on voter sentiment based on emoji reactions (love, care, like, sad, wow, ha-ha, and angry), enabling dynamic adjustment to campaign messaging. Political strategists can employ analytics tools like Hootsuite or Sprout Social to monitor sentiment shifts, craft emotionally resonant content that fosters stronger voter connections, and adjust messaging in real time. For instance, they might emphasize empathetic narratives when "sad" reactions spike to build trust and engagement effectively.

Second, Balen Shah's success as an independent candidate, driven by strong emotional appeal, demonstrates that affective engagement is not tied to political party affiliation, aligning with AIB's findings on the role of emotional authenticity in shaping public trust and engagement (Savolainen, 2015). This trend suggests that modern electorates, particularly those active in digital space, respond more positively to candidates who communicate clearly to show genuine concern on social issues, rather than relying on traditional party-based rhetoric. Campaign managers can focus on sincere, issue-based communication that centers a candidate's vision and relatable appeal. Practical steps include posting a video about their effort to address local healthcare access or hosting live Q&A sessions to directly engage with voters' concerns.

Third, this study signals a broader shift from tradition-

al clientelist and identity-driven politics to more personality-driven nonpartisan politics. Emotional expressions on platforms like Facebook indicate that voters increasingly value accountability, transparency, and competence over traditional loyalty-based appeals. Political actors or communication agencies should adapt messaging strategies, accordingly, focusing on relevant emotionally grounded narratives that build trust and engagement. This could involve crafting authentic stories around candidates' values and actions, emphasizing competence and responsiveness to public concerns, and leveraging emotional cues to connect with audiences on a personal level.

## 8. LIMITATIONS AND CONCLUSION

This research has several limitations that warrant consideration. First, Facebook users do not always truthfully represent their sentiments and behaviors in political posts, thus raising doubts about whether reactions to Facebook posts truly reflect their voting behavior. Moreover, citizens often criticize politicians regarding accountability, scandals, and campaigns. Online activities tend to elicit more impulsive and aggressive behavior, which may influence sentiments towards political parties. Secondly, contextual factors such as the political inclinations of voters, Internet accessibility, and candidates' political backgrounds were not considered, factors that could influence voters' decisions and sentiments. Future studies could explore the varied effects of these contextual elements.

Despite these limitations, this study contributes to a deeper understanding of affective engagement in digital political spaces, using the 2022 Kathmandu metropolitan election as a case to explore how Facebook reactions function as behavioral signals of voter sentiment. By analyzing over 300 campaign-period posts and millions of emoji reactions, the study provides evidence that Facebook engagement, particularly affective expressions, can offer insight into public sentiment and predict electoral outcomes.

As the use of social media in political communication continues to expand, this study highlights the significance of interpreting emotional engagement not only as online behavior, but as part of broader affective information practices. Facebook reactions, when analyzed thoughtfully, can serve as indicators of political alignment, campaign resonance, and shifting voter priorities, making them a vital tool for political analysis and strategy in diverse democratic contexts.

## CONFLICTS OF INTEREST

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

## REFERENCES

Agarwal, A., & Bansal, V. (2020, May 5-7). Exploring sentiments of voters through social media content: A case study of 2017 assembly elections of three states in India. In J. Filipe, M. Smialek, A. Brodsky, & S. Hammoudi (Eds.), *Proceedings of the 22nd International Conference on Enterprise Information Systems - Volume 1: ICEIS* (pp. 596-602). SciTePress. https://doi.org/10.5220/0009517105960602

Barberá, P. (2015). *How social media reduces mass political polarization. Evidence from Germany, Spain, and the U.S.* Paper presented at 2015 APSA Conference, San Francisco, CA, USA. http://pablobarbera.com/static/barbera_polarization_APSA.pdf

Bhattarai, P. (2023). The use of social media for political socialization in Nepal: An effectiveness analysis of platforms. *The Outlook: Journal of English Studies*, 14(1), 46-59. https://doi.org/10.3126/ojes.v14i1.56656

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415), 295-298. https://doi.org/10.1038/nature11421

Boulianne, S. (2020). Twenty years of digital media effects on civic and political participation. *Communication Research*, 47(7), 947-966. https://doi.org/10.1177/0093650218808186

Boulianne, S., & Larsson, A. O. (2023). Engagement with candidate posts on Twitter, Instagram, and Facebook during the 2019 election. *New Media & Society*, 25(1), 119-140. https://doi.org/10.1177/14614448211009504

Ceron, A., Pagano, G., & Bordignon, M. (2023). Facebook as a media digest: User engagement and party references to hostile and friendly media during an election campaign. *Journal of Information Technology & Politics*, 20(4), 454-468. https://doi.org/10.1080/19331681.2022.2157360

Chambers, N., Bowen, V., Genco, E., Tian, X., Young, E., Harihara, G., & Yang, E. (2015, September). Identifying political sentiment between nation states with social media. In L. Màrquez, C. Callison-Burch, & J. Su (Eds.), *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 65-75). Association for Computational Linguistics. https://doi.org/10.18653/v1/D15-1007

Cole, A. (2022). *Understanding Macron's political leadership and his impact on French politics*. https://blogs.lse.ac.uk/

europpblog/2022/05/03/understanding-macrons-political-leadership-and-his-impact-on-french-politics/

Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, 48(4), 384-392. https://doi.org/10.1037/0003-066X.48.4.384

Enli, G. S., & Skogerbø, E. (2013). Personalized campaigns in party-centred politics. *Information, Communication & Society*, 16(5), 757-774. https://doi.org/10.1080/1369118X.2013.782330

Fulton, C. (2009). The pleasure principle: The power of positive affect in information seeking. *Aslib Proceedings*, 61(3), 245-261. https://doi.org/10.1108/00012530910959808

Gautam, S. (2022). *Kathmandu's 'Balen Phenomeno*'. https://nepalitimes.com/opinion/comment/kathmandu-s-balen-phenomenon

Giuntini, F. T., Ruiz, L. P., Kirchner, L. D. F., Passarelli, D. A., Reis, M. D. J. D. 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

Guess, A. M., Nyhan, B., & Reifler, J. (2020). Exposure to untrustworthy websites in the 2016 US election. *Nature Human Behaviour*, 4(5), 472-480. https://doi.org/10.1038/s41562-020-0833-x

Haryanto, B., Ruldeviyani, Y., Rohman, F., Julius Dimas, T. N., Magdalena, R., & Muhamad, Y. F. (2019). Facebook analysis of community sentiment on 2019 Indonesian presidential candidates from Facebook opinion data. *Procedia Computer Science*, 161, 715-722. https://doi.org/j.procs.2019.11.175

Hauthal, E., Burghardt, D., & Dunkel, A. (2019). Analyzing and visualizing emotional reactions expressed by emojis in location-based social media. *ISPRS International Journal of Geo-Information*, 8(3), 113. https://www.mdpi.com/2220-9964/8/3/113

Jaidka, K., Ahmed, S., Skoric, M., & Hilbert, M. (2019). Predicting elections from social media: A three-country, three-method comparative study. *Asian Journal of Communication*, 29(3), 252-273. https://doi.org/10.1080/01292986.2018.1453849

Jungherr, A., Jürgens, P., & Schoen, H. (2012). Why the pirate party won the German election of 2009 or the trouble with predictions: A response to Tumasjan, A., Sprenger, T. O., Sander, P. G., & Welpe, I. M. "Predicting elections with Twitter: What 140 characters reveal about political sentiment". *Social Science Computer Review*, 30(2), 229-234. https://doi.org/10.1177/0894439311404119

Khan, A., Majumdar, D., & Mondal, B. (2025). Sentiment analysis of emoji fused reviews using machine learning and Bert. *Scientific Reports*, 15(1), 7538. https://doi.org/10.1038/s41598-025-92286-0

Kharel, A. B. (2024). Cyber-politics: Social media's influence on political mobilization. *Journal of Political Science*, 103-115. https://consensus.app/papers/cyberpolitics-social-medias-influence-on-political-kharel/292e80fcbd4959e39fb553450e1ea662/

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, 42(5), 361-371. https://doi.org/10.1002/(SICI)1097-4571(199106)42:5<361::AID-ASI6>3.0.CO;2-#

Lopatovska, I., & Arapakis, I. (2011). Theories, methods and current research on emotions in library and information science, information retrieval and human–computer interaction. *Information Processing & Management*, 47(4), 575-592. https://doi.org/10.1016/j.ipm.2010.09.001

Muraoka, T., Montgomery, J., Lucas, C., & Tavits, M. (2021). Love and anger in global party politics: Facebook reactions to political party posts in 79 democracies. *Journal of Quantitative Description: Digital Media*, 1. https://doi.org/10.51685/jqd.2021.005

Nahl, D. (2007). Social–biological information technology: An integrated conceptual framework. *Journal of the American Society for Information Science and Technology*, 58(13), 2021-2046. https://doi.org/10.1002/asi.20690

Nahon, K., & Hemsley, J. (2013). *Going viral*. Polity Press.

Oh, C., & Kumar, S. (2017, July 16-20). *How Trump won: The role of social media sentiment in political elections.* Paper presented at PACIS 2017 Proceedings, Langkawi Island, Malaysia. https://aisel.aisnet.org/pacis2017/48

Persily, N., & Tucker, J. A. (2020). Conclusion: The challenges and opportunities for social media research. In N. Persily & J. A. Tucker (Eds.), *Social Media and Democracy* (pp. 313-331). Cambridge University Press. https://www.cambridge.org/core/product/232F88C00A1694FA25110A318E9CF300

Plutchik, R. (1965). What is an emotion? *The Journal of Psychology*, 61(2), 295-303. https://doi.org/10.1080/00223980.1965.10543417

Rita, P., António, N., & Afonso, A. P. (2023). Social media discourse and voting decisions influence: Sentiment analysis in tweets during an electoral period. *Social Network Analysis and Mining*, 13(1), 46. https://doi.org/10.1007/s13278-023-01048-1

Saifeddine, S., & Abdellatif Chakor, M. (2024, April 23-25). Systematic literature reviews in political marketing: Behavior, influence, and trust in the era of big data and artificial intelligence. In Y. Farhaoui (Ed.), *Artificial intelligence, big*

*data, IOT and block chain in healthcare: From concepts to applications* (pp. 112-126). Springer, Cham. https://doi.org/10.1007/978-3-031-65018-5_11

Sandoval-Almazan, R., & Valle-Cruz, D. (2020). Sentiment analysis of Facebook users reacting to political campaign posts. *Digital Government: Research and Practice*, 1(2), Article 12. https://doi.org/10.1145/3382735

Savolainen, R. (2015). Cognitive barriers to information seeking: A conceptual analysis. *Journal of Information Science*, 41(5), 613-623. https://doi.org/10.1177/0165551515587850

Setopati Team (2022). *Setopati projection: Balen Shah to be elected Kathmandu Mayor*. https://en.setopati.com/political/158478

Theocharis, Y., Lowe, W., van Deth, J. W., & García-Albacete, G. (2015). Using Twitter to mobilize protest action: Online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements. *Information, Communication & Society*, 18(2), 202-220. https://doi.org/10.1080/1369118X.2014.948035

Tian, Y., Galery, T., Dulcinati, G., Molimpakis, E., & Sun, C. (2017, April). Facebook sentiment: Reactions and emojis. In L. W. Ku & C. T. Li (Eds.), *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media* (pp. 11-16). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-1102

Tumasjan, A., Sprenger, T., Sandner, P., & Welpe, I. (2010). Predicting elections with Twitter: What 140 characters reveal about political sentiment. *Proceedings of the International AAAI Conference on Web and Social Media*, 4(1), 178-185. https://doi.org/10.1609/icwsm.v4i1.14009

Upreti, A., Pokhrel, L., Ban, H., & Gurung, K. (2025). Decoding digital campaigns: A multi-method analysis of Facebook posts during 2022 parliamentary elections in Nepal. *Asia Pacific Journal of Information Systems*, 35(1), 195-215. https://doi.org/10.14329/apjis.2025.35.1.195

Wagle, P. R. (2024). *The effect of social media on result of election in metropolitan city: Special focus on Kathmandu, Nepal* (master's thesis). Arcada University, Helsinki, Finland. https://www.theseus.fi/handle/10024/865306

Wisniewski, P., Badillo-Urquiola, K., Ashtorab, Z., & Vitak, J. (2020). Happiness and fear: Using emotions as a lens to disentangle how users felt about the launch of Facebook reactions. *ACM Transactions on Social Computing,* 3(4), Article 20. https://doi.org/10.1145/3414825

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

RESEARCH PAPER
J Inf Sci Theory Pract 13(3): 85-104, 2025

eISSN : 2287-4577 pISSN : 2287-9099

https://doi.org/10.1633/JISTaP.2025.13.3.6

# Exploring the Knowledge Structure and Thematic Evolution of NTIS-Related Research in South Korea (2008-2024): A Bibliometric Analysis

**Moon-Koo Kim**
Department of Business Consulting, Daejeon University, Daejeon, Korea
E-mail: wisemkkim@naver.com

**Byung-Hwan Hyun***
Department of Business Consulting, Daejeon University, Daejeon, Korea
E-mail: hyunvv@gmail.com

## ABSTRACT

This study represents the first systematic analysis of research related to the National Science & Technology Information Service (NTIS)—South Korea's comprehensive national Research and Development (R&D) information portal—from 2008 to 2024, utilizing bibliometric analysis, thematic analysis, and selective literature review. The study aims to identify key research trends, collaborative patterns, and thematic evolution in NTIS-related research to understand its knowledge structure and provide directions for scholarly and system development. The analysis reveals that despite limited publications (169 articles), NTIS research demonstrates steady growth with 13.9% annual growth. Distribution analysis using Herfindahl-Hirschman Index (HHI) indicates low concentration levels across authors, institutions, and journals, though citations show higher concentration. Korea Institute of Science and Technology Information (KISTI) leads in all performance indicators and serves as the central hub in inter-institutional collaboration networks. Research scope has expanded into diverse academic fields, with growth in interdisciplinary studies, business administration, and engineering. While studies using NTIS data have increased, research on system improvement has declined. Thematic evolution analysis reveals nine research clusters and shows a shift from system development towards R&D trend analysis, performance evaluation, and advanced analytical methodologies, including Artificial Intelligence (AI) technologies. Major research themes include NTIS system implementation, data quality management, R&D investment-performance analysis, economic impact assessment, and analytical techniques. This study indicates that NTIS's function is evolving beyond R&D information service into a data-driven academic analysis infrastructure supporting evidence-based knowledge production across disciplines. To address limited research volume and domestic concentration, this study recommends activating NTIS-related academic communities, expanding international participation through compatibility with global systems, and building an intelligent platform for customized services.

**Keywords:** National Science & Technology Information Service (NTIS), bibliometric analysis, research trends, thematic evolution, selective literature review

**Received:** April 1, 2025 **Revised:** July 3, 2025
**Accepted:** July 14, 2025 **Published:** September 30, 2025

***Corresponding Author:** Byung-Hwan Hyun
https://orcid.org/0009-0000-8339-1286
**E-mail:** hyunvv@gmail.com

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 (AccessON). All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/). Under this license, the authors retain full ownership of their work, while permitting anyone to use, distribute, and reproduce the content in any medium, as long as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

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

## 1. INTRODUCTION

Science and Technology (S&T) information systems are increasingly important as critical infrastructure for the production, sharing, and utilization of knowledge within national Research and Development (R&D) ecosystems. By integrating research information, these systems enhance national innovation capabilities, support the efficient allocation of resources, and provide critical evidence for R&D policy decisions (Oh, 2008; Won et al., 2023). Major countries operate comparable R&D information services that reflect their policy environments and institutional structures to enhance public R&D investment efficiency and strengthen performance management. The United States operates the Research Portfolio Online Reporting Tools, Expenditures, and Results (RePORTER), the European Union (EU) maintains the Community Research and Development Information Service (CORDIS), and Japan administers the Grants-in-Aid for Scientific Research Database (KAKEN).

In South Korea, the National Science & Technology Information Service (NTIS)—a comprehensive knowledge portal providing lifecycle information on national R&D activities—was established in 2008 to manage information across the entire national R&D lifecycle (Korea Institute of Science and Technology Evaluation and Planning [KISTEP], 2013). NTIS has contributed to improving investment efficiency and research productivity by systematically linking data from R&D planning to outcome dissemination (KISTEP, 2025; Park, 2018). Recently, it has evolved into an intelligent platform leveraging big data and Artificial Intelligence (AI) technologies, thereby playing a critical role in enhancing South Korea's S&T innovation and global competitiveness (Korea Institute of Science and Technology Information [KISTI], n.d.; Lee & Yang, 2021; Lee et al., 2020). Alongside this expanding role of NTIS, the active promotion of NTIS utilization within South Korea's national R&D ecosystem and its enhanced role as a central national S&T information platform have led to an increasing trend of conducting analyses based on NTIS data, resulting in the steady expansion of NTIS-related research.

However, despite this clear growth in activity, a significant research gap remains: No comprehensive and systematic analysis has been conducted to date that maps the research trends, traces the evolution of academic collaboration networks, or identifies the major thematic areas within this body of NTIS-related studies. This limitation hinders a holistic understanding of the knowledge production dynamics and impact patterns specific to this research area, thereby limiting the ability to derive meaningful insights that could advance scholarly research focused on NTIS itself or enhance the broader utilization of NTIS-based data.

To address this deficiency, this study undertakes a multi-faceted analysis. We employ bibliometric and network analyses specifically to uncover the underlying knowledge structures and collaborative relationships that have formed within this research area. Concurrently, thematic analysis and selective review are utilized to identify key research topics and track their thematic evolution over time, thereby fostering an understanding of the research area's qualitative growth and informing future research directions. This analytical approach provides insights beyond mapping the research landscape; it offers valuable insights into the broader socio-informational impacts of a national R&D information system like NTIS on its academic ecosystem, particularly concerning knowledge influence and scholarly concentration versus diversity. This case analysis of South Korea's NTIS furnishes transferable policy and practical implications for other countries currently in the process of developing or enhancing similar R&D information infrastructures.

This study systematically examines the knowledge structure and major research themes through bibliometric analysis, network analysis, thematic analysis, and literature review of NTIS-related research literature from 2008 to 2024. This study addresses the following research questions:

■ RQ1: How have knowledge production and impact in NTIS-related research evolved over time? This study analyzes temporal patterns in published studies to identify these trends.

■ RQ2: To what extent do disparities exist among research entities in publication and citation performance? This study assesses concentration in publications and citations across authors, institutions, and journals.

■ RQ3: Which research entities lead NTIS-related studies? This work identifies prominent authors, institutions, and journals based on publication volume and citation impact.

■ RQ4: How are collaborative relationships formed among institutions? This study analyzes institutional collaboration networks and key players driving inter-organizational cooperation.

■ RQ5: What themes characterize NTIS-related research, and how have they evolved? This study identifies

 https://doi.org/10.1633/JISTaP.2025.13.3.6

major research themes through keyword analysis, academic field analysis, and thematic clustering, and explores their contextual meaning and development through literature review.

## 2. NTIS: DEVELOPMENT AND COMPARISON

### 2.1. Background and Development of NTIS in South Korea

The NTIS of South Korea, established in 2008, serves as a knowledge portal providing lifecycle information on national R&D activities, spanning planning, implementation, performance evaluation, and outcome dissemination (KISTEP, 2013; Kwak et al., 2022). Under the Ministry of Science and ICT (MSIT), the KISTI has led NTIS development to strengthen national R&D competitiveness by integrating scientific information across government departments and institutions (KISTI, n.d.). KISTI's responsibilities include maintaining the NTIS system, establishing data linkage frameworks between departments and institutions, managing data quality, improving services based on user needs, and advancing services through emerging technologies (KISTI, 2019). KISTI has improved the NTIS system through data quality enhancement, user-centered functionality, and analytical capability development (Lee & Yang, 2021). NTIS is pursuing service advancement through intelligent service provision using big data analysis and AI technologies (KISTEP, 2025; Lee et al., 2020). As of January 2025, NTIS collaborates with 18 specialized government agencies that oversee R&D programs to collect research outcome data. The system supports over 22,000 registered users and provides access to more than 1.15 million project data entries, serving as a hub for national R&D information in South Korea (KISTI, n.d.).

### 2.2. Role of NTIS in National R&D Ecosystem

NTIS functions as a cornerstone of South Korea's national R&D ecosystem, delivering services across the R&D lifecycle, and serves within the national research data governance framework (KISTEP, 2025; Kwak et al., 2022; Lee & Yang, 2021). It provides integrated access to information—encompassing national R&D programs, projects, researchers, facilities, equipment, and research outcomes—enabling stakeholders to make evidence-based decisions (KISTI, n.d.; Lee & Kim, 2021; Shmagun et al., 2022; Won et al., 2023). NTIS enhances efficiency through multiple mechanisms: preventing redundant project investments, encouraging shared utilization of research infrastructure, enabling performance verification, and ensuring transparent dissemination of R&D information (KISTEP, 2025; KISTI, n.d.). It promotes the practical application and impact of research by increasing the accessibility of R&D outcomes. The economic contributions of NTIS are substantial, generating multiplier effects that stimulate production, create added value, and promote job growth within the South Korean economy (Huh et al., 2009; Park, 2018; Park et al., 2012). From a policy standpoint, NTIS enhances transparency in national R&D investments and facilitates evidence-based policy formulation (KISTEP, 2025; Won et al., 2023). This data-driven approach improves science policies by providing information for evaluating researchers and institutions, identifying promising technological areas, and informing R&D budget allocation. Beyond domestic contributions, NTIS has expanded its international cooperation since the mid-2010s, collaborating with Costa Rica, Belarus, and Vietnam to establish S&T information systems and improve R&D information management capabilities (KISTI, n.d.).

### 2.3. Comparison with Major International Systems

Major countries operate national R&D information systems reflecting their policy environments and institutional structures to enhance public R&D investment efficiency and strengthen performance management (KISTEP, 2013; KISTI, n.d.; Won et al., 2023).

The United States operates the RePORTER system, managed by the National Institutes of Health (NIH), to transparently disclose the status and outcomes of its federal government R&D investments and enhance public access. This database encompasses research projects funded not only by NIH but also by other major agencies and operating divisions within the Department of Health and Human Services (HHS). RePORTER provides integrated information covering the entire research lifecycle, including research objectives, budget details, principal investigators and their affiliated institutions, published papers, patents, and clinical trial results. Users can perform in-depth searches and analyses based on criteria such as diseases, research areas, institutions, and research types. This system serves as a core infrastructure that supports researchers in identifying research trends and discovering collaboration opportunities, and aids policymakers in evaluating R&D investments and formulating strategies (NIH, n.d.). Japan administers the Grants-in-Aid for Scientific Research (KAKENHI) program through the Japan Society for the Promotion of Science (JSPS) under the Ministry of Education, Culture, Sports, Science and Technology (MEXT) guidance (JSPS, n.d.). The National

Institute of Informatics (NII) operates the KAKEN Database to enhance transparency (NII, n.d.). This system provides data on research projects, including objectives, investigators, institutions, funding amounts, durations, and outcomes. Within the EU, the CORDIS, operated by the European Commission (EC) under the Directorate-General for Research and Innovation (DG RTD), provides information on projects funded by EU programs, including Horizon Europe and Horizon 2020. The platform publishes project descriptions, organizations, funding levels, and deliverables, promoting collaboration within the EU research community (EC, n.d.).

Compared to international systems, South Korea's NTIS exhibits several distinctive features (Oh, 2008; Shmagun et al., 2022; Won et al., 2023). First, unlike systems focusing on specific funding bodies (e.g., NIH/HHS in RePORTER, KAKENHI in KAKEN) or programs (e.g., EU framework programs in CORDIS), NTIS integrates nationwide information on government-funded R&D activities—including investments, projects, outcomes, researchers, research infrastructure, and technology statistics—across ministries and scientific domains (KISTEP, 2013; KISTI, n.d.). Second, NTIS supports information access across the R&D lifecycle by collecting and linking data from government systems—from project planning and evaluation to execution and utilization (KISTEP, 2025; Lee & Yang, 2021). Third, NTIS is structured to support national R&D policy formulation, strategic budgeting, and performance evaluation by offering integrated data for decision-making (KISTEP, 2025; Won et al., 2023). It thus plays a key role in evidence-based R&D policymaking in South Korea. Lastly, while platforms like CORDIS focus on disseminating research results, NTIS facilitates R&D outcome utilization by linking to technology transfer platforms, industry matching services, and commercialization channels (KISTI, n.d.; Park, 2018). These functions maximize the socio-economic impact of public R&D investments beyond information disclosure. In summary, NTIS serves as a comprehensive platform, playing a central role in South Korea's R&D ecosystem by providing centralized information, integrating lifecycle data for policy support, and promoting practical utilization of research outcomes.

## 3. RESEARCH METHODOLOGY

This study applied an integrated research methodology that can be characterized as comprehensive bibliometric analysis combined with thematic analysis and selective review, to systematically examine the knowledge structure and thematic evolution of NTIS-related academic literature. The methodology encompasses data collection and preprocessing, bibliometric analysis, network analysis, concentration analysis, thematic analysis, and targeted literature examination, as shown in Fig. 1.

Bibliographic data required for the research were collected from the Korea Citation Index (KCI) and international scholarly databases, including Scopus and Web of Science (WOS), on January 5, 2025. The search terms "NTIS" and "National Science & Technology Information Service" were used, targeting the title, abstract, and keyword fields of publications. For KCI, both Korean and English search terms were applied to expand the search scope. The research period was set from 2008, when NTIS was established, to 2024, and only articles published in

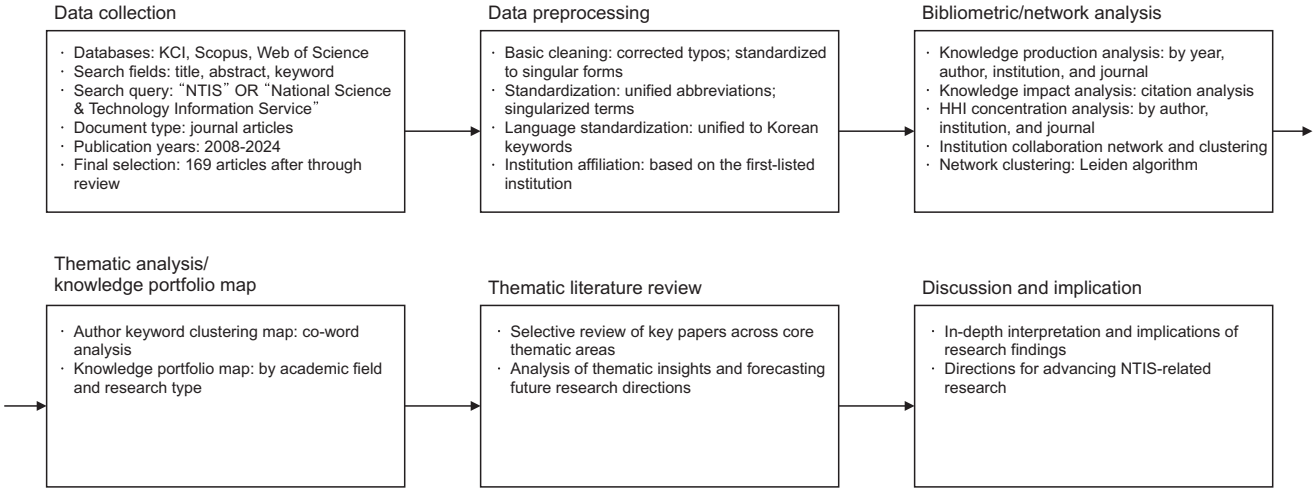



**Fig. 1.** Research framework for analysis. KCI, Korea Citation Index; NTIS, National Science & Technology Information Service; HHI, Herfindahl-Hirschman Index.

 https://doi.org/10.1633/JISTaP.2025.13.3.6

academic journals were included as analysis targets to ensure data consistency. The initially collected data were thoroughly reviewed for relevance to the research objectives, excluding inappropriate literature, and nine duplicate articles across databases were removed according to priority order (KCI → Scopus → WOS). Considering concerns regarding the credibility of certain Scopus-indexed journals and the limited number of WOS articles, the final analysis was conducted exclusively on 169 articles indexed in KCI.

The following data preprocessing procedures were performed on the final selected 169 articles. Basic cleaning work was conducted, including correcting typographical errors and standardizing plural expressions to singular forms. Author keywords written in both Korean and English were unified into Korean to ensure analytical consistency, and when both full terms and abbreviations were used, such as 'Research and Development' and 'R&D,' terminology was standardized using abbreviations. Additionally, when multiple institutional affiliations were listed for an author in a paper, the first listed institution was classified as that author's affiliation.

Based on the preprocessed data, the following analyses were performed.

First, bibliometric analysis was conducted to evaluate the productivity and scholarly impact of NTIS-related research. This method is used to identify trends and key actors in specific research areas through quantitative analysis of bibliographic information from published literature (Donthu et al., 2021; Kim & Park, 2025). In this study, publication patterns by year, performance indicators by author/institution/journal, and citation relationships were analyzed for 169 articles.

Second, concentration analysis using the Herfindahl-Hirschman Index (HHI) was conducted to measure the concentration of publications and citations among authors, institutions, and journals (Chi, 2016; Hirsch, 2005). The HHI, calculated as the sum of squares of each entity's market share ($HHI=\sum(si)^2$), is an indicator showing whether resources in a specific field are concentrated among a few entities or distributed diversely (Evren et al., 2021). Values below 1,500 were interpreted as low concentration, 1,500-2,500 as moderate concentration, and above 2,500 as high concentration (Carlton & Perloff, 2005).

Third, network analysis based on co-authorship data was conducted to understand inter-institutional collaboration structures. This analysis is effective in visually mapping relationships between research entities and revealing central actors and hidden collaboration patterns within the entire network, particularly in examining how collaboration structures evolve in specific research domains (Hyun et al., 2022). Collaboration networks were constructed targeting institutions that conducted joint research more than twice, and clustering analysis using the Leiden algorithm was applied to identify subgroups within the network (Traag et al., 2019).

Fourth, thematic analysis was performed to analyze the flow and characteristics of research topics. Co-word analysis was conducted on author keywords to visualize connections between research topics, targeting keywords that co-occurred more than twice. To identify dynamic changes in academic fields and research types, a knowledge portfolio map was developed and applied, classifying them into four categories based on occurrence frequency and growth rates: 'core' (high frequency/high growth), 'emerging' (low frequency/high growth), 'matured' (high frequency/low growth), and 'declining' (low frequency/low growth) (Lee et al., 2008; Yoon et al., 2018).

Finally, this study employed a selective literature review approach to gain a deeper understanding of NTIS-related research themes. This selective review, also known as a narrative approach or semi-systematic literature review, provides a flexible and comprehensive analysis of studies carefully selected by the authors from key research areas, organizing them thematically to understand research developments and patterns within their broader context (Snyder, 2019). This approach has been widely applied in combination with bibliometric analysis to provide both quantitative insights and qualitative understanding of research areas. To select appropriate literature for each research theme, this study consulted with five doctoral-level experts in technology policy, including two university professors and three researchers from government research institutes.

During the research process, the Bibliometrix R package was utilized for bibliographic information analysis, and Python was used for specific analytical tasks such as network visualization. Generative AI (Claude 3.7 Sonnet or Gemini 2.5 Pro) was employed solely as a tool for language translation assistance and refining textual clarity.

## 4. RESEARCH RESULTS

### 4.1. Knowledge Production, Impact, and Collaboration

NTIS-related research articles have grown steadily since it was established in 2008. As shown in Fig. 2, initially only three articles were published in 2008, increasing to 11 articles in 2017, 22 articles in 2020, and reaching a

peak of 24 articles in 2024. The compound annual growth rate (CAGR) of NTIS-related research has been 13.9%. In summary, despite its relatively limited number of publications, the field of NTIS-related research has shown long-term growth, albeit with some fluctuations.

Table 1 shows the publication and citation patterns

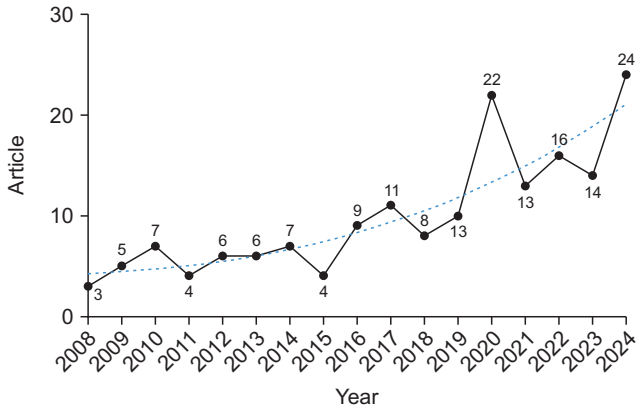



**Fig. 2.** Publication distributions by year.

of all articles. The 169 articles received a total of 635 citations, with an average of 3.8±6.7 citations per article. 104 articles (61.5% of 169) were cited at least once, and the most highly cited article accumulated 57 citations. Eighteen papers (10.7%) received 10 or more citations, accounting for 53.1% of total citations. The *h*-index, which considers both the quantity and quality of papers, was 12, indicating that 12 articles were cited at least 12 times each. The HHI was 245.3, which, according to Carlton and Perloff (2005), does not indicate a high level of concentration.

Among 391 contributing authors, 235 (60.1%) had at least one citation. The average number of authors per article was 3.0±1.8. The average *h*-index of authors was 0.7±0.7, indicating that individual scholarly impact is not particularly high. We calculated the HHI to assess concentration levels, resulting in scores of 41.5 for publications and 114.4 for citations. While citation concentration was marginally higher, both HHI scores are low, suggesting a diverse distribution of contributing authors rather than dominance by a few.

Table 2 presents the distribution of leading authors in

**Table 1.** Publication and citation metrics

| Index | Articles (A) | Cited articles (B) | Total citations | Average citations (B/A) | *h*-index | HHI |
|---|---|---|---|---|---|---|
| Value | 169 | 104 | 635 | 3.8±6.7 | 12 | 245.3 |

HHI, Herfindahl-Hirschman Index.

**Table 2.** Leading authors in publications and citations

| Publication | | Citation | | *h*-index | |
|---|---|---|---|---|---|
| Rank (n, %) | Author | Rank (n, %) | Author | Rank (n) | Author |
| 1 (11, 6.5%) | Tae-Hyeon Kim, Byeong-Hee Lee | 1 (57, 9.0%) | Sangki Jeong, Kyunghwa Kim, Seokjong Park | 1 (4) | Byeong-Hee Lee, Jungsun Lim |
| 3 (7, 4.1%) | Kwang-Nam Choi, Jungsun Lim | | | 3 (3) | Kwang-Nam Choi, Ju-Yeon Gang, Jae-Soo Kim, Sang-Kuk Kim, Tae-Hyeon Kim, Yeon-Hwa Nam, Hyo-Jung Oh, Seunghwan Oh, Dongkyu Won, Myungseok Yang |
| | | 4 (53, 8.3%) | Byeong-Hee Lee | | |
| 5 (6, 3.6%) | Myungseok Yang | 5 (50, 7.9%) | Seunghwan Oh | | |
| 6 (4, 2.4%) | Seoung Hun Bae, Hoekyung Jung, Jae-Soo Kim, Sang-Kuk Kim, Jongwon Lee, Seunghwan Oh, Kang-Ryul Shon, Sanghyuk Suh | 6 (45, 7.1%) | Pilseong Jang | | |
| | | 7 (40, 6.3%) | Jae-Soo Kim | | |
| | | 8 (29, 4.6%) | Ki-Suk Choi, Eunsil Kim, Woojung Shim | | |

 https://doi.org/10.1633/JISTaP.2025.13.3.6

NTIS-related research based on publication count, citation frequency, and *h*-index. Among the top contributors by publication, Tae-Hyeon Kim and Byeong-Hee Lee tied for first place, each with 11 articles. In citations, Sangki Jeong, Kyunghwa Kim, and Seokjong Park shared the highest citation count with 57 citations each. Regarding scholarly impact, as measured by the *h*-index, Byeong-Hee Lee and Jungsun Lim ranked highest, each with an *h*-index of 4. Three authors—Jae-Soo Kim, Byeong-Hee Lee, and Seunghwan Oh—ranked among the top across all three metrics. Several authors also demonstrated strong performance in both publication and *h*-index: Kwang-Nam Choi, Sang-Kuk Kim, Tae-Hyeon Kim, Jungsun Lim, and Myungseok Yang. Across these metrics, a significant number of the authors were affiliated with KISTI, while others were associated with institutions such as Gyeonggi Business & Science Accelerator (GBSA), Gyeongsang National University, Hoseo University, Jeonbuk National University, Korea University, Science and Technology Policy Institute (STEPI), and University of Science and Technology (UST).

In summary, a diverse group of researchers has contributed to NTIS-related studies, and a significant number of the authors who demonstrated outstanding quantitative and qualitative performance in this study were affiliated with KISTI.

NTIS-related research involved 118 institutions, with an average of 2.3±5.8 publications per institution. Among these institutions, 74.6% (88 institutions) received citations, with an average citation count of 7.8±21.8. The average *h*-index of all institutions participating in knowledge production was 1.0±1.0, which was not particularly high. The HHI for publications was 611.6 and for citations was 754.7, indicating that while absolute concentration levels were not high, citation concentration was relatively higher than publication concentration.

Table 3 shows the distribution of major institutions in NTIS-related research based on publication, citation, and *h*-index. Among the top-ranking institutions, KISTI (63 articles), UST (9), and Chungbuk National University/STEPI (8 each) ranked 1-3 in terms of publication. For citations, KISTI (206 citations), STEPI (76), and KISTEP (61) were the top three institutions, while in *h*-index, KISTI

**Table 3.** Leading institutions in publications and citations

| Publication | | Citation | | *h*-index | |
|---|---|---|---|---|---|
| Rank (n, %) | Institution | Rank (n, %) | Institution | Rank (n) | Institution |
| 1 (63, 37.3%) | KISTI | 1 (206, 32.4%) | KISTI | 1 (7) | KISTI |
| 2 (9, 5.3%) | UST | 2 (76, 12.0%) | STEPI | 2 (4) | STEPI, UST |
| 3 (8, 4.7%) | Chungbuk National Univ., STEPI | 3 (61, 9.6%) | KISTEP | | |
| | | 4 (57, 9.0%) | GBSA | 4 (3) | Chungbuk National Univ., Jeonbuk National Univ. |
| 5 (6, 3.6%) | Jeonbuk National Univ. | 5 (43, 6.8%) | UST | | |
| 6 (5, 3.0%) | Chungnam National Univ., Korea Univ., Seoul National Univ. | 6 (34, 5.4%) | Chungbuk National Univ. | 6 (2) | Chungnam National Univ., Gumi Univ., Hansung Univ., Hoseo Univ., KISTEP, Konkuk Univ., Kookmin Univ., LX-SIRI, Pukyong National Univ., Semyung Univ. |
| | | 7 (29, 4.6%) | Korea Univ. | | |
| | | 8 (23, 3.6%) | Kookmin Univ., Pukyong National Univ. | | |
| 9 (4, 2.4%) | Gyeongsang National Univ., Hanyang Univ., Hoseo Univ., Konkuk Univ., Pai Chai Univ., Sungkyunkwan Univ. | | | | |
| | | 10 (22, 3.5%) | Hansung Univ. | | |

KISTI, Korea Institute of Science and Technology Information; UST, University of Science and Technology; Univ., University; STEPI, Science and Technology Policy Institute; KISTEP, Korea Institute of Science and Technology Evaluation and Planning; GBSA, Gyeonggi Business & Science Accelerator; LX-SIRI, LX-Spatial Information Research Institute.

**Table 4.** Leading journals in publications and citations

| Publication | | Citation | | *h*-index | |
|---|---|---|---|---|---|
| Rank (n, %) | Journal | Rank (n, %) | Journal | Rank (n) | Journal |
| 1 (17, 10.1%) | *J. of the Korea Contents Association* | 1 (167, 26.3%) | *J. of Korea Technology Innovation Society* | 1 (7) | *J. of Korea Technology Innovation Society* |
| 2 (15, 8.9%) | *J. of Korea Technology Innovation Society* | 2 (51, 8.0%) | *J. of the Korea Contents Association* | 2 (4) | *J. of the Korea Contents Association* |
| 3 (5, 3.0%) | *JISTaP*, *J. of Korea Academia-Industrial Cooperation Society* | 3 (32, 5.0%) | *J. of the Korean Society for Information Management* | 3 (3) | *J. of the Korean Society for Information Management* |
| | | 4 (30, 4.7%) | *J. of the Korea Convergence Society* | 4 (2) | *J. of Industrial Innovation*, *J. of Intelligence and Information Systems*, *J. of Korea Academia-Industrial Cooperation Society*, *J. of Korean Library and Information Science Society*, *J. of Korean Society of Archives and Records Management*, *J. of Technology Innovation*, *J. of the Korea Convergence Society*, *Korean J. of Business Administration* |
| 5 (4, 2.4%) | *J. of Korean Society of Industrial and Systems Engineering*, *J. of the Korean Society for Information Management*, *J. of the Korean Society of Manufacturing Technology Engineers* | 5 (23, 3.6%) | *Innovation Studies* | | |
| | | 6 (21, 3.3%) | *J. of the Korean Society of Hazard Mitigation* | | |
| | | 7 (20, 3.1%) | *J. of Korea Academia-Industrial Cooperation Society* | | |
| 8 (3, 1.8%) | *J. of Convergence on Culture Technology*, *J. of Intelligence and Information Systems*, *J. of the Korea Institute of Information and Communication Engineering*, *J. of the Korea Society of Computer and Information*, *J. of the Korean Society for Library and Information Science*, *J. of the Korean Society of Hazard Mitigation*, *Korean J. of Business Administration* | 8 (19, 3.0%) | *Korean J. of Business Administration* | | |
| | | 9 (18, 2.8%) | *J. of Information Technology Applications & Management*, *J. of Intelligence and Information Systems* | | |

J., journal; JISTaP, Journal of Information Science Theory and Practice.

 https://doi.org/10.1633/JISTaP.2025.13.3.6

(7) and STEPI/UST (4 each) recorded the highest scores. While KISTI ranked first across all three metrics, Chungbuk National University STEPI and UST also demonstrated high performance, placing within the top ten in each category. Institutions appearing in the top rankings across two metrics include Chungnam National University, Hansung University, Hoseo University, Jeonbuk National University, KISTEP, Konkuk University, Kookmin University, Korea University, and Pukyong National University.

In summary, research in this field involves participation from diverse institutions, maintaining varied distribution among institutions in both publications and citations. While KISTI significantly leads this research across all metrics, government research institutions (KISTEP, STEPI) and various Korean universities are contributing to knowledge production and impact in this domain.

A total of 96 journals published NTIS-related research, with an average of 1.8±2.3 articles per journal. Among these, 64.6% (62 journals) received citations, with an average citation count of 6.6±18.4. The average *h*-index across all journals was 0.8±1.0, further indicating a relatively modest level of impact. The concentration of journals was higher in citations than in publications. The HHI was 277.6 for publications and 906.7 for citations.

Table 4 presents the distribution of leading journals in NTIS-related research based on publication, citation, and *h*-index. The *Journal of the Korea Contents Association*, and *Journal of Korea Technology Innovation Society* ranked first and second across all metrics. The *Journal of the Korea Contents Association* ranked first in article publications (17), second in citations (51), and second in *h*-index (4). The *Journal of Korea Technology Innovation Society* ranked second in article publications (15), first in citations (167), and first in *h*-index (7).

Following these leading journals, the *Journal of Intelligence and Information Systems*, *Journal of Korea Academia-Industrial Cooperation Society*, *Journal of the Korean Society for Information Management*, and *Korean Journal of Business Administration* also ranked highly across the metrics. Journals appearing in the top rankings across two metrics include the *Journal of the Korea Convergence Society*, *Journal of the Korean Society for Hazard Mitigation*, and *Journal of the Korean Society for Library and Information Science*.

In summary, a wide range of journals contributed to the dissemination, with citation concentration being relatively higher than publication. Journals related to contents, technology innovation, library and information science, convergence, and business administration led the research in this field.

This study analyzed institutional collaboration patterns and found that, among 169 articles, 81 (47.9%) exhibited collaboration between institutions. Fig. 3 presents a network constructed based on institutions that collaborated with each other more than twice. A clustering analysis identified a total of three collaborative groups.

■ Cluster 1 (red): This is the most extensive cluster, with KISTI occupying a central position in the network and demonstrating the highest degree centrality. KISTI maintains close collaborations with nine institutions: Chungnam National University, Gumi University, Hoseo

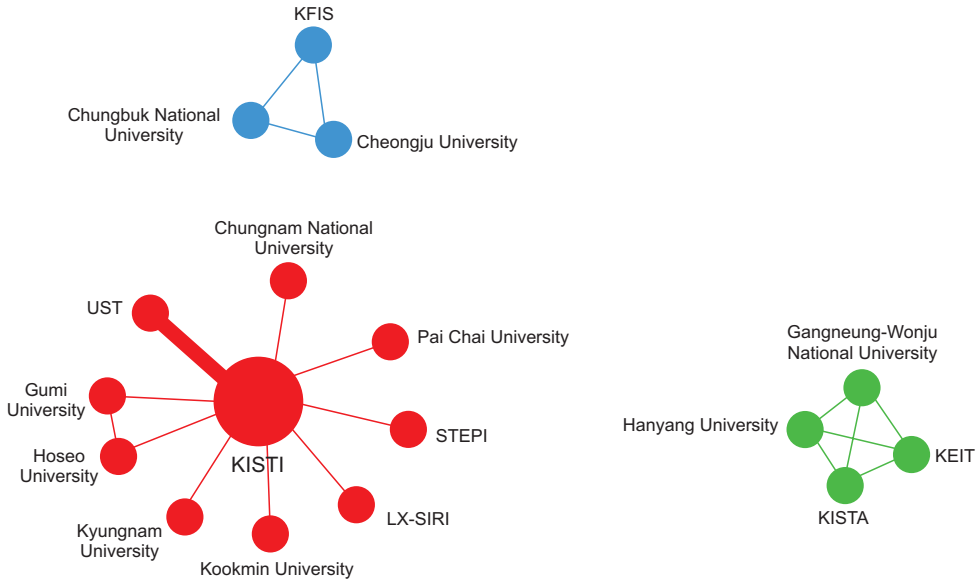



**Fig. 3.** Institutional collaboration clustering map. KISTI, Korea Institute of Science and Technology Information; UST, University of Science and Technology; STEPI, Science and Technology Policy Institute; LX-SIRI, LX-Spatial Information Research Institute; KISTA, Korea Intellectual Property Strategy Agency; KEIT, Korea Evaluation Institute of Industrial Technology; KFIS, Korea Finance Information Service.

University, Kookmin University, Kyungnam University, LX-Spatial Information Research Institute (LX-SIRI), Pai Chai University, STEPI, and UST.

■ Cluster 2 (green): This cluster consists of Gangneung-Wonju National University, Hanyang University, Korea Evaluation Institute of Industrial Technology (KEIT), and Korea Intellectual Property Strategy Agency (KISTA). It reflects collaborations primarily between universities and government R&D-related institutions.

■ Cluster 3 (blue): This is a smaller collaborative group composed of Cheongju University, Chungbuk National University, and Korea Finance Information Service (KFIS).

In summary, nearly half of the NTIS-related studies involve institutional collaboration, with KISTI playing a central role in fostering cooperation between universities and government-funded research institutes.

### 4.2. Thematic Analysis and Literature Review

For the thematic analysis, we examined the distribution and growth of NTIS-related research across academic fields based on journal classifications provided by the KCI. To ensure consistency in disciplinary categorization, we adopted the sub-discipline level (e.g., business administration or economics within the broader field of social sciences). Considering the distribution of published papers over time, we divided the study period into two phases for comparative analysis: Period 1 (2008-2019, 80 articles) and Period 2 (2020-2024, 89 articles).

Our analysis identified NTIS-related research across 36 academic fields, expanding from 19 in the earlier period to 31 in the later period. For the 12 major fields (≥2% of total articles), we constructed a knowledge portfolio map based on publication volume and growth rate (Period 2 vs. Period 1), as shown in Fig. 4. In terms of volume, interdisciplinary research leads with 27 articles, followed by technology policy and library and information science (19 each), and business administration (16). In terms of growth rate, general engineering (an outlier) shows the highest increase, followed by business administration and mechanical engineering (both 200.0%), S&T studies (100.0%), and computer science (75.0%). Academic fields were mapped onto a 2×2 matrix: High Growth–High Volume (interdisciplinary research, business administration); High Growth–Low Volume (general engineering, mechanical engineering, S&T studies, computer science, electronic information communication engineering); Low Growth–High Volume (technology policy, library and information science); and Low Growth–Low Volume (industrial engineering, economics, miscellaneous engineering).

In summary, NTIS-related research has significantly expanded in scope over time. While the relative share of technology policy and library and information science is declining, research is increasingly branching into areas such as engineering, computer science, and S&T studies. Interdisciplinary studies and business administration continue to maintain high publication volumes and growth rates, establishing themselves as core research areas.

This study conducted a literature review of major NTIS-related papers in interdisciplinary and business administration fields. While both areas utilized NTIS data to analyze research trends and propose service improvements, interdisciplinary research focused on R&D systems and collaboration strategies, whereas business administration research emphasized the analysis of government

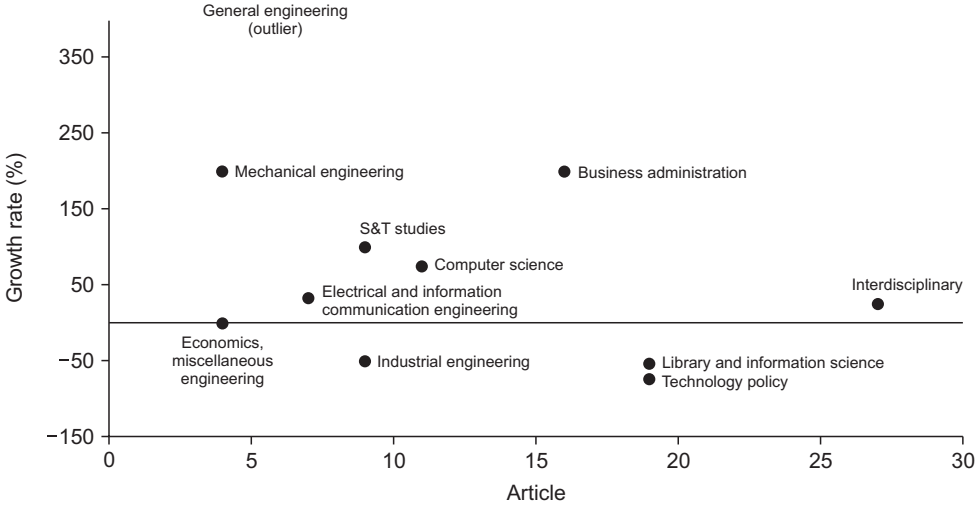



**Fig. 4.** Knowledge portfolio map by academic field. S&T, Science and Technology.

 https://doi.org/10.1633/JISTaP.2025.13.3.6

R&D investment performance.

An in-depth literature review of interdisciplinary papers revealed three main categories of research topics: (i) Enhancement of NTIS services through improvements in quality, accessibility, and user experience. Shon and Lim (2009) suggested methods to improve R&D data quality, Yu et al. (2020) examined how curation services can improve R&D information accessibility, and Lee et al. (2022) proposed an interactive search system using AI techniques to boost search efficiency. (ii) Analysis of research characteristics and trends in specific fields using NTIS data. Choi et al. (2021), for example, explored the knowledge structure of fuel cell electric vehicles within hydrogen economy projects. (iii) Investigation of R&D collaboration and technology transfer. Kim et al. (2016) studied factors influencing technology transfer in government research institutions, while Min et al. (2024) categorized R&D collaboration models between Korea and developing countries based on innovation performance.

In the realm of business administration, NTIS-related research is also grouped into three main categories: (i) Investigation of R&D characteristics through data and text analysis. Yoo et al. (2020) analyzed national convergence R&D characteristics using NTIS data, and Choi and Lee (2020) identified structural features of regional technology development R&D. (ii) Evaluation of NTIS service quality and usability improvement. Heo and Kim (2020) proposed enhancements to NTIS through an online information service quality evaluation model, and Choi and Son (2022) identified factors influencing the continued use of the NTIS platform based on user surveys. (iii) Analysis of government R&D investment performance and its influencing factors. Ro et al. (2018) empirically assessed government R&D investment performance in industrial technology fields, including patents, technology fee income, and commercialization, while Oh and Park (2020) systematically evaluated the impact of government R&D support for medium-sized enterprises, focusing on asset growth and R&D investment expansion.

Based on input from five doctoral-level experts in technology policy (two university professors and three government research institute researchers), we classified NTIS-related research into three categories: (1) studies analyzing NTIS itself, including its introduction, characteristics, and effects; (2) studies proposing improvements to NTIS; and (3) studies utilizing NTIS or its data. Our analysis revealed that NTIS utilization research dominated the literature (114 articles, 67.5%), followed by NTIS improvement research (47 articles, 27.8%), and NTIS itself research (8 articles, 4.7%). Comparing Period 1 and Period 2, utilization-focused research increased significantly by 42.6%, while NTIS improvement-focused research decreased by 25.9% and research on NTIS itself declined by 66.7%.

To provide more granular insights, we further refined our classification based on expert opinions: NTIS itself was subdivided into NTIS characteristics and NTIS effects; NTIS improvement was categorized into process/system improvement, user usability improvement, data quality improvement, and legal/institutional improvement; and NTIS utilization was classified into R&D trend identification, R&D performance analysis, and R&D applications. We also constructed a knowledge portfolio map (Fig. 5) to visualize these research categories. R&D trend identification (65 articles; 142.1% growth) falls into the High Growth–High Volume quadrant. The High Growth–Low Volume quadrant includes process/system

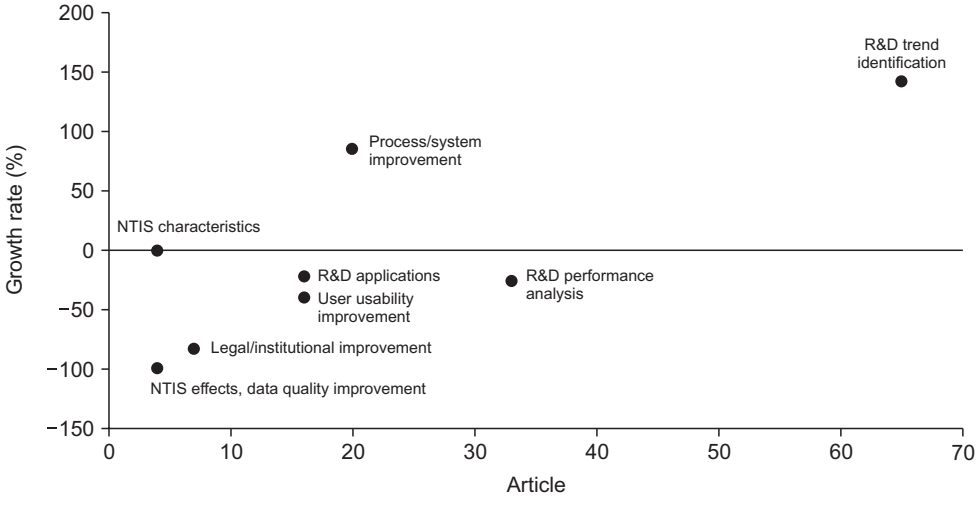



**Fig. 5.** Knowledge portfolio map by research type. NTIS, National Science & Technology Information Service; R&D, Research and Development.

improvement (20 articles; 85.7% growth). R&D performance analysis (33 articles; -26.3%) is positioned in the Low Growth–High Volume quadrant. The Low Growth–Low Volume quadrant encompasses legal/institutional improvement (-83.3%), user usability improvement (-40.0%), NTIS effects/data quality improvement (-100.0%), and NTIS characteristics (0.0%).

In summary, NTIS-related research has been growing, particularly in studies that utilize NTIS data, with R&D trend identification emerging as a core area. In contrast, research interest has declined in several areas, including legal and institutional improvement and data quality analysis.

This study conducted a literature review on national R&D trend analysis using NTIS data, which has emerged as the most prominent research area. The analysis revealed that research trend analysis is being conducted centered around 10 fields: (i) Health·Medicine·Bio (e.g., Korean medicine, healthcare, women's health research); (ii) Defense·Disaster Safety·Security (e.g., promising defense technologies, disaster safety response systems); (iii) Energy·Environment (e.g., hydrogen economy, carbon-neutral fuels); (iv) Materials·Components·Equipment (e.g., display materials, membrane materials, optical communication components); (v) ICT·AI·Smart Technologies (e.g., AI, autonomous driving, cloud security); (vi) Content·Cultural Technologies (e.g., immersive media, metaverse, webtoons); (vii) Agriculture·Food (e.g., healing agriculture, food safety, agricultural technology); (viii) Machinery·Equipment (e.g., construction machinery); (ix) Interdisciplinary Fields; and (x) Other Fields (e.g., regional S&T). Among these, research trend analyses in the ICT·AI·Smart Technologies, Health·Medicine·Bio, and Defense·Disaster Safety·Security fields were found to be conducted most actively.

To analyze these research trends, various techniques such as (i) text mining (e.g., graphene technology, agricultural R&D), (ii) topic modeling (e.g., metaverse, AI), (iii) network analysis (e.g., bio-healthcare, disaster safety projects), (iv) patent and citation analysis (e.g., display, all-solid-state batteries), and (v) metadata analysis (e.g., energy technology) have been applied either individually or in combination. The application of topic modeling is showing a significantly increasing trend, and this is projected to continue in the future.

Table 5 shows the changes in top keywords over time based on author keyword frequency. Across all periods, R&D (23 articles) and R&D project (17) were the most

**Table 5.** Leading keywords by period

| Overall (2008-2024) | | Period 1 (2008-2019) | | Period 2 (2020-2024) | |
|---|---|---|---|---|---|
| Rank (n, %) | Keyword | Rank (n, %) | Keyword | Rank (n, %) | Keyword |
| 1 (23, 13.6%) | R&D | 1 (8, 10.0%) | R&D program | 1 (16, 18.0%) | R&D |
| 2 (17, 10.1%) | R&D project | 2 (7, 8.8%) | R&D | 2 (13, 14.6%) | R&D project |
| 3 (11, 6.5%) | Text mining | 3 (5, 6.3%) | R&D performance | 3 (9, 10.1%) | Research trend, Text mining |
| 4 (10, 5.9%) | Research trend | 4 (4, 5.0%) | R&D project, R&D information, R&D investment | | |
| 5 (9, 5.3%) | Network analysis | | | 5 (8, 9.0%) | Network analysis |
| 6 (8, 4.7%) | R&D program | | | 6 (7, 7.9%) | AI |
| 7 (7, 4.1%) | AI, R&D investment, Topic modeling | 7 (3, 3.8%) | Data cleansing, DEA, Database integration, GRI, Information quality, IPR, Ontology, Technology transfer | 7 (6, 6.7%) | Topic modeling |
| | | | | 8 (4, 4.5%) | Deep learning |
| | | | | 9 (3, 3.4%) | Big data, Cluster analysis, LDA, Machine learning, R&D investment, R&D support, S&T standard classification |
| 10 (6, 3.6%) | R&D performance | | | | |

R&D, Research and Development; AI, Artificial Intelligence; DEA, Data Envelopment Analysis; GRI, Government Research Institute; IPR, Intellectual Property Rights; LDA, Latent Dirichlet Allocation; S&T, Science and Technology.

 https://doi.org/10.1633/JISTaP.2025.13.3.6

frequently appearing keywords. Other frequently observed topics included research methodology terms such as text mining (11), network analysis (9), and topic modeling (7); and R&D lifecycle implementation terms like R&D program (8), R&D investment (7), and R&D performance (6). Furthermore, AI (7) was also a high-ranking keyword, with the literature analysis results showing this comprised papers on AI research trends (5) and the use of AI as an analysis method (2).

A comparison of these top keywords by period reveals shifts in research focus and methodology application. In Period 1, R&D lifecycle terms such as R&D program/project and R&D information/investment/performance were the most prominent, followed by terms related to NTIS system development, including database integration, data cleaning, ontology, and information quality. In Period 2, which covers the most recent five years, the focus shifted slightly, with terms related to the R&D lifecycle—such as R&D, R&D projects, and R&D investment/support—remaining prominent. At the same time, new research trend areas emerged among the top keywords, alongside data- and text-analysis methods such as big data, text mining, network analysis, topic modeling, cluster analysis, and Latent Dirichlet Allocation (LDA). AI-related terms, including machine learning and deep learning, have also gained increased attention as key research topics.

In summary, keywords related to the R&D lifecycle have remained central throughout all periods as core NTIS terms. Meanwhile, the prominence of system development issues, which were important in earlier stages, has decreased. However, various analysis techniques for NTIS, and the application of advanced AI techniques, are emerging as key keywords.

Fig. 6 visualizes the results of network and clustering analysis based on author keywords that co-occurred more than twice in all articles. Through this analysis, nine clusters were identified.

■ Cluster 1 (red): The largest cluster by keyword volume, links core R&D concepts (e.g., R&D information/project) with relevant data types (e.g., big data, patent data), advanced analytical methodologies (e.g., text mining, network analysis, topic modeling, LDA), and research goals like trend analysis, demonstrating their interconnectedness. Characteristically, studies within this cluster typically involve collecting datasets (papers, research projects, patent data) from sources such as NTIS, focusing on specific fields (e.g., autonomous driving, cloud security, immersive media), analyzing these data using various techniques, deriving research topics, and suggesting future research directions based on the findings. For instance, Woo and Lee (2020) applied LDA-based topic modeling to a Korean national R&D project dataset in the ICT sector to analyze key research topics related to intelligent information technology. Kim et al. (2023) compared research trends in the ICT field of South and North Korea by applying data mining methods. In another study representative of this cluster, Lee and Lee (2022) collected national R&D project data in the metaverse field to analyze temporal shifts in research themes.

■ Cluster 2 (blue): This is the second largest cluster. It primarily includes research focused on analyzing research

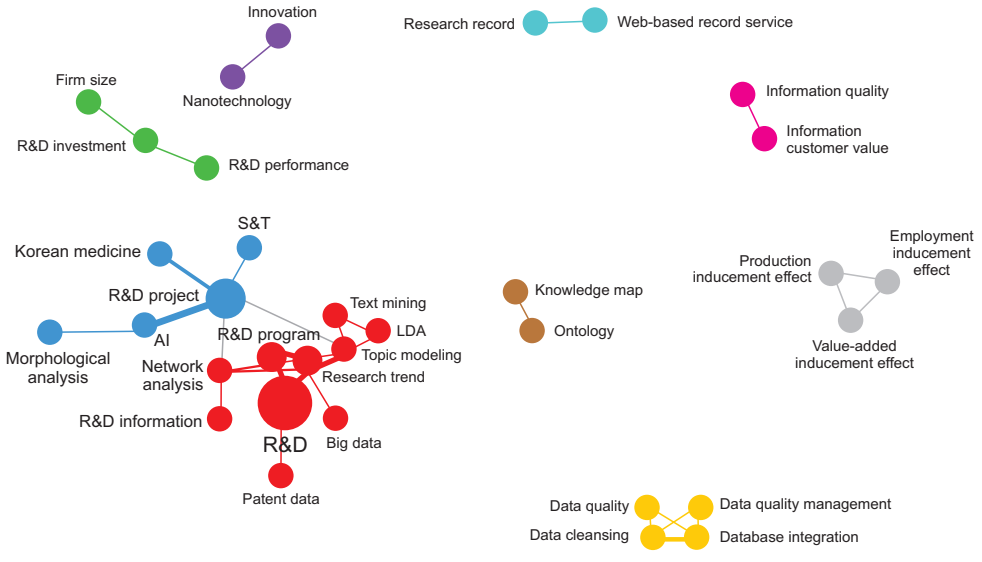



**Fig. 6.** Keyword clustering map. R&D, Research and Development; AI, Artificial Intelligence; S&T, Science and Technology; LDA, Latent Dirichlet Allocation.

trends in specific fields centered on R&D projects, as well as improving the NTIS search system using AI technologies or morphological analysis. Kwon et al. (2017) notably analyzed the public investment status in Korean medicine by utilizing NTIS data, aiming to support future strategic planning for Korean medicine R&D. Lee et al. (2020) proposed the design and implementation of an interactive search system that analyzes users' input words or sentences morphologically and leverages the KoBERT model to better identify user search intent and provide more accurate search results within the NTIS system.

■ Cluster 3 (green): Comprising keywords such as R&D investment, R&D performance, and firm size, this cluster explores the relationship between R&D investment and outcomes, taking organizational scale into account. This cluster consists of literature that measures investment performance or efficiency at the national level, specific industries, or specific technology fields based on paper analysis or patent analysis using NTIS data. For example, Kim (2017) utilized NTIS data to evaluate the performance of government investment in national IT R&D from the perspectives of papers, patents, and commercialization, and proposed ways to improve investment efficiency and strategic budget allocation.

■ Cluster 4 (purple): Including keywords like nanotechnology and innovation, this cluster represents innovation studies within specific high-tech fields. This cluster primarily consists of papers that analyze the innovation characteristics of Korean government R&D investments using NTIS data, with a focus on the nanotechnology field. Lim et al. (2019) quantified the innovation characteristics of R&D investment in the nanotechnology field using data envelopment analysis (DEA). Lim and Kim (2020) identified innovation patterns by government departments using NTIS data on government R&D investment outcomes (papers, patents, revenue, employment, etc.) in the nanotechnology.

■ Cluster 5 (orange): With keywords such as database integration, data cleansing, data quality, and data quality management, this cluster addresses the dimensions of NTIS database management and data quality management. This cluster consists mostly of literature that suggested directions for database integration and data quality improvement during the initial construction phase of NTIS. After the mid-2010s, no further research has been conducted. Shon and Lim (2009) examined the integrated database construction and linkage methods for NTIS's human resources, projects, and performance information. Shin et al. (2011) used NTIS as a case study to consider factors affecting data quality, such as quality policies, organization, business processes, and business rules.

■ Cluster 6 (brown): Featuring an ontology and knowledge map, this cluster focuses on methodologies for knowledge organization and visualization. The literature included in this cluster was conducted in the early to mid-2010s. These studies aimed to systematically structure national R&D information stored in NTIS using ontology and intelligent analysis techniques and to build knowledge map services, thereby enhancing the usability of NTIS data. Jeong (2015) developed a knowledge map service visualizing relationships between researchers, keywords, institutions, and journals by constructing a lightweight ontology and using topic modeling with national R&D data.

■ Cluster 7 (pink): Including information quality and information customer value, this cluster examines factors related to customer satisfaction and value in the context of NTIS services. This literature analyzed quality and customer value factors that influence customer satisfaction and user performance among users of the NTIS system, aiming to improve information service quality since the initial service provision phase. Suh et al. (2014) found that information service quality positively affects customer value and information source recognition, which influences usage performance, recommending service strategies through user analysis to enhance NTIS's value.

■ Cluster 8 (gray): With keywords such as value-added inducement effect, production inducement effect, and employment inducement effect, this cluster analyzes the economic impacts resulting from NTIS. Park (2018) conducted an inter-industry analysis using the Bank of Korea's input-output tables to analyze the economic ripple effects of NTIS. The results revealed substantial production-inducing effects, value-added effects, and employment-inducing effects compared to the government R&D budget investment, demonstrating the system's socioeconomic contributions.

■ Cluster 9 (mint): Containing a research record and web-based record information service, this cluster addresses research documentation and the delivery of online information services. The literature in this cluster primarily consists of studies from the late 2010s that suggest directions for activating web-based research record information services centered on NTIS. Nam et al. (2017) developed evaluation indicators reflecting research record characteristics through service quality indicator analysis, applying these to NTIS as a case study to identify service problems and propose improvement measures for web-

 https://doi.org/10.1633/JISTaP.2025.13.3.6

based research record information services.

In summary, the core themes of NTIS-related research converge on: data and system management, analysis of R&D investment and performance, enhancement of user value and analysis of economic ripple effects, and application of diverse analytical methodologies. These themes illuminate the multifaceted role of NTIS in promoting national R&D and innovation, alongside the development process of the NTIS system itself. Furthermore, synthesizing the thematic analysis of this study reveals that research related to Korea's NTIS holds significant meaning beyond a simple exploration of knowledge structures. It serves as an important case from information science, sociological, and R&D policy perspectives, concretely demonstrating how national investment in R&D information infrastructure influences the generation, direction, and diversification of academic knowledge. In other words, this suggests that NTIS acts as a driver promoting the propagation and differentiation of various research topics within the academic ecosystem, moving beyond the mere role of an R&D information service provider.

## 5. CONCLUSIONS

### 5.1. Discussion

This study reveals that NTIS-related research has fundamentally shifted from system-oriented development toward broader scholarly application, demonstrating that NTIS is increasingly functioning as academic infrastructure within South Korea's research landscape rather than merely serving as an information portal.

This comprehensive bibliometric analysis of NTIS-related academic literature in South Korea from 2008 to 2024 goes beyond simply quantifying a research area to demonstrate how national R&D information infrastructure dynamically co-evolves with its surrounding academic and policy ecosystems. Our findings provide insights into how such platforms both shape and are shaped by broader currents of governance, knowledge production, and technology adoption.

This study reveals that despite steady growth, the limited research volume coincides with pronounced institutional concentration around KISTI, the system development and operation entity. HHI analysis shows relatively low concentration in publications but higher concentration in citations, confirming that influential research clusters among specific institutions. This pattern reflects important dynamics: While KISTI's central role has clearly driven knowledge production and established it as a hub in inter-institutional collaboration networks, it simultaneously fosters knowledge path dependency and tends to channel academic participation in specific directions. For NTIS to develop as a truly national research platform, gradual evolution is needed from the current supplier-centered model led by core operating institutions toward an expanded demand-driven ecosystem that actively embraces diverse perspectives from universities and other research disciplines.

Furthermore, changes in research themes reflect a significant transformation in South Korea's national R&D paradigm. The decline in research focused on system development and data management, coupled with a surge in trend analysis and performance evaluation studies utilizing NTIS data, demonstrates that NTIS has matured beyond a simple information repository. Academic fields are also expanding significantly from traditional technology policy and library and information science to interdisciplinary research, business administration, and engineering domains. Nine thematic clusters identified through keyword cluster analysis concretely illustrate this diversification, particularly confirming that while R&D lifecycle-related terms maintain their central role across all periods, advanced analytical techniques such as big data, text mining, and AI are emerging as new research drivers. This transformation reflects the institutionalization of data-driven governance in S&T policy, positioning NTIS beyond simple service provision as a core tool for national innovation strategy. However, such changes bring new challenges. The relative decline of foundational disciplines such as library and information science and R&D policy, which traditionally provide theoretical foundations, may potentially weaken the long-term intellectual base for such applied research.

Finally, this analysis reveals important knowledge gaps that represent both challenges and opportunities for future research. The overwhelming domestic concentration of research leads to insufficient international comparative studies, and the pattern of publications centered on domestic journals limits the global dissemination of research outcomes. While the application of advanced analytical methods such as AI and text mining is increasing, in-depth research on national data governance, the specific impact of NTIS data on policy decisions, and the role of information science in shaping technology policy remains insufficient. Addressing these gaps is crucial. Through this, NTIS-related research can advance beyond its current focus on data utilization to develop as a core research

area exploring the complex interactions among national R&D infrastructure investment, knowledge production dynamics, and policy innovation.

### 5.2. Practical Implications for Enhancing NTIS-Related Research

Our analysis reveals limited research volume (169 articles over 16 years) and pronounced institutional concentration around KISTI (37.3% of total publications), underscoring the urgent need to broaden participation beyond the current supplier-centered model. As practical measures, research support programs such as thesis competitions utilizing NTIS data or workshops for early-career researchers should be expanded, and regular NTIS-related sessions should be operated in collaboration with relevant academic societies to form a focal point for the researcher community. Additionally, efforts to plan and support interdisciplinary convergence research projects should be undertaken to accelerate the positive trend of research expansion into various fields such as business administration and engineering. In particular, while maintaining the current KISTI-centered research network, it is necessary to gradually expand participation from universities and private research institutions, promoting a transition from a supplier-centered to a demand-centered research ecosystem.

Thematic analysis demonstrates a concerning narrowing of focus toward data utilization, with such studies increasing by 42.6% while foundational system improvement research declined by 25.9%, pointing to the need for deeper theoretical engagement to strengthen the intellectual foundation. This trend shows that while NTIS data utilization research has increased quantitatively, it is concentrated on specific topics, and research addressing theoretical foundations has actually diminished. Therefore, it is necessary to establish and guide research agendas that can add depth to research from a long-term perspective, beyond short-term trend analysis. Specifically, research in technology policy and library and information science fields should be reactivated to strengthen the theoretical foundation of NTIS research, and exploration of challenging topics should be encouraged, including the knowledge gaps identified in this study: national data governance frameworks, evaluation of NTIS's policy decision influence, and comparative studies with similar overseas systems. Additionally, empirical analysis of NTIS data's role in evidence-based policy formation and its actual impact on the national innovation system, which are currently insufficient, should also be expanded.

The findings reveal overwhelming domestic concentration, with NTIS studies concentrated primarily within South Korean academic institutions and journals, and a notable absence of international comparative studies, indicating significant opportunities for global integration. To address this limitation, it is necessary to activate international comparative research and expand cooperation with overseas researchers. Through publishing NTIS special issues in international academic journals such as the *Journal of Information Science Theory and Practice* (JISTaP) or operating NTIS sessions at international academic conferences, South Korea's experience with national R&D information systems can be shared with the international academic community. Such efforts will enable comparative analysis with similar systems in other countries and provide opportunities to objectively evaluate NTIS's relative strengths and weaknesses.

The increasing use of advanced analytical methods, with AI, big data, and topic modeling emerging as dominant keywords, points to the need for NTIS to expand its data openness and technical capabilities to support sophisticated analytical approaches. Since the development of the research ecosystem directly depends on the platform's openness and functionality, system innovation should be pursued for NTIS to leap beyond a domestic information service to become a global research infrastructure. To this end, compatibility with international standards should be secured by applying global identifiers and providing multilingual APIs to enhance accessibility for overseas researchers, and data sharing and reuse policies based on Open Science principles should be strengthened. Additionally, AI-based intelligent analysis functions should be introduced to seek evolution into a 'knowledge discovery platform' where researchers can discover new insights through data, beyond simple information retrieval. Development of customized services reflecting the requirements of various stakeholders including researchers, policymakers, industry, and the general public should be undertaken, along with continuous improvement through regular user experience evaluations.

### 5.3. Contributions and Limitations

This study makes a major contribution by expanding the academic scope of existing quantitative bibliometric analysis, which had been limited to specific technologies or academic fields, to a national R&D information system (NTIS) itself. Through this approach, the study systematically identifies the unique characteristics and developmental process of the NTIS system, as well as its associated re-

 https://doi.org/10.1633/JISTaP.2025.13.3.6

search landscape and knowledge structure. Furthermore, it contributes to laying a foundation for future research on similar R&D information systems or platforms by integrating bibliometric, network, thematic analysis, and selective review methods. This study also makes a significant academic contribution by empirically demonstrating that NTIS has evolved beyond a simple information provider into a core academic infrastructure for data-driven research, and by clarifying the process through which a national information infrastructure influences the direction and form of academic knowledge production. Based on the diversification of research topics, the expansion of academic fields, and the sophistication of analytical methodologies, the study identifies that NTIS functions not merely as an information repository but as a driving force that promotes and differentiates research across various disciplines. This study further contributes from a policy and practical standpoint by diagnosing the inherent limitations of an academic field centered on a specific information system and proposing practical directions for overcoming them. By addressing problems such as a quantitative lack of research, knowledge production centered on specific institutions, and a research scope limited to domestic boundaries, the study proposes concrete strategies, including the vitalization of research communities, diversification of research areas, global expansion through international collaborative research, and strengthening of inter-institutional cooperation. Through these proposals, the study provides important policy implications for the sustainable development of national R&D information infrastructures, including NTIS.

While this study provides comprehensive insights into NTIS-related research patterns, several methodological constraints and scope limitations inevitably influence our findings and their broader applicability. The most fundamental limitation stems from our exclusive analysis of Korean literature, which, while providing deep understanding of the domestic research landscape, constrains our ability to position NTIS within the broader international context of national R&D information systems. This domestic-focused perspective limits our understanding of whether the observed patterns represent universal tendencies in national R&D information ecosystems or reflect Korea-specific developmental trajectories. Furthermore, the methodology applied in this study was effective in systematically analyzing research trends and collaboration patterns, but has limitations in capturing the substantial impact of NTIS on the academic ecosystem—namely, the deeper mechanisms such as actual policy decision processes or new R&D project planning. While our bibliometric analysis successfully identified macroscopic knowledge structure patterns, it could not elucidate the complex mechanisms through which NTIS data has specifically influenced R&D policy-making processes, national innovation systems, and academic communities. This gap between observable publication patterns and actual impact on the academic-policy ecosystem represents a significant constraint in fully assessing NTIS's role as policy infrastructure. Moreover, the selective literature review approach employed in this study, while offering the advantage of providing qualitative insights into major research themes, operates within the constraints of subjective judgment rather than systematic synthesis. Although this approach enabled us to capture in-depth analysis of research focus areas, it lacks the methodological rigor that systematic literature review methods can provide.

These limitations collectively point toward promising directions for future research. International comparative studies examining multiple national R&D information systems, including the United States RePORTER and Japan's KAKEN, would provide crucial context not only for interpreting the Korean experience but also for understanding within a global framework. Mixed-methodological approaches combining bibliometric analysis with surveys and case studies could bridge the gap between observable publication patterns and actual impact on the academic ecosystem. In particular, the application of systematic literature review methodologies would enable more rigorous and structured synthesis of research methodologies, research topics, and results, supporting clearer judgments about the relative importance of different research streams.

## CONFLICTS OF INTEREST

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

## ACKNOWLEDGEMENTS

This research was supported by the Daejeon University Research Grants (2022).

## REFERENCES

Carlton, D. W., & Perloff, J. M. (2005). *Modern industrial organization* (4th ed.). Pearson/Addison Wesley.

Chi, P. S. (2016). Differing disciplinary citation concentra-

tion patterns of book and journal literature? *Journal of Informetrics*, 10(3), 814-829. https://doi.org/10.1016/j.joi.2016.05.005

Choi, E. B., & Son, D. H. (2022). Determinants of continuous intention-to-use on NTIS: Perspectives of UTAUT and TTF Model. *The Journal of Information Systems*, 31(2), 197-216. https://doi.org/10.5859/KAIS.2022.31.2.197

Choi, J. W., Lee, J. Y., Lee, B. H., & Kim, T. H. (2021). Exploring the knowledge structure of fuel cell electric vehicle in national R&D projects for the hydrogen economy. *The Journal of the Korea Contents Association*, 21(6), 306-317. https://doi.org/10.5392/JKCA.2021.21.06.306

Choi, W. S., & Lee, J. H. (2020). A study on the characteristics of R&D for technology development in Chungbuk area : Focused on text analysis and network analysis. *Journal of Industrial Innovation*, 36(2), 1-19. https://doi.org/10.22793/indinn.2020.36.2.001

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285-296. https://doi.org/10.1016/j.jbusres.2021.04.070

European Commission (EC). (n.d.). *Community Research and Development Information Service (CORDIS)*. https://cordis.europa.eu

Evren, A., Tuna, E., Ustaoglu, E., & Sahin, B. (2021). Some dominance indices to determine market concentration. *Journal of Applied Statistics*, 48(13-15), 2755-2775. https://doi.org/10.1080/02664763.2021.1963421

Heo, J. Y., & Kim, S. K. (2020). An inductive study on the service quality evaluation model of online information service : Focusing on the national online information service NTIS and NDSL. *Journal of Information Technology Services*, 19(1), 71-87. https://doi.org/10.9716/KITS.2020.19.1.071

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569-16572. https://doi.org/10.1073/pnas.0507655102

Huh, T. S., Choi, K. S., Jhun, S. J., & Park, M. H. (2009). An economic feasibility analysis on the establishment of dedicated management system for the national R&D reports. *Journal of the Society of Korea Industrial and Systems Engineering*, 32(2), 45-56. https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART001353679

Hyun, M. H., Lee, H. J., Lim, S. J., & Lee, K. D. (2022). Changes in the structure of collaboration network in artificial intelligence by national R&D stage. *Journal of Information Science Theory and Practice*, 10(Special), 12-24. https://doi.org/10.1633/JISTaP.2022.10.S.2

Japan Society for the Promotion of Science (JSPS). (n.d.). *Grants-in-aid for scientific research*. https://www.jsps.go.jp/english/e-grants/

Jeong, H. (2015). A study on ontology and topic modeling-based multi-dimensional knowledge map services. *Journal of Intelligence and Information Systems*, 21(4), 79-92. https://doi.org/10.13088/jiis.2015.21.4.079

Kim, H. K. (2017). Performance evaluation of national R&D investment in information technology areas. *Journal of The Korea Society of Computer and Information*, 22(2), 119-125. https://doi.org/10.9708/jksci.2017.22.02.119

Kim, J., Choi, H., & Mo, J. (2023). A comparative analysis of research trends in the information and communication technology field of South and North Korea using data mining. *Journal of Information Science Theory and Practice*, 11(1), 14-30. https://doi.org/10.1633/JISTaP.2023.11.1.2

Kim, M. K., & Park, M. J. (2025). Exploring and comparing the knowledge structures, collaborations and key themes of electric vehicle research in the environment domain. *International Journal of Sustainable Transportation*, 19(3), 195-210. https://doi.org/10.1080/15568318.2024.2448002

Kim, S., Yeon, S., Kang, I., & Lee, B. H. (2016). Analysis of the determinants influencing technology transfer in government funded research institutes: Focusing on features of national R&D projects. *The Journal of the Korea Contents Association*, 16(8), 624-639. https://doi.org/10.5392/JKCA.2016.16.08.624

Korea Institute of Science and Technology Evaluation and Planning (KISTEP). (2013). *A study on discovering themes related to policy of R&D performance and publishing issue papers* (TRKO201300031910). KISTEP, Eumseong, Korea. https://scienceon.kisti.re.kr/srch/selectPORSrchReport.do?cn=TRKO201300031910#;

Korea Institute of Science and Technology Evaluation and Planning (KISTEP). (2025). *A study on the construction of national R&D planning and management supporting service* (TRKO202500001862). KISTEP, Eumseong, Korea. https://scienceon.kisti.re.kr/srch/selectPORSrchReport.do?cn=TRKO202500001862#;

Korea Institute of Science and Technology Information (KISTI). (2019). *NTIS user guide*. KISTI.

Korea Institute of Science and Technology Information (KISTI). (n.d.). *NTIS introduction*. https://www.ntis.go.kr/ThAbout.do

Kwak, S. J., Park, J. H., Kang, B. S., Kim, D. G., & Oh, S. (2022). A study on research support service of the national R&D full cycle. *Journal of the Korean Society for Library and Information Science*, 56(1), 405-423. https://doi.org/10.4275/KSLIS.2022.56.1.405

Kwon, S. H., Kim, D., Ahn, M. Y., & Lim, B. (2017). The recent trend of R&D investment in Korean medicine by research steps and fields. *Journal of Society of Preventive Korean Medicine*, 21(2), 69-78. https://doi.org/10.25153/spkom.2017.21.2.008

Lee, D. Y., & Kim, K. H. (2021). Information analysis framework for supporting evidence-based research and development policy: Practical considerations for rationality in the policy process. *Informatization Policy*, 28(1), 77-93. https://doi.org/10.22693/NIAIP.2021.28.1.077

Lee, J., & Lee, S. (2022). An analysis of national R&D trends in the metaverse field using topic modeling. *Smart Media Journal*, 11(8), 9-20. https://doi.org/10.30693/SMJ.2022.11.8.9

Lee, J. W., Jo, W. S., & Kim, T. H. (2022). Design and implementation of an interactive search system based on KoBERT. *Journal of Knowledge Information Technology and Systems (JKITS)*, 17(5), 1081-1088. https://doi.org/10.34163/jkits.2022.17.5.028

Lee, J. W., Kim, T. H., & Choi, K. N. (2020). Design and implementation of interactive search service based on deep learning and morpheme analysis in NTIS system. *Journal of Convergence for Information Technology*, 10(12), 9-14. https://doi.org/10.22156/CS4SMB.2020.10.12.009

Lee, S., & Yang, M. (2021). A study on provision of national R&D information: Perspective of research conducting institutions. *Journal of D-Culture Archives*, 4(1), 149-158. https://doi.org/10.23089/jdca.2021.4.1.010

Lee, S., Lee, S., Seol, H., & Park, Y. (2008). Using patent information for designing new product and technology: Keyword based technology roadmapping. *R&D Management*, 38(2), 169-188. https://www.scilit.com/publications/393a2cf0b761f72cc4ade915fa45876e

Lim, J. S., & Kim, S. G. (2020). Quantifying innovation characteristics of national R&D investment through cross-cutting data analysis : Case study on NTIS nanotechnology. *Journal of The Korean Operations Research and Management Science Society*, 45(2), 13-23. https://doi.org/10.7737/JKORMS.2020.45.2.013

Lim, J. S., Hahn, H., Won, D. K., & Kim, S. (2019). Study on innovation measurement of national R&D investments for nanotechnology using data envelopment analysis. *Journal of Korea Technology Innovation Society*, 22(2), 207-219. https://koreascience.kr/article/JAKO201919867048948.page

Min, J. H., Song, Y., Shin, J., & Son, J. H. (2024). Classification of strategy types for global R&D collaboration: An analysis of Korea's partnerships with developing countries. *Journal of Climate Change Research*, 15(6), 1085-1098. https://doi.org/10.15531/KSCCR.2024.15.6.1085

Nam, Y. H., Gang, J., & Oh, H. J. (2017). Developing and applying the evaluation measures for web based information services of research records archives - focused on NTIS -. *Journal of Korean Library and Information Science Society*, 48(2), 345-372. https://doi.org/10.16981/kliss.48.201706.345

National Institutes of Health (NIH). (n.d.). *Research Portfolio Online Reporting Tools, Expenditures, and Results (RePORTER)*. https://reporter.nih.gov

National Institute of Informatics (NII). (n.d.). *KAKEN database*. https://kaken.nii.ac.jp/en

Oh, H. S. (2008). Comparative analysis of R&D information status: Korea, US, and Japan. *Journal of the Korean Society for Library and Information Science*, 42(3), 207-231. https://doi.org/10.4275/KSLIS.2008.42.3.207

Oh, S., & Park, C. (2020). The effect of government R&D support for middle-sized firm. *Productivity Research: An International Interdisciplinary Journal*, 34(1), 231-257. https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART002573200

Park, S. U. (2008). An analysis of economic effect of national knowledge information system development. *Journal of Information Science Theory and Practice*, 39(3), 73-94. https://www.kci.go.kr/kciportal/ci/sereArticleSearch/ciSereArtiView.kci?sereArticleSearchBean.artiId=ART001282585

Park, S. U. (2018). An economic ripple effect analysis of national science & technology information service : Focusing an input-output analysis. *Journal of Korea Technology Innovation Society*, 21(4), 1296-1312. https://koreascience.kr/article/JAKO201819867049447.page

Park, S. U., Lee, P. W., Hahn, S. H., Lee, S. H., & Hwang, S. W. (2012). A study on models of economic performance analysis for science and technology information management distribution project. *Journal of Information Management*, 43(3), 43-59. https://doi.org/10.1633/JIM.2012.43.3.043

Ro, Y. J., Cho, Y., & Sung, Y. (2018). A study on the performance of the Korean government R&D investment in the area of industrial technology. *Journal of Industrial Innovation*, 34(1), 163-189. https://doi.org/10.22793/indinn.2018.34.1.007

Shin, S. H., Yoon, Y. J., Yang, M. S., Kim, J. M., & Shon, K. R. (2011). A data cleansing strategy for improving data quality of national R&D information - case study of NTIS. *Journal of The Korea Society of Computer and Information*, 16(6), 119-130. https://doi.org/10.9708/jksci.2011.16.6.119

Shmagun, H., Shim, J., Oppenheim, C., Choi, K., Chun, K., & Kim, J. (2022). Open Science-friendly national R&D

knowledge and information infrastructure: A case study from Korea. *Septentrio Conference Series,* (1). https://doi.org/10.7557/5.6591

Shon, K. R., & Lim, J. T. (2009). Research of quality improvement by factors analysis data quality problem: Focus on national R&D information linking structure. *The Journal of the Korea Contents Association*, 9(1), 1-14. https://doi.org/10.5392/JKCA.2009.9.1.001

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039

Suh, S. H., Lee, S. Y., & Lee, B. H. (2014). Factors influencing the customer value and the performance of national R&D information services of the R&D managers. *The Journal of the Korea Contents Association*, 14(7), 480-494. https://doi.org/10.5392/JKCA.2014.14.07.480

Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, 9(1), 5233. https://doi.org/10.1038/s41598-019-41695-z

Won, Y., Kim, H., & Park, S. (2023). Case studies and a taxonomy of science of STI policy programs for exploring the evidence-based approach. *Innovation Studies*, 18(1), 63-96. https://doi.org/10.46251/INNOS.2023.2.18.1.63

Woo, C. W., & Lee, J. Y. (2020). Investigation of research topic and trends of national ICT research-development using the LDA model. *Journal of the Korea Convergence Society*, 11(7), 9-18. https://doi.org/10.15207/JKCS.2020.11.7.009

Yoo, K. C., Lee, T. H., Choi, S. H., & Lee, J. H. (2020). Feature analyze and research of national convergence R&D: With focus on the text mining. *Journal of Information Technology Applications and Management*, 27(1), 59-73. https://doi.org/10.21219/jitam.2020.27.1.059

Yoon, Y. S., Zo, H., Choi, M., Lee, D., & Lee, H. W. (2018). Exploring the dynamic knowledge structure of studies on the Internet of things: Keyword analysis. *ETRI Journal*, 40(6), 745-758. https://doi.org/10.4218/etrij.2018-0059

Yu, E. J., Choi, K. N., & Hwang, Y. (2020). Curation service to improve user's access to national R&D information : Focusing on issues R&D service. *The Journal of the Korea Contents Association*, 20(9), 1-10. https://doi.org/10.5392/JKCA.2020.20.09.001

# Call for Paper

Journal of Information Science Theory and Practice (JISTaP)

We would like to invite you to submit or recommend papers to Journal of Information Science Theory and Practice (JISTaP, eISSN: 2287-4577, pISSN: 2287-9099), a fast track peer-reviewed and no-fee open access academic journal published by Korea Institute of Science and Technology Information (KISTI), which is a government-funded research institute providing STI services to support high-tech R&D for researchers in Korea. JISTaP marks a transition from Journal of Information Management to an English-language international journal in the area of library and information science.

JISTaP aims at publishing original studies, review papers and brief communications on information science theory and practice. The journal provides an international forum for practical as well as theoretical research in the interdisciplinary areas of information science, such as information processing and management, knowledge organization, scholarly communication and bibliometrics.

We welcome materials that reflect a wide range of perspectives and approaches on diverse areas of information science theory, application and practice. Topics covered by the journal include: information processing and management; information policy; library management; knowledge organization; metadata and classification; information seeking; information retrieval; information systems; scientific and technical information service; human-computer interaction; social media design; analytics; scholarly communication and bibliometrics. Above all, we encourage submissions of catalytic nature that explore the question of how theory can be applied to solve real world problems in the broad discipline of information science.

Co-Editors in Chief : Gary Marchionini & Dong-Geun Oh

Please click the "Online Submission" link in the JISTaP website (http://www.jistap.org), which will take you to a login/account creation page. Please consult the "Author's Guide" page to prepare your manuscript according to the JISTaP manuscript guidelines.

Any question? Hyekyong Hwang (managing editor) : jistap@kisti.re.kr

JISTaP
Journal of Information Science Theory and Practice

# Information for Authors

The Journal of Information Science Theory and Practice (JISTaP), which is published quarterly by the Korea Institute of Science and Technology (KISTI), welcomes materials that reflect a wide range of perspectives and approaches on diverse areas of information science theory, application and practice. JISTaP is an open access journal operated under an Open Access Policy. See the section on Open Access for detailed information on the Open Access Policy.

## A. Originality and Copyright

All submissions must be original, unpublished, and not under consideration for publication elsewhere. Once an article is accepted for publication, all papers are accessible to all users at no cost. If used for other researches, its source should be indicated in an appropriate manner and the content can only be used for uncommercial purpose under Creative Commons license.

## B. Peer Review

All submitted manuscripts undergo a single-blind peer review process in which the identities of the reviewers are withheld from the authors.

## C. Manuscript Submission

Authors should submit their manuscripts online via AccessON Peer Review Management System Plus (ACOMS$^{+}$). Online submission facilitates processing and reviewing of submitted articles, thereby substantially shortening the paper lifecycle from submission to publication. After checking the manuscript's compliance to the Manuscript Guidelines, please follow the "Online Submission" hyperlink in the top navigation menu to begin the online manuscript submission process.

## D. Open Access and Creative Commons License Statement



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

## E. Manuscript Guidelines

Manuscripts that do not adhere to the guidelines outlined below will be returned for correction. Please read the guidelines carefully and make sure the manuscript follows the guidelines as specified. We strongly recommend that authors download and use the manuscript template in preparing their submissions.

1. Page Layout :

All articles should be submitted in single column text on standard Letter Size paper (21.59 × 27.94 cm) with normal margins.

2. Length :

Manuscripts should normally be between 4,500 and 9,000 words (10 to 20 pages).

3. File Type :

Articles should be submitted in Microsoft Word format. To facilitate the manuscript preparation process and speed up the publication process, please use the manuscript template.

4. Text Style :

- Use a standard font (e.g., Times New Roman) no smaller than size 10.
- Use single line spacing for paragraphs.
- Use footnotes to provide additional information peripheral to the text. Footnotes to tables should be marked by superscript lowercase letters or asterisks.

5. Title Page :

The title page should start with a concise but descriptive title and the full names of authors along with their affiliations and contact information (i.e., postal and email addresses). An abstract of 150 to 250 words should appear below the title and authors, followed by keywords (4 to 6).

**Author1**
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**

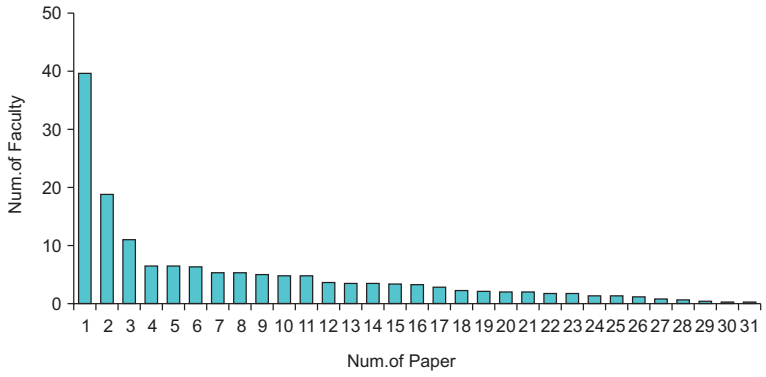



**Fig. 1.** Distribution of authors over publication count.

**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**

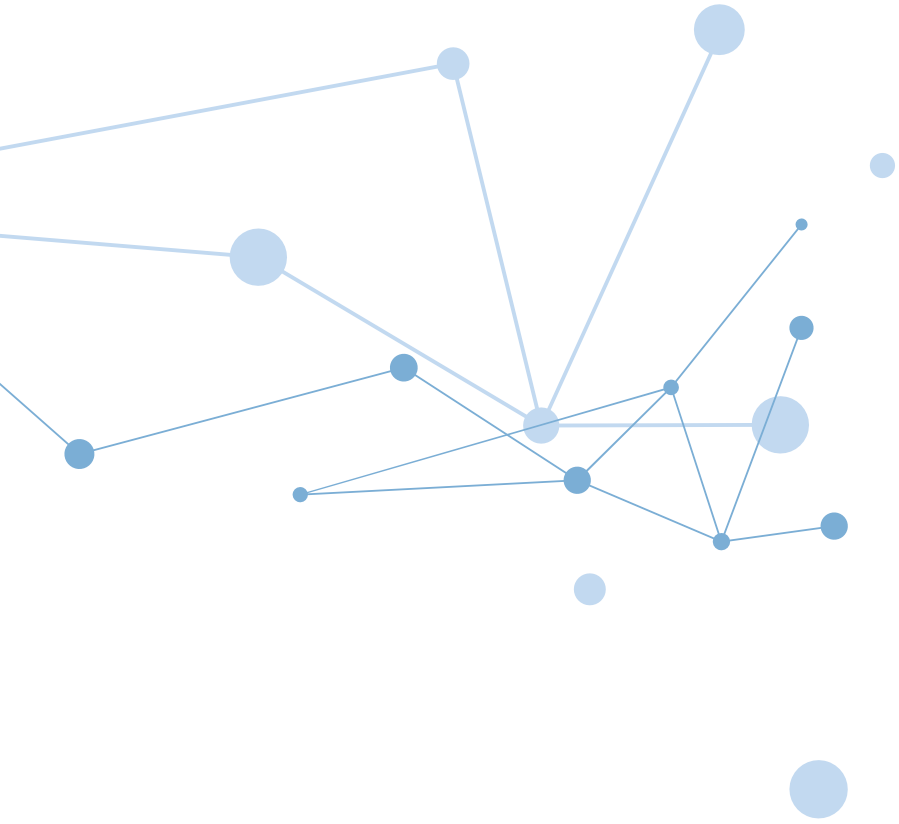
- Author, A., & Author, B. (Year). Chapter title. In A. Editor, B. Editor, & C. Editor (Eds.), *Book title* (pp. xx-xx). Publisher Name.
- Smith, K. & Brown, M. (2012). Author impact factor by weighted citation counts. In G. Martin (Ed.), *Bibliometric approach to quality assessment* (pp. 101-121). Springer.

**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
https://www.jistap.org