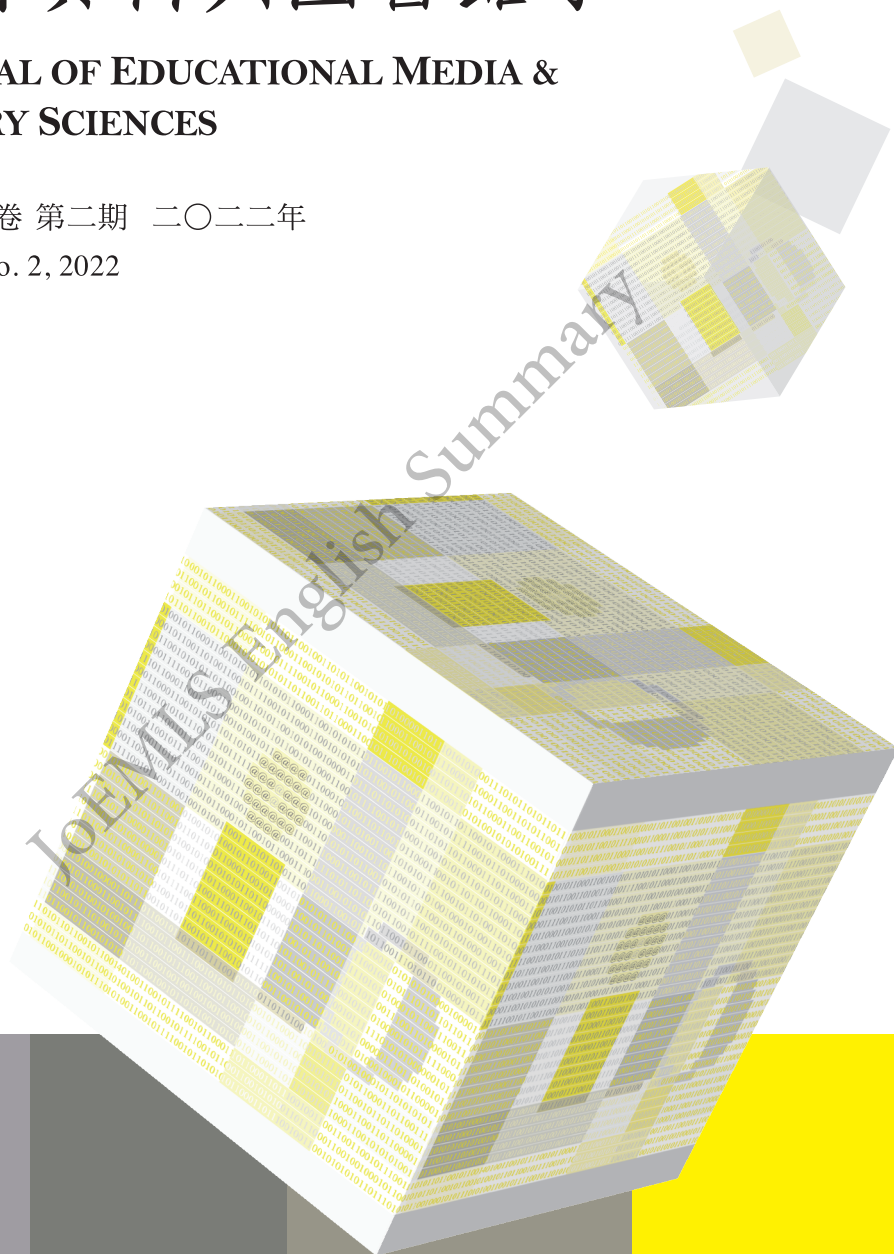


# 教育資料與圖書館學

JOURNAL OF EDUCATIONAL MEDIA &  
LIBRARY SCIENCES

第五十九卷 第二期 二〇二二年

Vol. 59, No. 2, 2022





教育資料與圖書館學，始於1970年3月創刊之教育資料科學月刊，其間於1980年9月更名為教育資料科學，並改以季刊發行。自1982年9月起易今名。另自2016年11月起，改以一年出版三期（3月、7月、11月）。現由淡江大學出版中心出版，淡江大學資訊與圖書館學系和覺生紀念圖書館合作策劃編輯。本刊為國際學術期刊，2008年獲國科會學術期刊評比為第一級，2015年獲科技部人文社會科學研究中心評定為教育學門專業類一級期刊。並廣為海內外知名資料庫所收錄(如下英文所列)。

**The JOURNAL OF EDUCATIONAL MEDIA & LIBRARY SCIENCES (JoEMLS)**, published by the Tamkang University Press and co-published with the Department of Information & Library Science (DILS) and Chueh Sheng Memorial Library, was formerly the **Bulletin of Educational Media Science** (March 1970 – June 1980) and the **Journal of Educational Media Science** (September 1980 – June 1982). In 2015, The JoEMLS is acknowledged as the first class scholarly journal in Taiwan by Ministry of Science and Technology (MOST). Since November 2016, the JoEMLS has been changed from quarterly to a tri-annual journal, published in March, July, and November.

The JoEMLS is indexed or abstracted in  
*Cabell's Directory of Publishing Opportunities*  
*Chinese Electronic Periodicals Service (CEPS)*  
*H.W. Wilson Database*  
*Index to Chinese Periodicals*  
*Library, Information Science & Technology Abstract (LISTA)*  
*Library & Information Sciences Abstracts (LISA)*  
*Library Literature & Information Science (LLIS)*  
*Scopus*  
*Taiwan Social Sciences Citation Index (TSSCI)*  
*Ulrich's Periodicals Directory*

# 教育資料與圖書館學

## JOURNAL OF EDUCATIONAL MEDIA & LIBRARY SCIENCES

### 主編 (Chief Editor)

邱炯友 (Jeong-Yeou Chiu)

政治大學圖書資訊與檔案學研究所教授

Professor, Graduate Institute of Library, Information and

Archival Studies, National Chengchi University, Taiwan

淡江大學資訊與圖書館學系兼任教授

Adjunct Professor, Department of Information and Library

Science, Tamkang University, Taiwan

### 執行編輯 (Executive Editor)

陳亞寧 (Ya-Ning Chen)

淡江大學資訊與圖書館學系副教授

Associate Professor, Department of Information and

Library Science, Tamkang University, Taiwan

### 名譽主編 (Editor Emeritus)

黃世雄 榮譽教授 (Professor Emeritus)

Shih-Hsion Huang

### 歷任主編 (Former Editors)

李華偉 教授 (Professor Hwa-Wei Lee)

李長堅 教授 (Professor Chang C. Lee)

### 編輯 (Managing Editor)

高禎熹 (Sz-Shi Kao)

林瑞慧 (Chang-Huei Lin)

### 編輯助理 (Editorial Assistants)

陳姿靜 (Tzu-Ching Chen)

陳思潔 (Sih-Jie Chen)

### 協同主編 (Associate Editor)

張瓊穗 (Chiung-Sui Chang)

淡江大學教育科技學系教授

Professor, Department of Educational Technology,

Tamkang University, Taiwan

### 英文協同主編 (English Associate Editor)

賴玲玲 (Ling-Ling Lai)

淡江大學資訊與圖書館學系副教授

Associate Professor, Department of Information and

Library Science, Tamkang University, Taiwan

### 地區協同主編 (Regional Associate Editors)

#### 大陸地區 (Mainland China)

張志強 (Zhiqiang Zhang)

南京大學出版科學研究所教授

Professor, Institute of Publishing Science at Nanjing

University, China

#### 歐洲地區 (UK and Europe)

Dr. Judith Broady-Preston

Emeritus Professor of Learning and Teaching,

Department of Information Studies,

University of Wales, Aberystwyth, UK

#### 美洲地區 (USA)

Dr. Jin Zhang

Professor, School of Information Studies,

University of Wisconsin-Milwaukee, USA

### 編務諮詢委員會 (Editorial Board)

林雯瑤 (Wen-Yau Cathy Lin)

淡江大學資訊與圖書館學系系主任

Chair, Department of Information and Library Science,

Tamkang University, Taiwan

宋雪芳 (Sheue-Fang Song)

淡江大學覺生紀念圖書館館長

Director, Chueh Sheng Memorial Library,

Tamkang University, Taiwan

陳雪華 (Hsueh-Hua Chen)

臺灣大學圖書資訊學系名譽教授

Professor Emeritus, Department of Library and Information

Science, National Taiwan University, Taiwan

梁朝雲 (Chaoyun Chaucer Liang)

臺灣大學生物產業傳播暨發展學系教授

Professor, Department of Bio-Industry Communication and

Development, National Taiwan University, Taiwan

曾元顯 (Yuen-Hsien Tseng)

臺灣師範大學圖書資訊學研究所教授

Professor, Graduate Institute of Library & Information

Studies, National Taiwan Normal University, Taiwan

黃鴻珠 (Hong-Chu Huang)

淡江大學資訊與圖書館學系榮譽教授

Professor Emeritus, Department of Information and

Library Science, Tamkang University, Taiwan

蔡明月 (Ming-Yueh Tsay)

政治大學圖書資訊與檔案學研究所講座教授

Chair Professor, Graduate Institute of Library, Information and

Archival Studies, National Chengchi University, Taiwan

薛理桂 (Li-Kuei Hsueh)

政治大學圖書資訊與檔案學研究所名譽教授

Emeritus Professor, Graduate Institute of Library, Information

and Archival Studies, National Chengchi University, Taiwan

方卿 (Qing Fang)

武漢大學信息管理學院教授

Professor, School of Information Management,

Wuhan University, China

沈固朝 (Guchao Shen)

南京大學信息管理學院教授

Professor, School of Information Management,

Nanjing University, China

Pia Borlund

Professor, Faculty of Social Sciences, Department of

Archivistics, Library and Information Science, Oslo

Metropolitan University, Norway

Sam Hastings

Professor (Retired), Faculty, School of Library &

Information Science, University of South Carolina, USA

Edie Rasmussen

Professor, School of Library, Archival and Information

Studies, University of British Columbia, Canada

Josephine Sche

Professor, Information and Library Science Department,

Southern Connecticut State University, USA

Peter Sidorko

Senior Consultant, The University of Hong Kong Libraries,

The University of Hong Kong, Hong Kong

Hong Xu

University Librarian, Duke Kunshan University, China

## JoEMLS 編輯政策

本刊係採開放存取 (Open Access) 與商業資料庫付費途徑，雙軌發行之國際學術期刊，兼具電子版與紙本之平行出版模式。本刊除秉持學術規範與同儕評閱精神外，亦積極邁向InfoLibrary寓意之學域整合與資訊數位化理念，以反映當代圖書資訊學研究趨勢、圖書館典藏內容與應用服務為本；且以探討國內外相關學術領域之理論與實務發展，包括圖書館學、資訊科學與科技、書業與出版研究等，並旁及符合圖書資訊應用發展之教學科技與資訊傳播論述。

### 開放取用政策

在作者授權之基礎下，讀者可經由本刊網站立即取得本刊全文資料。本刊遵從DOAJ對於開放取用 (Open Access) 的定義以及布達佩斯宣言。讀者無須註冊即可閱讀本刊內容，本刊亦不向投稿者及作者收取文章處理費 (article processing charge, APC)。

### 典藏政策

JoEMLS向來以「綠色期刊出版者」(Green Publisher / Journal) 自居，同意且鼓勵作者將自己投稿至JoEMLS之稿件，不論同儕評閱修訂稿與否，都能自行善加利用處理，但希望有若干限制：

- (1)勿將已刊登之修訂稿 (post-print) 再自行轉為營利目的之使用；
- (2)典藏版以期刊排印之PDF檔為首選；
- (3)任何稿件之典藏版本皆須註明其與JoEMLS之關係或出版後之卷期出處。

## JoEMLS Editorial Policy

The JoEMLS is an Open Access (OA) Dual, double-blind reviewed and international scholarly journal dedicated to making accessible the results of research across a wide range of Information & Library-related disciplines. The JoEMLS invites manuscripts for a professional information & library audience that report empirical, historical, and philosophical research with implications for librarianship or that explore theoretical and practical aspects of the field. Peer-reviewed articles are devoted to studies regarding the field of library science, information science and IT, the book trade and publishing. Subjects on instructional technology and information communication, pertaining to librarianship are also appreciated. The JoEMLS encourages interdisciplinary authorship because, although library science is a distinct discipline, it is in the mainstream of information science leading to the future of InfoLibrary.

### Open Access Policy

With author's authorization, readers can immediately obtain full texts for free online via our journal website. We comply with the DOAJ definition on Open Access, and the Budapest Open Access Initiative (BOAI) which means we allow and provide. There is no requirement for readers to register to read the contents of JoEMLS, and JoEMLS does not charge article processing charge (APC) to contributors and authors.

### Archiving Policy

The JoEMLS, as a role of "OA green publisher/journal," provides free access online to all articles and utilizes a form of licensing, similar to Creative Commons Attribution license, that puts minimal restrictions on the use of JoEMLS's articles. The minimal restrictions here in the JoEMLS are:

- (1) authors can archive both preprint and postprint version, the latter must be on a non-commercial base;
- (2) publisher's PDF version is the most recommend if self-archiving for postprint is applicable; and
- (3) published source must be acknowledged with citation.



Contents

EDITORIAL

- Opportunities for Deepening the Reform of  
the Metadata Elements of Journal Articles**  
Jeong-Yeou Chiu 97

RESEARCH ARTICLES

- To Use or Not to Use? Exploring the Factors  
Influencing Professional Reusers'  
Intention to Adopt and Utilize  
Governmental Open Data in Taiwan**  
Tung-Mou Yang & Yi-Jung Wu 101
- Effects of Afterschool Programmes in  
Public Libraries on Disadvantaged  
Children: A Case of New Taipei City  
Library in Taiwan**  
Ya-Wei Chuang & Hui-Yun Sung 137
- Exploring Article Process Charge of  
Open Access Journals from the  
Perspectives of Publication  
Characteristics and Citation Impact  
Indicators: A Case Study in the  
Medical Field**  
Chia-Yu Lin & Wen-Yau Cathy Lin 169
- Estimation of Topic Similarity and  
Its Application to Measuring Stability  
of Topic Modeling**  
Sung-Chien Lin 201



## EDITORIAL

# Opportunities for Deepening the Reform of the Metadata Elements of Journal Articles

The metadata of journal articles generally refers to information regarding journal name, volume and issue number, article's title, name(s) of author(s), keywords, page range, etc. However, in order to help readers enhance their understanding of article contents when searching for articles, most journal publishers and database organizations present the above metadata together with the abstract of the article to facilitate readers to quickly identify its content. For scholarly articles, in addition to the academic value of the text itself, the references cited by the authors in the texts have significant research value in all the issues in scholarly communication and publication. The references not only present the author's exploration and recognition of previous scholarship, but also assist future generations to further discover scientific trajectories and disseminate scholarly knowledge through the cited references.

Although open access (OA) to academic journals continues to flourish in the global academic community, most academic publishers focus on the open access and availability of full-text digital files of articles. In contrast to the opening of digital files, the release of metadata and their subsequent use are still awaiting the discussion and efforts of the industry, government, and academia. In order to implement the concept of OA, some groups have advocated opening up the abstracts, together with the references, in articles in recent years, as a result of which two initiatives, the Initiative for Open Citations (I4OC) and the Initiative for Open Abstracts (I4OA), have come into being. Of the two, the I4OC even hopes that academic publishers or journal publishers, under the premise of providing free bibliographic data, can provide references at the same time, so as to promote the release of citation data in a structured, separable, and open-access condition, and then proceed to build a copyright-free academic citation database.

At the present stage, most of the search fields provided to readers in the academic literature databases or official websites established by journals in Taiwan are such metadata as journal name, volume and issue number, article's title, name(s) of author(s), keywords, and page range, but not including the element of references. If readers want to retrieve the reference literature, they must first confirm whether full-text downloading is permitted before further downloading the electronic file of a single article, from which they make the

searches and checking; or readers must leave the search interface of the literature database or journal website and navigate to the Taiwan Humanities and Social Sciences Citation Index Database (free search) or Scopus and Web of Science (subscription required) to do their searches and checking, without being able to obtain the research information needed at a one-stop location. It is hoped that academic publishers and journal publishers can work together with the government and private databases, striving to integrate the complete metadata of Taiwan's academic journals and gear them to international standards, so as to enhance the professional image of Taiwan's academic publishing.

In this issue (Issue 2, Volume 59), 10 manuscripts have gone through the review process. Two manuscripts were rejected at the internal review process for format evaluation. Eight manuscripts have gone through the whole review process, and four were accepted, with a rejection rate of 50% (4 out of 8). The articles published in this issue include: "To Use or Not to Use? Exploring the Factors Influencing Professional Reusers' Intention to Adopt and Utilize Governmental Open Data in Taiwan" by Tung-Mou Yang and Yi-Jung Wu, "Effects of Afterschool Programs in Public Libraries on Disadvantaged Children: The Case of the New Taipei City Library" by Ya-Wei Chuang and Hui-Yun Sung, "Exploring Article Process Charge of Open Access Journals from the Perspectives of Publication Characteristics and Citation Impact Indicators: A Case Study in the Medical Field" by Chia-Yu Lin and Wen-Yau Cathy Lin, and "Estimation of Topic Similarity and Its Application to Measuring Stability of Topic Modeling" by Sung-Chien Lin. Special thanks are dedicated to the reviewers and authors.

Jeong-Yeou Chiu  
*JoEMLS* Chief Editor



# To Use or Not to Use? Exploring the Factors Influencing Professional Reusers' Intention to Adopt and Utilize Governmental Open Data in Taiwan

Tung-Mou Yang<sup>a\*</sup> Yi-Jung Wu<sup>b</sup>

## Abstract

*In recent years, open government data has become one of the prevailing policy implementations among government administrations around the world. Researchers maintain that open data providers and users play critical roles in forming a sound open data ecosystem. However, recent studies have found that open data use has not kept up with expectations, with the number of open data applications increasing slowly. Therefore, using a qualitative research approach to focus on professional reusers, this study explores the determinants that influence professional reusers' intention to use governmental open data. With qualitative empirical data support, the identified determinants include perceived usefulness, perceived effort, external influence, facilitating condition, legislation and license, self-efficacy, and perceived risk. In addition, the determinants are incorporated into the theory of planned behavior to investigate how the determinants act as behavioral, normative, and control beliefs in influencing professional reusers' intentions. Further, this study discusses related suggestions that can strengthen the sustainability of an open data ecosystem. The discussion and practical implications of this study are expected to provide insights to both practitioners and policymakers for further developing open data policies and enriching the current open data-related literature.*

**Keywords:** *Open data, Open government data, Open data use, Open data users, Influential factors, Taiwan*

<sup>a</sup> Associate Professor, Department of Library and Information Science, National Taiwan University, Taipei, Taiwan

<sup>b</sup> Associate Professor, Department of Public Policy and Management, Shih Hsin University, Taipei, Taiwan

\* To whom all correspondence should be addressed. E-mail: tmyang@ntu.edu.tw

## Introduction

Government agencies are perceived as entities possessing various resources that can be shared and utilized across the boundary of the public and private sectors. In particular, while government agencies are usually considered the largest data creators and collectors across a nation's different domains, open government data in recent years has gradually become an important policy implementation adopted by government administrations around the world (World Wide Web Foundation, 2018). The number of countries with open data portals has increased significantly from 46 in 2014 to 153 in 2020 (United Nations Department of Economic and Social Affairs, 2020). Open data refers to the concept in which datasets are published online in electronic forms that are machine-readable and with a non-proprietary format, and the datasets can be freely accessed, used, modified, and shared by anyone at any place for any purpose (The World Bank, n.d.). The goal of open government data is to enhance the fundamental principles of open government, including transparency, participation, and collaboration, by ensuring public engagement in using governmental open data. The public is expected to utilize governmental open data to generate innovative data usage and applications, spur economic growth, and bring positive impacts to society (Attard et al., 2015; Janssen et al., 2012; Janssen & Zuiderwijk, 2014). A common belief maintains that when government agencies publish more datasets on open data platforms, the public will be motivated to reuse the datasets. However, recent research indicates that open data use has not kept up with expectations, although related infrastructure, such as open data platforms, has been available online to provide services (Najafabadi & Luna-Reyes, 2017; Zuiderwijk et al., 2016). In addition, there are difficulties in obtaining people's participation in using governmental open data. The public seems to agree with an open data movement but does not necessarily engage in reusing the published datasets (Hellberg & Hedström, 2015). While open data users play a critical role in forming a sound open data ecosystem, if the projected data users do not use the published datasets, the objective of open data initiatives can be futile (Attard et al., 2015; Hivon & Titah, 2017). Therefore, it is necessary to understand the factors that influence people's engagement in using governmental open data.

Researchers have recently classified open data users into two categories: direct reusers and end users (Abella et al., 2019). Direct reusers can be professional reusers and social reusers. Professional reusers refer to entrepreneurs and private companies that use open data to innovate and create for-profit products and services. Social reusers represent non-profit organizations such as NGOs providing services to others. On the other hand, end users mean entities, organizations, or citizens consuming open data-related products and



services provided by direct reusers. In addition, researchers have suggested that there should be more open data studies conducted in different contexts, such as sources, areas, and countries, to explore the use of governmental open data and how related open data policies can be further developed (Magalhães & Roseira, 2016; Susha et al., 2015; Zuiderwijk & Hinnant, 2019). Even to this day, there is still limited research on open data use focusing on professional reusers in the context of Taiwan's open government data. Therefore, this research explores and discusses open data use from the perspective of Taiwan's professional reusers using governmental open data to develop for-profit business applications. Specifically, the following research questions are investigated in this exploratory study: 1. What are the critical factors influencing professional reusers' intention to use governmental open data? 2. What is the nature of the impact of these factors? 3. What strategies can be employed to sustain an open data ecosystem based on the perspective of professional reusers?

The rest of the paper is organized as follows. In the following section, the global open data movement is first briefly introduced. Then, the recent literature on open data research and open data use is reviewed, and the proposed research is subsequently presented. Next, the paper describes the research design and methods of this study, in which the employed case, data collection, and data analysis are presented. Then, the paper discusses the findings and implications of the empirical data analysis. Lastly, the conclusion section expounds on the contribution and limitations of the current study. Future research directions are also suggested.

## **Literature Review**

### **The Movement of Open Government Data**

In this global movement, the U.S. government first established its open data portal in 2009 to provide data access to its federal government agencies. Open government data became an important pillar in supporting the information-centric strategy of the U.S. digital government (Digital Government, n.d.). The OPEN Government Data Act (Open, Public, Electronic, and Necessary Government Data Act) was further enacted, requiring U.S. federal agencies to make datasets open online by default in a form that is machine-readable and freely reusable (H.R.4174 –115th Congress [2017-2018]; Foundations for Evidence-Based Policymaking Act of 2018, 2019). Similarly, in 2010 the U.K. government initiated its open data portal to release datasets of its central and local government agencies. A government report, the Open Data White Paper, was published in 2012 by the U.K. government to emphasize the potential of publishing governmental open data to achieve transparency and accountability (UK Minister of State for the Cabinet

Office, 2012). By the end of 2012, The European Union (EU) also established its open data portal, allowing the public to access data released from EU agencies and institutions. An Open Data Charter was signed in 2013 by the Group of Eight (G8) leaders to promote the principles of open government data, including open data by default, quality and quantity, usable by all, improving governance, and enhancing innovation. In 2014, the leaders of the largest G20 industrial economies agreed to use governmental open data as a tool against corruption. The Open Government Partnership, dedicated to making government administrations more open, accountable, and responsive to citizens, also enacted its Open Data Charter by the end of 2015 to provide open data principles to its member countries. Notably, the Charter has been adopted by 79 national and local governments from around the world (Open Data Charter, 2021). According to the 2020 United Nations E-Government Survey, the number of countries implementing open government data policies has been growing steadily, and 153 countries have established their respective open data portals (United Nations Department of Economic and Social Affairs, 2020). Likewise, intergovernmental organizations such as the United Nations, OECD, and World Bank have initiated open data sites and projects to provide the public access to their datasets.

### **The Recent Research on Open Government Data**

Researchers from different disciplines, such as information science, information management, public administration, computer science, and law, have explored the complexity of open government data from various perspectives and knowledge areas (Hossain et al., 2016; Zuiderwijk & Hinnant, 2019). Government agencies appear to have a risk-averse culture and act conservatively in opening their datasets (Peled, 2011). As the World Wide Web Foundation (2018) indicates in its Open Data Barometer report, the vast majority of governmental datasets remain closed to the public. The progress of embedding open data policies is slow, and governments still treat open data as a side project. Another recent study also indicates that most of the published government resources on open data sites are informational data rather than granular data, and only a small number of the datasets advertised as open data are actually open (V. Wang & Shepherd, 2020).

Therefore, from the perspectives of technology, organization, legislation and policy, and environment, researchers have studied and discussed related social and technical enablers and impediments that have an impact on government agencies in implementing open data policies (Barry & Bannister, 2014; Conradie & Choenni, 2014; Janssen et al., 2012; Yang et al., 2015). It is indicated that the reuse of governmental data can be an obscure policy domain, and the interrelation between open data and other policy areas is complicated (Bates, 2014).

Meanwhile, government agencies tend to lack clear guidelines for publishing open data to external parties. Researchers suggest that it is important to create a system and enact policies to promote the release of open data by reducing the social, economic, and political impacts that government agencies may encounter (Gerunov, 2017; Nugroho et al., 2015). For instance, Zuiderwijk and Janssen (2014) proposed a framework to improve the development of both new and existing open data policies by taking in perspectives of environment, policy content, and performance evaluation. Dawes et al. (2016) developed a general model for planning open data programs using an ecosystem approach to address various perspectives of open data initiatives, including policy and strategy, data publication and use, feedback and communication, and stakeholder interactions.

In addition, researchers investigated the emerging open data business models to understand their characteristics, patterns, and strategies used to create economic opportunities (Zeleti et al., 2014). Different types of infomediary business models driven by governmental datasets were identified (Janssen & Zuiderwijk, 2014). In terms of open data assessment, measurement frameworks were developed to evaluate the maturity and progress of open data initiatives in government administrations (Chu & Tseng, 2016; Solar et al., 2012; Yang & Wu, 2019; Zhu & Freeman, 2019). Further, related indexes were also developed for the evaluation and categorization of open data portals and their metadata qualities (Kubler et al., 2018; Thorsby et al., 2017). Similarly, the Open Knowledge Foundation and World Wide Web Foundation also built their own benchmarks—the Global Open Data Index and the Open Data Barometer—to evaluate the open data developments of global government administrations.

### **The Use of Governmental Open Data**

Researchers have indicated that publishing data alone is not enough to enable the life cycle of open government open (Attard et al., 2015). After data publication, data users must be able to discover and access data for data consumption. Likewise, according to Dawes et al.'s (2016) open data ecosystem model, one of the key stakeholders are users who utilize open data in conducting analyses and developing applications. In addition to government agencies acting as data providers, data users play a critical role in the success of open data initiatives. Therefore, there is a need to better understand the interaction between government agencies and open data users, considering its sustainability is crucial to the development of an open data ecosystem (Hivon & Titah, 2017).

A recent study has identified five major motives for open data use, including exploring creativity, creating business value, enabling local citizen value, addressing global societal challenges, and advocating the open data agenda (Lassinanti et al., 2019). It is indicated that individuals' open data use can be

influenced by policies, processes, and infrastructure used to provide open data (Susha et al., 2015). In particular, impediments may exist while users attempt to access and use datasets (Zuiderwijk et al., 2012). A recent investigation points out that many open data portals are still in a very early stage of development, and a great deal of work is needed to help the public understand and use data (Thorsby et al., 2017). While open data use includes activities to search, identify, and download datasets, open data portals should also make the related processes easier to facilitate users in obtaining datasets (Dawes et al., 2016). In particular, a machine-readable and non-proprietary format is essential to enable easier access and unrestricted use (Attard et al., 2016).

Researchers have indicated that users' trust in the quality of open data weighs on whether the users would engage in using datasets (Purwanto et al., 2020; Zhu & Freeman, 2019; Zuiderwijk et al., 2016). Open data without proper quality control may jeopardize dataset reuse and bring negative impacts on civic participation (Weerakkody et al., 2017). Thus, the importance of metadata of published datasets is also emphasized; it is expected to help in the use of open data by enhancing user experience (Zuiderwijk et al., 2016).

It is suggested that open data infrastructure should maintain mechanisms to respond to the questions, problems, and suggestions raised by open data users (Purwanto et al., 2020; Zhu & Freeman, 2019; Zuiderwijk et al., 2016). Communication channels allow open data users to request datasets and report errors for government agencies to improve the quality of the released datasets (Yang et al., 2015; Zuiderwijk et al., 2016).

Researchers have further argued that outdated laws and policies could prevent data from being used to create value. For instance, the inexistence or inconsistency of licensing in the datasets across different jurisdictions could have an impact on open data use (Attard et al., 2016; Magalhães & Roseira, 2016). Thus, detailed information should be provided for users to understand how open data can be used in compliance with related licenses and regulations (Kaasenbrood et al., 2015).

It is also found that the availability of citizens' resources matters in open data use (Purwanto et al., 2020). The lack of fundamental skills and expertise for processing and analyzing data could act as an impediment to utilizing open data for generating values (Magalhães & Roseira, 2016; Safarov et al., 2017). In order to promote the public adoption of open data, it is suggested that required skills and expertise be made available to help open data users participate further (Safarov et al., 2017; Susha et al., 2015).

### **Conceptualizing the Adoption of Open Data Use**

Open government data can be considered an innovative e-government service that allows the public to adopt and use open data for respective purposes

without restriction. The public using open data needs to go through the processes of data discovery, data exploration, and data exploitation (Attard et al., 2015). When accessing and utilizing open data, users have to adapt to various open data technologies within an open data infrastructure, including open data portals, related application programming interfaces, various data formats, linked data vocabularies, and metadata elements (Zuiderwijk et al., 2015). Accordingly, as an emerging innovative service in the public sector, open government data must obtain the public's participation and collaboration in forming a sound open data ecosystem; the public's adoption of using governmental open data can be conceptualized as a process of innovation acceptance. Specifically, the unified theory of acceptance and use of technology (UTAUT) is a commonly utilized framework that helps discuss technology and innovation acceptance. Its four constructs are performance expectancy, effort expectancy, social influence, and facilitating condition (Venkatesh et al., 2003). The four constructs are suggested to influence a user's intention to accept an innovative system or technology. Performance expectancy refers to the degree to which an individual believes that using a system or technology will help achieve gains in performance. Effort expectancy refers to a system or technology's degree of ease to use. Further, social influence means the degree to which an individual thinks that important others suggest them to use a system or technology. Facilitating condition refers to the degree to which an individual believes that an organizational and technical infrastructure or related resources are there to support the use of a system or technology.

As aforementioned, researchers suggest that while open government data is a complex phenomenon, investigations should also take into account the environment and legal context in which related stakeholders reside (Zuiderwijk & Hinnant, 2019). In addition, derived from the perspective of behavioral psychology, self-efficacy refers to an individual's perception of their ability to plan and execute certain actions to produce specific performance and reach a particular goal (Bandura, 1977). In the e-government literature, researchers have suggested that a user's self-efficacy matters in the adoption and use of e-government services (Hung et al., 2013; Rana et al., 2015). They emphasized that users could consider their capabilities, which is expected to affect whether or not they adopt and use innovative e-government services.

Furthermore, based on the field of social psychology, the theory of planned behavior represents a well-established framework that embraces and discusses various factors that affect an individual's intention to perform a certain behavior (Ajzen, 1991). In particular, the theory of planned behavior has also been extensively adopted by e-governance researchers in discussing users' intention



toward the adoption and use of innovative government services (Hung et al., 2013; Ozkan & Kanat, 2011; Rana et al., 2015, 2016; H.-J. Wang & Lo, 2013). For instance, Hung et al. (2013) used the theory of planned behavior to identify the factors determining users' intention to use governmental mobile services. Rana et al. (2015, 2016) utilized the theory's concepts to explore how influential factors affect citizens' adoption of e-government systems. H.-J. Wang and Lo (2013) also used the theory as a theoretical framework to investigate the factors influencing citizens' intention to use government websites. According to Ajzen (1991), the framework suggests that an individual's behavioral intention is a function comprising three antecedent constructs: attitude toward the behavior, subjective norm, and perceived behavioral control. Specifically, attitude toward the behavior refers to the degree to which the performance of a behavior is either positively or negatively evaluated. In addition, it can be determined by an individual's behavioral beliefs. A behavioral belief refers to the subjective probability that a given outcome or experience can be generated by performing a behavior. On the other hand, subjective norm means the perceived social pressure to engage or not to engage in a behavior. It is pointed out that subjective norms can be determined by an individual's normative beliefs. A normative belief is defined as the reflection of an individual's perceived expectations from other individuals, groups, and organizations to engage in a particular behavior. Lastly, perceived behavioral control refers to an individual's perception of their capability to engage in a given behavior. It is suggested to be determined by control beliefs, which are the perceived presence of factors that may help or hinder an individual's ability to perform a behavior.

### **The Proposed Research**

While the number of governmental datasets that have been opened to the public keeps growing, the number of applications using open data also increases slowly (Najafabadi & Luna-Reyes, 2017; Zuiderwijk et al., 2016). Researchers point out that there are difficulties in obtaining people's interests in using open data. Although the public seems to like the idea of open government data policies, they do not actively participate in the process of data use (Hellberg & Hedström, 2015). The lack of users exploiting open data resources indicates the need to understand what factors influence open data use and what strategies attract and stimulate users to participate (Attard et al., 2015; Zuiderwijk et al., 2016). Researchers also suggest that there should be more studies that explore open data adoption in different contexts, such as areas, sources, and countries, to understand the various factors that influence open data usage of different types of data reusers (Magalhães & Roseira, 2016; Susha et al., 2015). Findings within different socio-technical contexts can help extend the knowledge of how open

data policies could be further developed and implemented (Zuiderwijk & Hinnant, 2019). Therefore, this research explores the influential factors of open data use from users' perspectives in the context of Taiwan's open government data. The aforementioned frameworks are expected to help conceptualize the factors that explain the phenomenon of interest with empirical data support. In particular, while one of the major motives of open data use is to create business value, this study focuses on professional reusers who utilize governmental open data to develop innovative applications for business services.

### **Research Design and Method**

Since the inception of the Freedom of Government Information Law in Taiwan, the Taiwan government has gradually institutionalized related practices for the public to access government information and data. In particular, the open government data movement has received significant attention from both public and private sectors in Taiwan. Taiwan's central and local government agencies have implemented open data policies and established open data infrastructure, such as open data sites, to open their datasets. In November 2011, the Taipei City government established Taiwan's first open data website (<https://data.taipei>). On the other hand, New Taipei City also made the debut of its open data portal (<https://data.ntpc.gov.tw>) in December 2012. While recognizing the critical role and value of an open government, at the end of 2012, the Prime Minister of Taiwan required the Research, Development, and Evaluation Commission to devise and enact open data policies at the central government level. Subsequently, Data.gov.tw made its debut in April 2013 and acted as the open data portal for the central government agencies of Taiwan. Under an executive order from the Prime Minister, each central government agency of Taiwan was required to open at least fifty datasets by the end of 2013. The number of total datasets available on Data.gov.tw is currently 57,275 (as of June 12, 2022), which has grown significantly compared to the number of datasets several years ago (3,187 datasets as of December 10, 2014). Meanwhile, the other four city governments of the six major municipalities of Taiwan, including Taoyuan City, Taichung City, Tainan City, and Kaohsiung City, have enabled their respective open data portals to provide services. Most of the other local governments have also participated in the open data policy implementation. Noteworthy, the Taiwan government has received high rankings in the Global Open Data index for two consecutive years in 2016 and 2017—proof of the Taiwan government's efforts in implementing open data policies. Further, in forming a sound open data ecosystem, the Taiwan government has also engaged in promoting open data use through various activities, such as workshops, hackathons, and open data contests. Hence, given

its efforts and commitment to implementing open data policies and promoting the use of open data, the context of Taiwan's open government data presents one of the good cases for conducting this exploratory research.

In addition, regarding open data user studies in the context of Taiwan's open government data, Lo et al. (2014) first employed a survey approach to explore data users' perceptions of related legislations and policies, sharing approaches, technological standards and data formats, open data scopes, and open data promotion. Similarly, other researchers utilized the survey approach to investigate the factors that may influence data users' satisfaction and intention of open data adoption—that is, information quality, user interface quality, computer self-efficacy, and social influence (Chen, 2015; H.-J. Wang, 2020). H.-J. Wang and Lo (2019) also used a survey approach and indicated the importance of top management support and competitive pressure for firms' adoption of governmental open data. However, there is still limited research focusing on professional open data reusers in the context of Taiwan's open government data using a semi-structured interview approach to discuss in-depth qualitative findings. Accordingly, as previously mentioned, this study is expected to fill this research gap.

The major challenge in this research is the difficulty in identifying and connecting open data users. According to the open data principles, governmental open data sites can be accessed by anyone from any place at any time without registration. Therefore, the researchers of the study found it relatively difficult to identify and approach professional reusers who use governmental open data to develop for-profit business applications. Consequently, a purposive sampling approach was applied to locate relevant candidates for conducting interviews. Notably, purposive sampling is usually employed for selecting information-rich cases to conduct in-depth qualitative studies (Wengraf, 2001).

In promoting open data use, the Industrial Development Bureau of the Ministry of Economic Affairs (the IDB of MOEA) has been the responsible central government agency for holding an annual event for an open data contest. Held by the agency for several years, this contest is one of the major government-held events for promoting open data use, attracting professional reusers who use open data for conducting various business applications. The major aim of the contest is to invite and encourage entrepreneurs and private companies to develop innovative business applications using governmental open data. The award-winning companies of the contest receive high prizes and consultations from the government to support their business models and applications. Venture capital firms are also invited to the event to invest in the companies they are interested in.

Therefore, the researchers of this study considered this event a great opportunity to approach professional reusers for conducting interviews. After directly contacting the IDB of MOEA, the researchers had the opportunity to attend the convention of the annual contest held on December 6, 2019. The award-winning entrepreneurs and private companies were invited during the convention to present their applications using governmental open data. Thus, the researchers interacted with those award-winning contest participants to recruit interviewees for this study. There were a total of fourteen contest participants who agreed to participate in this study as interviewees.

All recruited interviewees had extensive experience using governmental open data in developing business applications. In their respective companies, they held key positions, such as product manager, chief executive officer, chief technology officer, chief operating officer, technical director, general manager, and vice president. Particularly, the interviewees were from various business industries and used a variety of open government data, such as geographic data, custom data, health, and medical data, weather data, economic data, traffic data, transportation data, tourism data, and legal case data. With various professional backgrounds and experience in using governmental open data, the interviewees provided rich information to the researchers in this exploratory research. In addition, while the interviewees of this study were recruited from private companies of different business industries, it also helped achieve the goal of multiple sources of evidence to enhance the quality of the obtained qualitative empirical data.

Subsequently, the interviews were conducted between March and April 2020. This study employed a semi-structured interview to collect qualitative empirical data, given its flexibility to follow up on new information and explore new findings (Bryman, 2004). The interview questions were designed to lead the interviewees toward helping the researchers identify the answers to the research questions. The interviews were recorded using digital recorders and field notes and were transcribed for later data analysis and report writing. The average duration of interviews was about one hour and thirty minutes.

During data analysis, the researchers adopted qualitative data analytic techniques to analyze data and identify common patterns. The utilized techniques were open coding, axial coding, and selective coding (Strauss & Corbin, 1998). The interview data were reviewed and analyzed line by line during the open coding process. The initial codes were associated with the text segments extracted from the interview data, representing the concepts derived from the data. Subsequently, axial coding refined, aligned and classified the initial codes generated in open coding. Conceptually similar codes were then grouped to

form categories. Lastly, selective coding was employed to select and interpret the relationships of the categories of concepts from axial coding and confirm whether the elaborations and interpretations were inherent in and comply with the observed phenomena of this study. Atlas.ti was utilized in this study to analyze the qualitative data.

## **Findings**

Based on the empirical data analysis, seven influential factors were identified in this study: perceived usefulness, perceived effort, external influence, facilitating condition, legislation and license, self-efficacy, and perceived risk, which are in line with the concepts in the aforementioned literature for conceptualizing the adoption of open data use. With qualitative empirical data support, the seven identified factors are discussed in the following subsections.

### **Perceived Usefulness**

The interviewees indicated that perceived usefulness is important when first adopting and using governmental open data. They expected that the obtained open datasets could benefit their existing business operations by reducing cost and increasing efficiency. In particular, open data is free for all to access, and the published datasets may include a variety of fields, such as weather data and geographic data, which can be extremely costly and difficult for small businesses or entrepreneurs to collect by themselves. For instance, an interviewee (Chief Executive Officer, P4\_04) explained:

...The most important help is that governmental open data helps us reduce significant operational costs. It was fairly difficult for us to obtain geographic data. For instance, it is impossible for my company to investigate the geographic distribution of fault lines by itself. We have to admit that we don't have the capability and resources to collect such data....  
(Chief Executive Officer, P4\_04)

Some interviewees further pointed out that, for some datasets, such as legal cases involving car accidents, ridership of subway stations, and company registrations, the government is the public authority to collect and generate the datasets and is usually the only data source that the interviewees claim they have to rely on.

While one of the major goals of open government data is to unleash the potential of governmental datasets to spur innovations and economic growth, some interviewees also pointed out that open data use could act as a core catalyst that drives innovative business models. They expected that open data could be integrated with their own datasets to generate applications through dataset mashups. For instance, an interviewee (Chief Technology Officer, P24\_06) stated:



...Open data usually represents raw datasets, which means there are plenty of opportunities that you can dig or mine some things from the datasets by combining other datasets or resources that you might already have at hand. For instance, we use the open data of xxx government agency to predict the trend of material supply and demand that can be very valuable information to our clients.... (Chief Technology Officer, P24\_06)

However, many interviewees indicated that government agencies had not opened their needed datasets. A significant gap remains between what the government agencies have opened and what the open data users expect to use. For instance, the interviewees said that some government agencies do not provide their datasets through open data approach. Instead, those agencies provide web information systems for the public to check governmental datasets online. Nevertheless, those systems need human operations to input search criteria by hand, and only limited data records can be retrieved at each query.

In addition, some interviewees pointed out that their perceived usefulness of open data gradually decreases as they gain more experience using some of the published datasets. The interviewees also claimed they were not satisfied with the data quality. It is indicated that the current data quality of some datasets, such as accuracy, completeness, granularity, and timeliness, could not meet data users' expectations. For instance, some datasets might quickly become obsolete and contain erroneous data, while their update frequencies remain relatively low. The interviewees said that using the datasets, they could generate wrong data analysis results or provide inaccurate services to their clients. The interviewees also pointed out that they need raw rather than processed datasets. Some interviewees even argued that they would decide not to use governmental open data and turn to look for other alternatives by themselves. For instance, some interviewees stated the following:

...We have been thinking whether we can use xxx agency's open data to drive innovation and enhance our business operation. However, it is a pity that the current datasets opened by the agency are really trivial. Honestly, the datasets are not useful to fit our need.... (Chief Technology Officer, P24\_06)

...In my opinion, the data update frequency seriously matters. If the datasets are updated on a daily basis, the data can be very useful to fit our needs. Nevertheless, if it takes two or three months to update the datasets just once, well, I don't see we can benefit from using the datasets.... (Chief Executive Officer, P3\_03)

...Government agencies have opened many datasets, which cover a variety of different areas such as transportation, health, weather, and tourism. However, the depths of most of the datasets are relatively shallow and do not really fit our needs. What we need may be just a specific area of data; nevertheless, we care much more about the depth of data such as its details and comprehensiveness.... (Chief Executive Officer, P2\_02)

### **Perceived Effort**

The interviewees suggested perceived effort as another important factor that must be considered. They mentioned that they have to evaluate whether they can afford the needed time and resources for using open data. As the interviewees claimed, the process of data access and data cleaning can be time-consuming, making it challenging for them to obtain and use the datasets they need. Governmental open datasets could be scattered in different places rather than on a single portal, so the interviewees had to spend significant time on searching government websites to look for datasets. For instance, an interviewee (Chief Executive Officer, P2\_02) explained:

...We attempt to retrieve the datasets from the unified open data portal, data.gov.tw. However, some government agencies may just have their datasets published on other open data-related sites or just on somewhere of their respective agencies' websites. Therefore, we have to spend lots of effort on finding those datasets from different channels, and sometimes we even need to develop web crawlers to parse data from agencies' websites or their online databases.... (Chief Executive Officer, P2\_02)

In addition, the interviewees suggested that most of the datasets in the open data portals are static data in open formats, such as CSV, JSON, or XML, which meets the three-star requirement of the open data schema. They also mentioned encountering the problem of determining when the datasets may be updated. Further, they pointed out that having no notification mechanism is relatively inconvenient, requiring them to revisit the sites frequently for possible updates. This situation also poses a challenge for the interviewees in obtaining up-to-date datasets. For instance, an interviewee (General Manager, P5\_07) stated:

...We retrieve many different datasets from the open data sites. However, we have no idea when the datasets may be updated. It doesn't really make sense for us to designate someone to frequently check the sites to see whether new datasets are available. In the industry, we are more used to connecting API for real-time data exchange rather than a traditional file downloading.... (General Manager, P5\_07)

Furthermore, the interviewees pointed out that government agencies opened their datasets with inconsistent formats and fields, a tremendous challenge in developing business applications that require processing and integrating open data of different central and local government agencies with the same core businesses. Some interviewees explained:

...According to our experience in processing transportation datasets of different local government agencies, we notice that some datasets have more data fields and some have very limited data fields. What is worse is that two datasets may contain data fields having the same names but with totally different definitions. Another problem is that the datasets may adopt different character encodings. Therefore, humans checking to view through the datasets become inevitable, and this process is very labor intensive. It is difficult for us to clean and integrate datasets through automatic machine processing.... (Vice President, P23\_05)

...We try to enhance our tourism application by using governmental open data. However, the tourism datasets actually come from agencies of different local governments, and the datasets have different formats and fields. I would say those datasets are really fragmented and fairly difficult for us to integrate. We are frustrated, and it is unrealistic for us to apply the datasets to develop application functions that could be available only in some cities. What if our app users plan to travel across cities?.... (General Manager, P5\_07)

Moreover, some interviewees pointed out that even an agency may open a dataset that has inconsistent data representations among the data records. An interviewee (Product Manager, P7\_09) gave an example:

...Even within the same dataset, a data field may have different representations. For instance, in the address data field, some records may use traditional Chinese characters to represent floor and address numbers, and some records may use English characters such as 'f' or 'F' and Arabic numerals. The address field also contains different symbols, punctuation marks, and abbreviations, which can be very confusing. There is no authority control.... (Product Manager, P7\_09)

## **External Influence**

External influences can be conceptualized as the driver derived from the professional reusers' surrounding environment, emphasizing the importance of open data policy and encouraging open data usage to establish a sound open data ecosystem. The interviewees indicated that external influences could foster their

perception of the importance of open data and draw their intention toward using open data. In particular, it was determined that external influences could be in the forms of international open data trends, social groups, government promotions, and peer organizations. Some interviewees pointed out that they have been following the international movement of open government data, which initially raised their interest in considering using governmental open data. An interviewee (Chief Technology Officer, P24\_06) said:

...We have been observing the international trend of open data development. For instance, we notice that the U.S. has done a very good job in opening some datasets that are related to our business, which then becomes a driver that we start observing the related development in Taiwan and attempting to initiate the following adoption and use.... (Chief Technology Officer, P24\_06)

The interviewees also pointed out that in the past few years, Taiwan non-governmental organizations and social groups have been promoting the transparency of government information and the use of open data through civic activities such as hackathons that attract both IT professionals and non-IT individuals. This trend forms a positive atmosphere that encourages business and non-profit organizations to adopt open data for innovative usage and social good. While most interviewees are from the software and information technology industries, they indicated they are influenced by related non-profit organizations and social groups through shared ideology and expectation to consider using governmental open data. An interviewee (Director, P8\_10) explained:

...We are indeed influenced by social communities. Particularly, as software engineers, we do like the ideas of an open system, open source, and open data that promote the principles of transparency and sharing. When we need to look for datasets to use, open data then becomes one of the options that draw our attention, and we would take it into consideration.... (Director, P8\_10)

Similarly, the Taiwan government has designated agencies to promote open data use through various activities such as workshops, international forums, and open data contests. While the interviewees are within the information technology industry, they indicated that they could be influenced by government policies and would attempt to see whether they could meet the expectations of the policies. Accordingly, some interviewees admitted that they are influenced by the government in evaluating the feasibility of using governmental open data in their business operations. An interviewee (Chief Executive Officer, P9\_11) stated:

...I would say that the promotion and encouragement from the government raised our initial interests toward open data use. Then, we became interested in knowing open data further. Nevertheless, it still takes time to see whether open data can really fit into our business model.... (Chief Executive Officer, P9\_11)

Some interviewees suggested that they can also be influenced by their peer group. It is indicated that when they know that other individuals or companies in their industries have been utilizing open data to enhance operation and productivity or develop innovative applications, they tend to act more seriously in considering whether they should use open data further.

### **Facilitating Condition**

Furthermore, the interviewees claimed that it would have a positive influence if they received resources and facilitation from others in their surrounding environment to support their open data use. It is indicated that the obtained resources and facilitation can help reduce the interviewees' perceived cost of using governmental open data. Accordingly, they become motivated to set higher priorities for using open data in developing business applications. An interviewee (Vice President, P23\_05) described:

...Well, for a company, using governmental open data can be viewed as a kind of trial and investment, and there is a cost for using open data. Through the contest, we are fortunate to receive monetary reward and consultation from the government, which becomes a driver for us to consider using open data further.... (Vice President, P23\_05)

Interviewees also explained that they would like to use open data further in their business applications. However, while one of the core parts of business applications relies on rich data sources, the interviewees pointed out that they looked forward to obtaining more facilitation from the government in terms of interacting with government agencies to access more open data that fit their needs.

### **Legislation and License**

As professional reusers, the interviewees indicated that their major purpose in using open data is to develop business applications. Therefore, the interviewees asserted that they needed to carefully review whether related regulations and licenses were clear and appropriate to guide their open data usage in their respective business domains. An interviewee (Chief Executive Officer, P3\_03) explained:

...Practically, when we apply any third-party dataset to our business application, we have to study its license carefully. We need to figure out to



what extent we can legally use the dataset for business purposes without any violation of regulations.... (Chief Executive Officer, P3\_03)

Furthermore, the interviewees suggested that the Taiwan government has gradually developed its open data license by following the principles of Creative Commons 4.0. Noteworthy, the license has been widely adopted by both central and local government agencies in open data initiatives. However, some interviewees pointed out that the Taiwan government has not yet enacted specific open data law. Government agencies are encouraged but not necessarily required to open their datasets. Without specific open data law, the interviewees claimed that when they attempted to request more open data from certain government agencies, they encountered challenges resulting from the complex bureaucratic system and other existing legislations, such as the Personal Data Protection Act and the Charges and Fees Act. The existing legislations and regulations became barriers, hindering their intentions to use open data further. An interviewee (Product Manager, P7\_09) gave an example:

...We have been requesting a dataset regarding the information of the nationally registered physicians and the clinics and hospitals where the physicians currently practice. However, the responsible government agency keeps turning down our request and asserts that this dataset could violate Personal Data Protection Act. However, clinics and hospitals usually publicly publish the information of their affiliated physicians on websites. Patients also need to know the information when they want to make appointments. We are really puzzled and still can't obtain the dataset to develop our application.... (Product Manager, P7\_09)

### **Self-Efficacy**

The interviewees suggested that not everyone has the capability to access and use open data directly. In addition to the commonly seen open formats such as CSV, XML, and JSON, open data can consist of other specialized formats and structures, and domain expertise is usually needed to clean, integrate, and interpret datasets. Accordingly, the interviewees pointed out that in order to use governmental open data, they must possess sufficient abilities in terms of data analysis and software engineering in their respective domains. In particular, as professional reusers running business services, the interviewees maintained that they are proficient IT developers and maintain high-level domain knowledge in their respective fields, implying their confidence in possessing the required capability to obtain governmental open data for developing business applications. Some of the interviewees stated the following:

...Running a data processing company, we would say that domain expertise and technical skills all have very important influences on our company's open data adoption. It can be directly related to our confidence in using open data..... (Chief Technology Officer, P24\_06)

...Well, because of our domain knowledge, we know how to interpret the datasets and apply them to our software applications. Without sufficient domain knowledge, we will have difficulty using open datasets. Similarly, if we don't have coding ability, we will not be able to analyze the datasets and don't know how to use the datasets for value-added purposes.... (Vice President, P23\_05)

### **Perceived Risk**

As professional reusers, the interviewees indicated that it could pose a great risk if they solely rely on open data as the only one or the major data source to develop business applications. The interviewees pointed out that their current open data had quality issues. In particular, the interviewees mentioned that using the database could provide inaccurate services to their clients that might incur consumer complaints and damage their companies' reputations. Similarly, using the datasets to develop strategic business plans may lead to wrong analyses and predictions and result in potential operational losses. The interviewees described the phenomenon:

...There is a very high possibility that inaccurate data records exist in the open datasets that we adopt to develop applications. Our company could suffer loss and receive complaints from the application users. Our clients can become less confident in using our applications, and our intention to use open data further is negatively influenced.... (Chief Executive Officer, P2\_02)

...Government agencies may update their internal databases at any time to reflect their daily operations. However, some datasets on the open data portals are not frequently updated, and those datasets can become outdated. By using the datasets, our system could generate poor services and inaccurate results to our app users.... (Product Manager, P7\_09)

Although one of the common purposes of open data is to promote innovative applications expected to spur economic growth, some interviewees claimed that they act hesitantly and have serious concerns about applying open data to develop their core business applications. They mentioned that they were uncertain whether their adopted datasets could be continually opened by government agencies. The interviewees further argued that it is possible that some datasets might stop being

opened, which could bring fatal impacts on startups and businesses if they solely rely on using open data to develop their core products. In addition, they pointed out that the potential risk exists, which they are not willing to take. Some of the interviewees stated the following:

...I personally have many questions regarding a company using governmental open data in its core business model even though the data could be used in a very innovative approach. I think it is extremely risky to do so. Open data can be applied to develop some added features of an application but should not become the core part of the application. If the data is stopped being opened, your business will be severely impacted.... (Director, P9\_11)

...Honestly, I am not sure whether the datasets can be continually opened and frequently updated. It will be a serious issue if the datasets are no longer opened. You can use open data as a supplement to your application. However, you should not use open data to develop the core part of your business model. Your business could be in jeopardy if the datasets become not opened.... (Chief Executive Officer, P3\_03)

## **Discussion**

### **The Identified Factors and Their Influences**

With empirical data support, seven factors were identified through the interviews with professional reusers. Perceived usefulness and perceived effort were found to have the most important influences when professional reusers evaluate whether they would adopt and use governmental open data. By using governmental open data, professional reusers expect to increase the efficiency and reduce the cost of developing business services and applications. Particularly, some open data are derived from the domains that private companies are not permitted to or are not capable of collecting by themselves. Therefore, professional reusers tend to have high expectations toward using governmental open data and expect to use the datasets to drive and develop new business models or enhance their existing business services. Still, a significant gap remains between what government agencies have opened and what professional reusers look for. It was found that professional reusers' perceived usefulness toward open data use could gradually decrease. At the same time, they remain unsatisfied with the quality of the retrieved datasets in terms of accuracy, completeness, granularity, and timeliness, which are critical data criteria that professional reusers must obtain to provide applications of high business value.

Similarly, perceived effort weighs when professional reusers find that they indeed have to spend considerable time and resources on the processes of data access, data cleaning, and data integration, which seriously contradicts their original expectation of increasing efficiency and reducing operational costs using governmental open data. As identified in this study, different central and local government agencies of the same core business usually open datasets with inconsistent data formats, data fields, and metadata information, which poses a huge challenge for professional reusers in integrating the published datasets of various government agencies for developing business applications. This issue in data inconsistency can greatly increase professional reusers' perceived efforts of using governmental open data. However, automatic and machine-operated data processing becomes very difficult to practice.

Furthermore, it was also found that professional reusers are still evaluating the long-term feasibility of using governmental open data in their business services. They are also concerned about its potential risks, which can further hinder the government's expectation of unleashing the potential of governmental datasets to spur business and economic growth. In the short term, professional reusers are afraid of creating inaccurate and flawed services resulting from data quality issues, which may negatively impact their companies' reputations. In the long term, professional reusers expect to use governmental open data to develop applications for long-lasting business services rather than just for a one-time side project. Therefore, professional reusers act hesitantly and are concerned about applying governmental open data to the core part of their business models and applications, as they are uncertain whether their current datasets can be continually opened in the foreseeable future.

Due to professional reusers' surrounding environment, external influences were also found to come from international open data trends, social groups, government promotions, and peer organizations to influence professional reusers' intention toward open data use. Nevertheless, external influences tend to act as catalysts for professional reusers' initial interest and intention toward using governmental open data. The strength of external influences can gradually decrease as professional reusers begin using the datasets. Then, their perceived usefulness, perceived effort, and perceived risk would weigh more.

In addition, facilitation from professional reusers' surrounding environment can help reduce costs and motivate their open data use. In particular, for professional reusers running startup companies, monetary reward and consultation from the government and experience-sharing from other organizations can encourage them to use governmental open data. Professional reusers expect more facilitation from government agencies to help them obtain more open datasets that meet their needs in developing business applications.

Furthermore, professional reusers also need to verify whether their open data usage complies with related legislations and licenses because their major purpose of using open data is to develop commercial applications. However, this study found that a complex bureaucratic system and existing legislations and regulations can sometimes become barriers that restrain professional reusers from requesting more open data from government agencies.

Moreover, professional reusers were found to maintain a high level of self-efficacy. In particular, they showed confidence in possessing sufficient domain expertise and technical skills for using governmental open data. This result may contradict the argument in some open data studies that technical skills and expertise should be available to facilitate and promote open data use. In fact, the general public represents end users, who usually do not directly utilize governmental open data. End users usually rely on the services and applications provided by professional reusers, who are the direct reusers and have the sufficient capability to use governmental open data. Therefore, for professional reusers, it is still more important to focus on enhancing their perceived usefulness and alleviating their perceived effort and perceived risk to promote the use of governmental open data.

### **The Implications and an Integrated View of the Factors**

Among the identified factors, perceived usefulness, perceived effort, external influences, and facilitating conditions correspond to the concepts of the UTAUT, a technology acceptance model. Particularly, according to the qualitative empirical data, the two most commonly discussed concepts of the technology acceptance model—perceived usefulness and perceived effort—were found to have the most important influences on professional reusers' intention to use governmental open data. Therefore, exploring the antecedent factors that may influence perceived usefulness and perceived effort is worth pursuing. As discussed in the literature review, professional reusers' use of governmental open data can be conceptualized as a process of innovation acceptance by adapting to the open data infrastructure. Accordingly, the quality aspects, such as data quality and system quality, of the utilized open data portals and related information systems can be considered the potential antecedent factors influencing professional reusers' perceived usefulness and perceived effort. The quality factors are also expected to affect professional reusers' perceived risk of using governmental open data—another factor identified in this study but not originally discussed in the technology acceptance model.

In addition, according to the empirical data analysis, professional reusers' perceived usefulness of using open data tends to decrease after using the obtained datasets. External influences also tend to act as an initial catalyst, and the strength of the factor gradually decreases after professional reusers use the datasets.

Therefore, it is assumed that the influences of the identified factors are expected to vary as the time of professional reusers' open data usage proceeds and the related using experience accumulates. Accordingly, a longitudinal study can be another potential research strategy that can be applied to study the phenomenon further.

Similarly, as discussed in the literature, other types of open data users still exist, such as social reusers and end users (Abella et al., 2019). It is also possible that the identified factors can have different influences on the other types of users. It is assumed that open data users—impelled by different motivations, such as exploring creativity, creating business value, enabling local citizen value, addressing global societal challenges, and advocating the open data agenda as described by Lassinantti et al. (2019)—may assign different weights to the identified influential factors. For instance, among the different types of open data users, professional reusers may tend to maintain higher self-efficacy toward their open data use than other users, such as social reusers and end users.

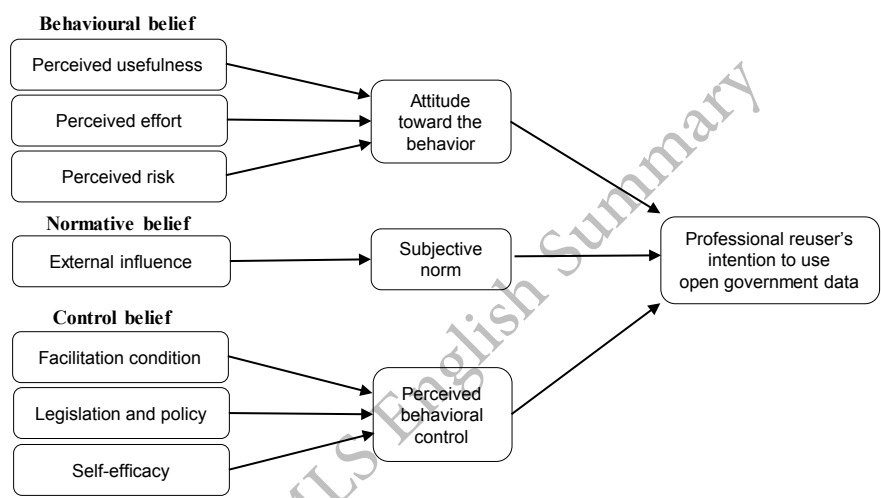
From an integrated view, the identified factors can be further incorporated with Ajzen's (1991) theory of planned behavior to represent behavioral beliefs, normative beliefs, and control beliefs that determine professional reusers' attitudes, subjective norms, and perceived behavioral control toward governmental open data use (see Figure 1). According to the theory of planned behavior, professional reusers' attitudes can be determined by different behavioral beliefs, which represent the subjective probabilities that given outcomes can be produced by the behavior. In this study, perceived usefulness, perceived effort, and perceived risk are conceptualized as the outcomes resulting from professional reusers' engagement in using governmental open data. The three factors can represent either positive or negative outcomes of performing the behavior. Accordingly, the three factors can act as professional reusers' behavioral beliefs in determining their attitude toward open data use.

In addition, subjective norms are determined by normative beliefs, which refer to the perceived behavioral expectations from other individuals, groups, or organizations. In this study, the identified external influences result from social groups and the government. As the interviewees indicated that they could be influenced by social groups' ideology and the government's open data policies, they would consider meeting their expectations. Therefore, external influences from social groups and the government can be related to professional reusers' perceived behavioral expectations and act as the normative belief that determines their subjective norms toward open data use.

Furthermore, perceived behavioral control is determined by different control beliefs, which refer to the factors that may either facilitate or hinder the behavior. In this study, facilitating conditions, existing legislations and regulations, and self-

efficacy can either benefit or impede professional reusers’ use of governmental open data. Therefore, the three factors are assumed to act as professional reusers’ control beliefs determining their perceived behavioral control over open data adoption. As shown in Figure 1, incorporating the theory of planned behavior, a synthesized view is presented to see how the identified factors act as behavioral, normative, and control beliefs that influence professional reusers’ intention toward using governmental open data. This synthesized view of the identified factors is expected to provide a foundation for later survey-based quantitative research.

**Figure 1. A Synthesized View Integrated with the Theory of Planned Behavior**



Moreover, the findings suggest that there is still a significant gap between what government agencies have opened and professional reusers’ expectations. Apparently, there are blurred areas that must be addressed further. Meanwhile, ambiguities and conflicts might exist between the concept of an open government, hindering the publication of more datasets to achieve transparency, participation and collaboration, and the concept of traditional stewardship in which government agencies are required to govern their datasets strictly. It is indicated in the literature that governmental open data is defined as non-privacy-restricted and non-confidential data, generated using public money and made available for the public to access without restrictions (Janssen et al., 2012). However, when different open datasets are merged to provide integrated information, information could infringe privacy and confidentiality (Yang et al., 2015). Researchers also suggest that it is critical to find a balance and resolve the potential contradictions between open government data policies and other public values, such as trust, transparency, privacy, and security (Meijer et al., 2014). Similarly, as Dawes (2010) has pointed out, stewardship and usefulness should be two broad and



complementary information principles. The principle of stewardship protects government information from damage, loss, or misuse and is concerned with the assurance of responsibility, validity, and legitimacy. On the other hand, the principle of usefulness encourages exploration, application, and innovation. It makes government information more accessible and easier to obtain and be used by various public and private users. It is suggested that the two principles should be considered to help balance the many considerations critical for achieving greater government transparency and realizing the public value of government information (Dawes, 2010).

### **The Practical Implications of This Study**

According to the empirical findings, government agencies may have dedicated efforts to open datasets. However, many of the opened datasets do not meet professional reusers' needs for developing business applications, and many datasets professional reusers expect to use are still not opened by government agencies. Therefore, there is a need to bridge the gap by enhancing professional reusers' perceived usefulness of using governmental open data. The following list presents the practical implications of this research:

- Government agencies should maintain appropriate communication channels for obtaining comments and feedback from professional reusers regarding their data usage experiences and expectations. For instance, government agencies of different business domains should hold public workshops, meetings, and forums to provide professional reusers with the opportunities to interact with the agencies and express their data needs. Through this approach, government agencies can also benefit from having a clearer direction in terms of what datasets to open with higher priority.
- According to the Pareto principle (the 80/20 rule; Koch, 1999), 20% of the opened datasets may represent those most professional reusers are interested in using. Therefore, government agencies can investigate what datasets are most frequently downloaded and utilized by open data users. Then, government agencies can explore whether those datasets meet data users' expectations and whether more related datasets can be opened further.
- There is a need to reduce professional reusers' perceived effort of using governmental open data. Government agencies should be continually encouraged to use the unified open data portal to publish their open data information, which can help simplify professional reusers' processes of data search and data access. In addition, a vertical coordination mechanism can be established among the central and local government agencies with the same core businesses. For each business domain, a responsible central government

agency can coordinate to create a standardized template that other government agencies may adopt in opening their related datasets. Through standardization, the barriers to open data use resulting from inconsistent data formats, data fields, and metadata information can be reduced when professional reusers use open datasets of different central and local government agencies with the same core businesses. Similarly, for some central government agencies that could open similar datasets, a horizontal coordination mechanism may be enabled to set up a unified window for opening datasets. Consequently, professional reusers' uncertainty regarding which datasets they should use can be decreased, while those similar datasets usually come with inconsistent data content and have different update frequencies.

- When opening their datasets, many government agencies merely put the links of their published datasets on open data platforms for users to download. However, this approach is still inconvenient for professional reusers because a manual operation is needed to download data files. This approach also poses difficulty in obtaining up-to-date data. Thus, it is preferred that government agencies consider using open API to open their datasets. In this way, professional reusers can automate the process of connecting their information systems to open API and assign customized query parameters to retrieve up-to-date data, which can help professional reusers create high business value applications. It is expected that the approach of an open API can greatly increase professional reusers' perceived usefulness and reduce their perceived effort in using governmental open data.
- Government agencies should continue improving the quality of their published datasets, which is critical to enhancing professional reusers' perceived usefulness of open data use. A collaborative data feedback model can also be established, which allows professional reusers to collaborate with government agencies to improve open data quality. In some data domains, professional reusers may be willing to clean, refine and supplement their retrieved open data to generate more comprehensive datasets with better data quality in terms of accuracy, completeness, and timeliness. The data feedback model allows professional reusers to provide enhanced datasets back to government agencies with some open licenses or subscription fees. Meanwhile, government agencies can benefit from obtaining better quality datasets and re-releasing the datasets to the general public. Further, this approach can encourage private sectors, such as entrepreneurs and startups, to participate in the process of open data preparation and refinement through the existing infrastructure and foundation of open government data.

- Another similar approach suggests that the government should facilitate the formation of a data market while professional reusers act as the role of data intermediary in their respective business domains, as some literature has also suggested (Mercado-Lara & Gil-Garcia, 2014; Schrock & Shaffer, 2017; Van Schalkwyk et al., 2016). Based on the foundation of open government data, professional reusers can directly provide their refined datasets to the general public and other businesses as a free public service or commercial service with fees. Through this approach, professional reusers acting as data intermediaries can help improve the quality of governmental open data with added values and make it easier for others to use open data.

While data intermediaries are capable of providing data services with enhanced data quality and guaranteed sustainability based on the existing infrastructure of open government data, this approach could be an alternative for direct data reusers who are willing to subscribe to the service, particularly for those who tend to have high perceived risks of directly using governmental open data. Data intermediaries can dedicate their efforts to interacting with the data reusers to assist their data usage requirements and problems. This strategy also helps reduce the loading that government agencies might have to deal with. Particularly, in addition to the current free open data service, it is also worth exploring whether government agencies can consider offering service contracts with reasonable fee charges as an alternative for professional reusers seeking government agencies to provide dedicated data services such as large volume and specialized data access.

- This study found that external influences and facilitation conditions matter. Thus, the government should continue promoting open data use through various activities, including workshops, forums, hackathons, and contests, which can spur professional reusers' initial interest in using governmental open data. Through the activities, government agencies can also provide more facilitation to professional reusers such as entrepreneurs and startups to assist their open data use. Furthermore, as aforementioned, while professional reusers usually maintain a high level of self-efficacy in their domain knowledge and technical skills, a well-established mechanism that professional reusers can interact with the agencies to express their feedback and data needs is the "true facilitation" that professional reusers look forward to receiving.
- In terms of legislation and policy, a designated open data law should be enacted. Such a law not only provides clearer guidance and acts as the foundation for government agencies to implement open data policies but also helps professional reusers reduce their perceived risk and retain their confidence

in using governmental open data concerning its long-term sustainability and that government agencies will commit themselves to continually open datasets. Enacting open data law can also help professional reusers counter the institutional barriers resulting from a complex bureaucratic system and other existing legislations and regulations when attempting to interact with government agencies to request more open datasets.

- Public and private collaboration is another approach that government agencies can consider to encourage professional reusers to use open data (Susha et al., 2017, 2019). Specifically, government agencies can invite professional reusers to help resolve the pressing challenges using governmental open data. Government agencies can also collaborate with professional reusers to open more datasets during the process. In this way, professional reusers can have the opportunity to develop solutions that meet the needs of government agencies. If government agencies are satisfied with the developed solutions, this public and private collaboration can also present an opportunity, allowing professional reusers to later market the developed solutions as business services to other government agencies or companies in the industry.
- While governmental open data should not include datasets that may infringe personal privacy and national security, another approach to bridging the gap between what government agencies have opened and what professional reusers expect to obtain is the help from open data committees, which have been set up in most central government agencies in Taiwan. The composition of open data committees can include representatives of government agencies, private and public sectors, social groups, academia, and individual citizens. The committees are expected to help determine whether some governmental datasets that professional reusers request should be opened. Therefore, the designated function of open data committees in respective government agencies should be well exercised in considering the principles of stewardship and usefulness—the two aforementioned complementary information principles proposed by Dawes (2010).

## **Conclusion**

A sound open government data ecosystem should include both open data providers and users. Meanwhile, this research explores and discusses the factors determining professional reusers' intention to use governmental open data. With qualitative empirical data support, the identified factors are perceived usefulness, perceived effort, external influence, facilitating condition, legislation and license, self-efficacy, and perceived risk. Particularly, perceived usefulness and

perceived effort are the two major factors derived from professional reusers' data usage experience. Professional reusers also consider the perceived risk of using governmental open data for developing business applications in terms of long-term data sustainability. As proficient IT developers in their respective domains, professional reusers maintain a high level of self-efficacy toward open data use. However, external influences and facilitation conditions from their surrounding environment must also be taken into consideration. Similarly, existing legislations and regulations within the bureaucratic system affect professional reusers when attempting to interact with government agencies to request more open data. In addition, the identified factors are integrated with the theory of planned behavior to present how the factors determine professional reusers' intention of using governmental open data through behavioral, normative, and control beliefs. Furthermore, the discussion and implications of this study can provide insights to researchers, practitioners, and policymakers regarding potential research directions and how open data policies can be further developed and implemented to attract open data users. This study conducted in the context of Taiwan's open government data is expected to enrich the current open data-related literature.

However, there are limitations in the current research. This study uses a qualitative approach, in which potential research biases could occur while some parts of the interview data may be emphasized more and some could be neglected. Similarly, this research uses a purposive sampling approach to recruit relevant interviewees for conducting interviews, through which potential sampling bias might also occur. Accordingly, there should be more research for exploring the complexity of open data use in different data areas, social contexts, and countries. In addition, as aforementioned, future research can explore the factors influencing different types of open data users, such as social reusers and end users. Furthermore, it would be interesting to investigate whether the factors may have different levels of impact on different types of open data users. Moreover, quantitative research can also be employed to evaluate the strengths of the factors identified in this study.

### **Acknowledgements**

The authors are grateful to the editors and the three anonymous reviewers for providing thoughtful comments to help the authors revise and improve the paper. The authors also thank all the interviewees for participating in this research and offering valuable qualitative empirical data to help the authors finish this research project. This study was financially supported by the Ministry of Science and Technology, Taiwan (107-2410-H-002-165-MY2) and by the Universities and Colleges Humanities and Social Sciences Benchmarking Project (Grant no. 111L9A002).

## References

- Abella, A., Ortiz-de-Urbina-Criado, M., & De-Pablos-Heredero, C. (2019). The process of open data publication and reuse. *Journal of the Association for Information Science and Technology*, 70(3), 296-300. <https://doi.org/10.1002/asi.24116>
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Attard, J., Orlandi, F., & Auer, S. (2016). Value creation on open government data. In T. X. Bui & R. H. Sprague, Jr. (Eds.), *Proceedings of the 49th annual Hawaii international conference on system sciences* (pp. 2605-2614). IEEE. <https://doi.org/10.1109/HICSS.2016.326>
- Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4), 399-418. <https://doi.org/10.1016/j.giq.2015.07.006>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Barry, E., & Bannister, F. (2014). Barriers to open data release: A view from the top. *Information Polity*, 19(1), 129-152. <https://doi.org/10.3233/IP-140327>
- Bates, J. (2014). The strategic importance of information policy for the contemporary neoliberal state: The case of Open Government Data in the United Kingdom. *Government Information Quarterly*, 31(3), 388-395. <https://doi.org/10.1016/j.giq.2014.02.009>
- Bryman, A. (2004). *Social research methods* (2nd ed.). Oxford University Press.
- Chen, H.-L. (2015). *Zhengfu kaifang ziliao pinzhi yu manyidu zhi yanjiu* [The analysis of user satisfaction for open government data quality] [Unpublished master's thesis]. Tamkang University. <https://doi.org/10.6846/TKU.2015.00169>
- Chu, P.-Y., & Tseng, H.-L. (2016). A theoretical framework for evaluating government open data platform. In *Proceedings of the international conference on electronic governance and open society* (pp. 135-142). ACM. <https://doi.org/10.1145/3014087.3014099>
- Conradie, P., & Choenni, S. (2014). On the barriers for local government releasing open data. *Government Information Quarterly*, 31(S1), S10-S17. <https://doi.org/10.1016/j.giq.2014.01.003>
- Dawes, S. S. (2010). Stewardship and usefulness: Policy principles for information-based transparency. *Government Information Quarterly*, 27(4), 377-383. <https://doi.org/10.1016/j.giq.2010.07.001>
- Dawes, S. S., Vidasova, L., & Parkhimovich, O. (2016). Planning and designing open government data programs: An ecosystem approach. *Government Information Quarterly*, 33(1), 15-27. <https://doi.org/10.1016/j.giq.2016.01.003>
- Digital Government. (n.d.). *Introduction*. <https://obamawhitehouse.archives.gov/sites/default/files/omb/egov/digital-government/digital-government.html>
- Gerunov, A. (2017). Understanding open data policy: Evidence from Bulgaria. *International Journal of Public Administration*, 40(8), 649-657. <https://doi.org/10.1080/01900692.2016.1186178>

- H.R.4174 –115th Congress (2017-2018): Foundations for Evidence-Based Policymaking Act of 2018. (2019, January 14). <https://www.congress.gov/bill/115th-congress/house-bill/4174>
- Hellberg, A.-S., & Hedström, K. (2015). The story of the sixth myth of open data and open government. *Transforming Government: People, Process and Policy*, 9(1), 35-51. <https://doi.org/10.1108/TG-04-2014-0013>
- Hivon, J., & Titah, R. (2017). Conceptualizing citizen participation in open data use at the city level. *Transforming Government: People, Process and Policy*, 11(1), 99-118. <https://doi.org/10.1108/TG-12-2015-0053>
- Hossain, M. A., Dwivedi, Y. K., & Rana, N. P. (2016). State-of-the-art in open data research: Insights from existing literature and a research agenda. *Journal of Organizational Computing and Electronic Commerce*, 26(1-2), 14-40. <https://doi.org/10.1080/10919392.2015.1124007>
- Hung, S.-Y., Chang, C.-M., & Kuo, S.-R. (2013). User acceptance of mobile e-government services: An empirical study. *Government Information Quarterly*, 30(1), 33-44. <https://doi.org/10.1016/j.giq.2012.07.008>
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, 29(4), 258-268. <https://doi.org/10.1080/10580530.2012.716740>
- Janssen, M., & Zuiderwijk, A. (2014). Infomediary business models for connecting open data providers and users. *Social Science Computer Review*, 32(5), 694-711. <https://doi.org/10.1177/0894439314525902>
- Kaasenbrood, M., Zuiderwijk, A., Janssen, M., de Jong, M., & Bharosa, N. (2015). Exploring the factors influencing the adoption of open government data by private organisations. *International Journal of Public Administration in the Digital Age*, 2(2), 75-92. <https://doi.org/10.4018/ijpada.2015040105>
- Koch, R. (1999). *The 80/20 principle: The secret of achieving more with less*. Currency.
- Kubler, S., Robert, J., Neumaier, S., Umbrich, J., & Le Traon, Y. (2018). Comparison of metadata quality in open data portals using the analytic hierarchy process. *Government Information Quarterly*, 35(1), 13-29. <https://doi.org/10.1016/j.giq.2017.11.003>
- Lassinantti, J., Stahlbrost, A., & Runardotter, M. (2019). Relevant social groups for open data use and engagement. *Government Information Quarterly*, 36(1), 98-111. <https://doi.org/10.1016/j.giq.2018.11.001>
- Lo, J., Yang, T.-M., Wang, H.-J., & Shiang, J. (2014). Zhengfu kaifang ziliao de celue yu tiaozhan: Shiyongzhe guandian de fenxi [Strategies and challenges of open government data: Analysis of user perspectives]. *Electronic Commerce Studies*, 12(3), 283-300.
- Magalhães, G., & Roseira, C. (2016). Exploring the barriers in the commercial use of open government data. In J. Bertot, E. Estevez, & S. Mellouli (Eds.), *ICEGOV '15-16: Proceedings of the 9th international conference on theory and practice of electronic governance* (pp. 211-214). ACM. <https://doi.org/10.1145/2910019.2910078>
- Meijer, R., Conradie, P., & Choenni, S. (2014). Reconciling contradictions of open data regarding transparency, privacy, security and trust. *Journal of Theoretical and Applied Electronic Commerce Research*, 9(3), 32-44. <https://doi.org/10.4067/S0718-18762014000300004>



- Mercado-Lara, E., & Gil-Garcia, J. R. (2014). Open government and data intermediaries: The case of AidData. In G. Puron-Cid & S. Robertson (Chairs), *dg.o '14: Proceedings of the 15th annual international conference on digital government research* (pp. 335-336). ACM. <https://doi.org/10.1145/2612733.2612789>
- Najafabadi, M., & Luna-Reyes, L. (2017). Open government data ecosystems: A closed-loop perspective. In T. Bui (Chair), *Proceedings of the 50th Hawaii international conference on system sciences*. IEEE. <https://doi.org/10.24251/HICSS.2017.327>
- Nugroho, R. P., Zuiderwijk, A., Janssen, M., & de Jong, M. (2015). A comparison of national open data policies: Lessons learned. *Transforming Government: People, Process and Policy*, 9(3), 286-308. <https://doi.org/10.1108/TG-03-2014-0008>
- Open Data Charter. (2021). *Government adopters*. <https://opendatacharter.net/government-adopters/>
- Ozkan, S., & Kanat, I. E. (2011). e-Government adoption model based on theory of planned behavior: Empirical validation. *Government Information Quarterly*, 28(4), 503-513. <https://doi.org/10.1016/j.giq.2010.10.007>
- Peled, A. (2011). When transparency and collaboration collide: The USA open data program. *Journal of the American Society for Information Science and Technology*, 62(11), 2085-2094. <https://doi.org/10.1002/asi.21622>
- Purwanto, A., Zuiderwijk, A., & Janssen, M. (2020). Citizen engagement with open government data: Lessons learned from Indonesia's presidential election. *Transforming Government: People, Process and Policy*. <https://doi.org/10.1108/TG-06-2019-0051>
- Rana, N., Dwivedi, Y., & Lal, B. (2015). Factors influencing citizen's adoption of an E-Government System: Validation of the decomposed theory of planned behavior. In *Proceedings of the UK academy for information systems conference 2015*. The UK Academy for Information Systems. <https://aisel.aisnet.org/ukais2015/14>
- Rana, N., Lal, B., & Slade, E. (2016). Adoption of two Indian e-government systems: Validation of extended theory of planned behavior (TPB). In *Proceedings of the Americas' conference on information systems 2016*. <https://aisel.aisnet.org/amcis2016/eGov/Presentations/7>
- Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity*, 22(1), 1-24. <https://doi.org/10.3233/IP-160012>
- Schrock, A., & Shaffer, G. (2017). Data ideologies of an interested public: A study of grassroots open government data intermediaries. *Big Data & Society*, 4(1). <https://doi.org/10.1177/2053951717690750>
- Solar, M., Concha, G., & Meijueiro, L. (2012). A model to assess open government data in public agencies. In H. Scholl, M. Janssen, M. Wimmer, C. Moe, & L. Flak (Eds.), *Electronic government* (Vol. 7443, pp. 210-221). Springer. [https://doi.org/10.1007/978-3-642-33489-4\\_18](https://doi.org/10.1007/978-3-642-33489-4_18)
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Sage.
- Susha, I., Grönlund, Å., & Janssen, M. (2015). Driving factors of service innovation using open government data: An exploratory study of entrepreneurs in two countries. *Information Polity*, 20(1), 19-34. <https://doi.org/10.3233/IP-150353>

- Susha, I., Grönlund, Å., & Van Tulder, R. (2019). Data driven social partnerships: Exploring an emergent trend in search of research challenges and questions. *Government Information Quarterly*, 36(1), 112-128. <https://doi.org/10.1016/j.giq.2018.11.002>
- Susha, I., Janssen, M., & Verhulst, S. (2017). Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development In T. Bui (Chair), *Proceedings of the 50th Hawaii international conference on system sciences*. IEEE. <https://doi.org/10.24251/HICSS.2017.325>
- The World Bank. (n.d.). *Open data essentials*. <https://opendatatoolkit.worldbank.org/en/essentials.html>
- Thorsby, J., Stowers, G. N. L., Wolslegel, K., & Tumbuan, E. (2017). Understanding the content and features of open data portals in American cities. *Government Information Quarterly*, 34(1), 53-61. <https://doi.org/10.1016/j.giq.2016.07.001>
- UK Minister of State for the Cabinet Office. (2012). *Open data: Unleashing the potential*. GOV. UK. <https://www.gov.uk/government/publications/open-data-white-paper-unleashing-the-potential>
- United Nations Department of Economic and Social Affairs. (2020). *United Nations e-government survey 2020: Digital government in the decade of action for sustainable development*. [https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2020-Survey/2020%20UN%20E-Government%20Survey%20\(Full%20Report\).pdf](https://publicadministration.un.org/egovkb/Portals/egovkb/Documents/un/2020-Survey/2020%20UN%20E-Government%20Survey%20(Full%20Report).pdf)
- Van Schalkwyk, F., Willmers, M., & McNaughton, M. (2016). Viscous open data: The roles of intermediaries in an open data ecosystem. *Information Technology for Development*, 22(Sup1.), 68-83. <https://doi.org/10.1080/02681102.2015.1081868>
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>
- Wang, H.-J. (2020). Adoption of open government data: Perspectives of user innovators. *Information Research*, 25(1), paper 849. <http://InformationR.net/ir/25-1/paper849.html>
- Wang, H.-J., & Lo, J. (2013). Determinants of citizens' intent to use government websites in Taiwan. *Information Development*, 29(2), 123-137. <https://doi.org/10.1177/0266666912453835>
- Wang, H.-J., & Lo, J. (2019). Factors influencing the adoption of open government data at the firm level. *IEEE Transactions on Engineering Management*, 67(3), 670-682. <https://doi.org/10.1109/TEM.2019.2898107>
- Wang, V., & Shepherd, D. (2020). Exploring the extent of openness of open government data – A critique of open government datasets in the UK. *Government Information Quarterly*, 37(1), 1-10. <https://doi.org/10.1016/j.giq.2019.101405>
- Weerakkody, V., Irani, Z., Kapoor, K., Sivarajah, U., & Dwivedi, Y. K. (2017). Open data and its usability: An empirical view from the Citizen's perspective. *Information Systems Frontiers*, 19(2), 285-300. <https://doi.org/10.1007/s10796-016-9679-1>
- Wengraf, T. (2001). *Qualitative research interviewing: Biographic narrative and semi-structured methods*. Sage.

- World Wide Web Foundation. (2018). *Open data barometer* (4th ed.). The Open Data Barometer. [https://opendatabarometer.org/4thedition/?\\_year=2016&indicator=ODB](https://opendatabarometer.org/4thedition/?_year=2016&indicator=ODB)
- Yang, T.-M., Lo, J., & Shiang, J. (2015). To open or not to open? Determinants of open government data. *Journal of Information Science*, 41(5), 596-612. <https://doi.org/10.1177/01655515155586715>
- Yang, T.-M., & Wu, Y.-J. (2019). Taiwan zhengfu kaifang ziliao tuixing zhi jinkuang diaocha yu tantao [The maturity assessment of the recent open data development in the context of Taiwan e-Government]. *Journal of Educational Media & Library Sciences*, 56(1), 7-44. [https://doi.org/10.6120/JoEMLS.201903\\_56\(1\).0041.RS.BM](https://doi.org/10.6120/JoEMLS.201903_56(1).0041.RS.BM)
- Zeleti, F. A., Ojo, A., & Curry, E. (2014). Emerging business models for the open data industry: Characterization and analysis. In G. Puron-Cid & S. Robertson (Chairs), *dg.o '14: Proceedings of the 15th annual international conference on digital government research* (pp. 215-226). ACM. <https://doi.org/10.1145/2612733.2612745>
- Zhu, X., & Freeman, M. A. (2019). An evaluation of U.S. municipal open data portals: A user interaction framework. *Journal of the Association for Information Science and Technology*, 70(1), 27-37. <https://doi.org/10.1002/asi.24081>
- Zuiderwijk, A., & Hinnant, C. C. (2019). Open data policy-making: A review of the state-of-the-art and an emerging research agenda. *Information Polity*, 24(2), 117-129. <https://doi.org/10.3233/IP-190160>
- Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, 31(1), 17-29. <https://doi.org/10.1016/j.giq.2013.04.003>
- Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012). Socio-technical impediments of open data. *Electronic Journal of e-Government*, 10(2), 156-172.
- Zuiderwijk, A., Janssen, M., & Dwivedi, Y. K. (2015). Acceptance and use predictors of open data technologies: Drawing upon the unified theory of acceptance and use of technology. *Government Information Quarterly*, 32(4), 429-440. <https://doi.org/10.1016/j.giq.2015.09.005>
- Zuiderwijk, A., Janssen, M., & Susha, I. (2016). Improving the speed and ease of open data use through metadata, interaction mechanisms, and quality indicators. *Journal of Organizational Computing and Electronic Commerce*, 26(1-2), 116-146. <https://doi.org/10.1080/10919392.2015.1125180>



# Effects of Afterschool Programs in Public Libraries on Disadvantaged Children: The Case of the New Taipei City Library<sup>ψ</sup>

Ya-Wei Chuang<sup>a</sup> Hui-Yun Sung<sup>b\*</sup>

## Abstract

*This study aims to investigate effects of afterschool programs in public libraries on disadvantaged children in Taiwan, which serves to demonstrate the value of public libraries for supporting social justice. This study was qualitative in nature. Interviews and observations were employed for data collection. Study participants included: 22 children, 13 parents, eight volunteers, and seven librarians. Findings of this study show that the reasons for parents allowing their children to participate in the program were a lack of time to take care of their children, an inadequate home learning environment, their children requiring assistance needed for to complete their schoolwork, and the meal voucher incentives. In addition, the main activities the children participated in in the public library afterschool programs were finishing their homework and independent reading; other activities, such as reading together, receiving help with their homework, art projects, and games, were arranged by volunteers. Finally, the effects of these programs were improvements in relation to reading habits, reading interests, academic progress, learning vision, interpersonal relationships, etiquette, and mental health.*

**Keywords:** Public libraries, Afterschool programs, Disadvantaged children, Qualitative research

## SUMMARY

### Introduction

Afterschool programs assist parents in caring for and teaching children,

<sup>ψ</sup> This article was adapted from a master's thesis under Ya-Wei Chuang "Effects of Afterschool Programs in Public Libraries on Disadvantaged Children: A Case of New Taipei City Library".

<sup>a</sup> Master Student, Graduate Institute of Library and Information Science, National Chung Hsing University, Taichung, Taiwan

<sup>b</sup> Associate Professor of Graduate Institute of Library and Information Science, and Director of International PhD Program in Taiwan and Transcultural Studies, National Chung Hsing University, Taichung, Taiwan

\* To whom all correspondence should be addressed. E-mail: [hsung@dragon.nchu.edu.tw](mailto:hsung@dragon.nchu.edu.tw)

alleviate children safety concerns after school, and support children's physical and mental development. Time arrangement after school is associated with children's growth. The informal learning process is one of the influential factors in creating children's achievement gap, apart from the formal school learning (Hartman, 2011). Nevertheless, insufficient resources among disadvantaged families make it difficult for some children to receive additional learning opportunities. These families are consequently more dependent on support from social institutions.

Social justice is a core public library mission (Pateman & Vincent, 2010). Public libraries provide free and equal access to afterschool programs for children who need the services to support their learning and development process. For children in comparatively disadvantaged economic and social areas, or those from schools and families which lack learning related resources, afterschool programs are particularly necessary with a view to compensating for the gap in education resources (Bevin & Goulding, 1999; Bundy, 2006). Previous studies have also shown how public libraries' afterschool programs could positively influence children (Bailey, 1999; Huffman & Rua, 2008; Train & Elkin, 2000).

New Taipei City Library is the first public library in Taiwan that incorporates afterschool programs to its regular services (S.-W. Huang, 2014), where services for disadvantaged primary school children are prioritized. Afterschool programs are executed by volunteers to provide homework help as well as reading companionship for children; diverse activities are designed and carried out during winter and summer vacations. Other than that, meal vouchers are distributed to disadvantaged children upon fulfilling the required time of participation. This study takes New Taipei City Library as an example to explore the effects of afterschool programs in public libraries on disadvantaged children.

## **Methodology**

This was a qualitative study, where interviews served as a primary data collection method and observations as a subsidiary one. This study chose seven branch libraries of New Taipei City Library as the research settings based on their nature in representation, scale, and time length of offering afterschool programs. Study participants involved 22 children from disadvantaged families, 13 parents from disadvantaged families, eight volunteers, and seven librarians who participated in the afterschool program. This study adopted thematic analysis for analyzing the data gathered, with a focus on reasons why parents let children participate in the program, what activities children participated in, as well as what effects the programs had on children involved.

## **Results and Discussion**

The result has shown that the reasons for parents allowing their children

to participate in the program included: lack of time to accompany children, inadequate learning environment at home, assistance needed for completion of schoolwork, and the incentives of meal vouchers. Different from previous research, this study found that meal vouchers served as a strong incentive for parents to let children participate in the program.

Main activities in the public library afterschool programs in which children participated were finishing homework and independent reading; other activities were arranged by volunteers, such as reading together, homework help, artwork, and games. The result has shown the effects of such programs on children, including: cultivation of reading habits, development of reading interests, progress in academic performance, enhancement of learning vision, boost in interpersonal relationships, etiquette acquisition, as well as improvement in mental state.

Two of the most significant effects are of the reading and learning perspectives. The library environment has increased children's exposure to books. The regular reading together time has helped children gradually develop their reading habits, and they had experienced the fun in reading as they were able to choose books to their preference. This study has found that activities provided by public libraries were the only activities in which many disadvantaged children (especially those from families lacking in skills and channels of information acquisition) participate during winter and summer vacations. Children broadened their learning vision as they read and participated in various activities, in which they gained more new experiences and learned more diversely in numerous theme activities.

## Conclusion

This study explored the practical operation of afterschool programs in public libraries and the effects of such programs on children involved, in hope of filling the research gap in this field. The results derived from this study have demonstrated how public libraries fulfilled the social justice mission through executing afterschool programs. These programs have brought about positive effects in closing the information gap and learning gap of disadvantaged children in the following aspects: cultivation of reading habits, development of reading interests, progress in academic performance, enhancement of learning vision, boost in interpersonal relationships, etiquette acquisition, as well as improvement in mental state.

## ROMANIZED & TRANSLATED REFERENCES FOR ORIGINAL TEXT

于錢寧娜 (2005)。美國紐約市皇后區公共圖書館的兒童服務。臺北市立圖書館館訊，22(4)，96-104。https://doi.org/10.29773/BTPL.200506.0009【Yu, Linna Chien (2005). Children's service at Queens Borough Public Library. *Bulletin of the Taipei Public Library*, 22(4), 96-104. https://doi.org/10.29773/BTPL.200506.0009 (in Chinese)】



- 孔詢媚 (2019)。梧棲敬拜教會實施課後輔導對梧南國小弱勢學生之影響〔未出版之碩士論文〕。南華大學國際事務與企業學系亞太研究碩士班。【Kung, Hsun-Mei (2019). *The impact of the learning performance of Wuchi Worship Church's after-school program on the disadvantaged students of Wu-Nan Elementary School* [Unpublished master's thesis]. M. A. Program of Asia-Pacific Studies, Department of International Affairs and Business, Nanhua University. (in Chinese)】
- 王育敏、邱靖惠 (2009)。窮孩子·低社交？家庭貧窮對兒童社會參與之影響。社會發展季刊，124，72-85。【Wang, Yu-Min, & Qiu, Jing-Hui (2009). Qiong haizi, di shejiao? Jiating pinqiong dui ertong shehui canyu zhi yingxiang. *Community Development Journal* (Quarterly), 124, 72-85. (in Chinese)】
- 何俊青 (2014)。台灣課後照顧的研究現況：期刊與學位論文的內容分析。南台人文社會學報，12(11)，27-60。【Ho, Chun-Ching (2014). Current status of after-school child care in Taiwan: A content analysis of journal papers and theses/dissertations. *STUST Journal of Humanities and Social Sciences*, 12(11), 27-60. (in Chinese)】
- 吳文炎 (2015)。課業輔導對弱勢兒童社會流動的長期影響：以博幼基金會埔里中心為例〔未出版之碩士論文〕。靜宜大學社會工作與兒童少年福利學系。【Wu, Wen-Yen (2015). *Long-term effects of tutoring for after school learning on disadvantage children about social mobility: Puli Center of Boyo Social Welfare Foundation as an example* [Unpublished master's thesis]. Department of Social Work & Child Welfare, Providence University. (in Chinese)】
- 李介麟 (2006)。弱勢家庭學生的假期失落現象。師友月刊，470，21-25。【Li, Jie-Lin (2006). Ruoshi jiating xuesheng de jiaqi shiluo xianxiang. *The Educator Monthly*, 470, 21-25. (in Chinese)】
- 李孟峰、連廷嘉 (2010)。「攜手計畫－課後扶助方案」實施歷程與成效之研究。教育實踐與研究，23(1)，115-143。【Li, Meng-Feng, & Lien, Ting-Chia (2010). A study of implementation process and results of the "Hand-in-Hand Project: Education Support for Disadvantaged Child." *Journal of Educational Practice and Research*, 23(1), 115-143. <https://doi.org/10.6776/JEPR.201006.0115> (in Chinese)】
- 李宜玟 (2012)。數學低成就學習動機之類型與區別分析：中小學弱勢學生與一般學生之比較。教育科學研究期刊，57(4)，39-71。【Lee, Yi-Mei (2012). Discriminating math low-achievement motivation patterns: Comparing disadvantaged and other students in elementary and junior high school. *Journal of Research in Education Sciences*, 57(4), 39-71. (in Chinese)】
- 周倩如 (2006)。公共圖書館推廣活動融入國民小學課程初探。臺北市立圖書館館訊，23(4)，1-15。【Chou, Chien-Ju (2019). A study of integrating public libraries' promoting activities into the curriculum of elementary school. *Bulletin of the Taipei Public Library*, 23(4), 1-15. (in Chinese)】
- 林士翔 (2003)。國小低收入戶與非低收入戶學生成就動機與生活適應之相關研究〔未出版之碩士論文〕。國立屏東師範學院國民教育研究所。【Lin, Shin-Hsiang (2003). *The research for differences between elementary school students whose from low income families and not low income families on their academic achievement motivation and life adjustment* [Unpublished master's thesis]. Graduate Institute of Compulsory Education, National Pingtung Teachers College. (in Chinese)】



- 林俊瑩、邱欣怡、葉芝君(2012)。社經背景，幼年受管教經驗與教養價值觀對學前教育階段家長管教子女方式的影響。中華輔導與諮商學報，32，123-149。https://doi.org/10.7082/CJGC.201204.0121【Lin, Chunn-Ying, Chiu, Hsin-Yi, & Yeh, Chih-Chun (2012). Impact of socio-economic backgrounds, experience of being disciplined in early childhood, and parenting value on parenting styles of preschool children's parents. *Chinese Journal of Guidance and Counseling*, 32, 123-149. https://doi.org/10.7082/CJGC.201204.0121 (in Chinese)】
- 林信志、簡瑋成(2019)。臺灣都會地區國小弱勢學生暑期學習活動資本之研究。教育研究與發展期刊，15(3)，23-58。https://doi.org/10.3966/181665042019091503002【Lin, Hsin-Chih, & Chien, Wei-Cheng (2019). A study on learning cultural capital of disadvantaged students during the summer vacation in urban elementary schools in Taiwan. *Journal of Educational Research and Development*, 15(3), 23-58. https://doi.org/10.3966/181665042019091503002 (in Chinese)】
- 邱瑜瑾、姜義雯(2010)。貧窮家庭研究－經濟、社會與心理面向之敘述分析。人文社會科學研究，4(1)，107-131。https://doi.org/10.6284/NPUSTHSSR.2010.4(1)6【Chiou, Yu-Chin, & Chiang, Yi-Wen (2010). A study on poverty families-narrative Analysis from Economic, Social and Psychological Dimensions. *NPUST Humanities and Social Science Research*, 4(1), 107-131. https://doi.org/10.6284/NPUSTHSSR.2010.4(1)6 (in Chinese)】
- 洪清一(2011)。夜光天使點燈計畫對原住民高危險群學生學業學習、行為表現成效及其相關問題之研究。台灣原住民研究論叢，9，37-70。https://doi.org/10.29763/TISR.201106.0002【Hong, Ching-I (2011). Yeguang tianshi diandeng jihua dui yuanzhumin gao weixian qun xuesheng xueye xuexi, hangwei biao xian cheng xiao ji qi xiangguan wenti zhi yanjiu. *Taiwan Indigenous Studies Review*, 9, 37-70. https://doi.org/10.29763/TISR.201106.0002 (in Chinese)】
- 唐盈棣(2011)。偏鄉遠距課輔補救教學實施現況之研究－以南投縣為例〔未出版之碩士論文〕。國立暨南國際大學公共行政與政策學系。https://doi.org/10.6837/NCNU.2011.00018【Tang, Ying Di (2011). Remedial instruction through long-distance tutoring in remote areas -- The case study of Nantou County [Unpublished master's thesis]. Department of Public Policy and Administration, National Chi Nan University. https://doi.org/10.6837/NCNU.2011.00018 (in Chinese)】
- 莊沛蓉(2019)。國小學童參與課後方案之受輔經驗與其自我概念影響之研究－以桃園市為例〔未出版之碩士論文〕。中原大學教育研究所。https://doi.org/10.6840/cycu201900217【Chuang, Pei-Jung (2019). A study on the elementary school students with the experience of after-school programs and their self-concept -- Two Taoyuan's elementary schools as an example [Unpublished master's thesis]. Graduate School of Education, Chung Yuan Christian University. https://doi.org/10.6840/cycu201900217 (in Chinese)】
- 陳怡文、林麗娟(2014)。數位學伴課業輔導之教學分析。教學科技與學習，2(2)，207-234。【Chen, Yi-Wen, & Chan Lin, Lih-Juan (2014). Analysis of tutor teaching in online afterschool learning. *Educational Technology & Learning*, 2(2), 207-234. (in Chinese)】
- 陳淑麗(2009)。弱勢學童讀寫希望工程：課輔現場的瞭解與改造。心理出版社。【Chen, Su-Li (2009). *Ruoshi xuetong duxie xiwang gongcheng: Kefu xianchang de liaojie yu gaizao*. Psychological Publishing. (in Chinese)】

- 陳雅鈴 (2006)。貧窮如何影響兒童。教育研究月刊, 146, 87-101。【Chen, Ya-Ling (2006). Pinqiong ruhe yingxiang ertong. *Journal of Education Research*, 146, 87-101. (in Chinese)】
- 湯維玲、蔡佩娟 (2013)。一位國小校長推動攜手計畫補救教學之行動研究。課程與教學, 16(1), 69-92。https://doi.org/10.6384/CIQ.201301\_16(1).0003 【Tang, Wei-Ling, & Tsai, Pei-Juan (2013). An action research on a principal's promotion of after school alternative programs using remedial instruction. *Curriculum & Instruction Quarterly*, 16(1), 69-92. https://doi.org/10.6384/CIQ.201301\_16(1).0003 (in Chinese)】
- 黃河 (2018)。106 年度新北市兒童及少年生活狀況與福利需求調查：兒童調查報告 (案號 sab10610)。新北市政府社會局。https://www.sw.ntpc.gov.tw/uploaddowndoc?file=download/201810031109131.pdf&filedisplay=兒童調查報告.pdf&flag=doc 【Huang, He (2018). 106 niandu New Taipei City ertong ji shaonian shenghuo zhuangkuang yu fuli xuqiu diaocha: Ertong diaocha baogao (Report No. sab10610). Social Welfare Department, New Taipei City Government. https://www.sw.ntpc.gov.tw/uploaddowndoc?file=download/201810031109131.pdf&filedisplay=%E6%96%B0%E5%8C%97%E5%B8%82%E5%85%92%E7%AB%A5%E8%AA%BF%E6%9F%A5%E5%A0%B1%E5%91%8A.pdf&flag=doc (in Chinese)】
- 黃思維 (2014)。「幸福閱讀·學習相伴」全國首創圖書館陪讀服務。新北市志工季刊, 13, 20-21。【Huang, Si-Wei (2014). "Xingfu yuedu, xuexi xiangban" quanguo shou chuang tushuguan peidu fuwu. *New Taipei City Volunteer*, 13, 20-21. (in Chinese)】
- 劉鎮寧 (2013)。課後照顧經營與管理。心理出版社。【Liou, Jen-Ning (2013). *Kehou zhaogu jingying yu guanli*. Psychological Publishing. (in Chinese)】
- 歐怡珍 (2010)。新竹市國民小學實施夜光天使點燈計畫之個案研究〔未出版之碩士論文〕。國立新竹教育大學教育學系碩士班。【Ou, I-Chen (2010). *A case study of Night Angel Illumination Program's implementation in primary schools at Hsinchu City* [Unpublished master's thesis]. Department of Education, National Hsinchu University of Education. (in Chinese)】
- 鄭勝耀 (2013)。弱勢學生與補救教學之研究。教育人力與專業發展, 30(1), 13-26。【Cheng, Sheng-Yao (2013). Ruoshi xuesheng yu bujiu jiaoxue zhi yanjiu. *Educators and Professional Development*, 30(1), 13-26. (in Chinese)】
- 謝依紋、張書儀、傅郁涵 (2017)。從弱勢家庭個案管理服務談資源整合-陳網兒童及少年服務中心為例。家庭教育雙月刊, 68, 57-64。【Xie, Yi-Wen, Zhang, Shu-Yi, & Fu Yu Han (2017). Cong ruoshi jiating gean guanli fuwu tan ziyuan zhenghe -- Chen Chou Ertong Ji Shaonian Fuwu Zhongxin wei li. *Journal of Family Education Bimonthly*, 68, 57-64. (in Chinese)】
- 顏國樑、宋美瑤 (2013)。苗栗縣國民小學攜手計畫課後扶助執行成效之分析與建議。學校行政, 83, 177-197。【Yen, Kuo-Liang, & Song, Mei-Yao (2013). A study of suggestions and effective analysis on implementing the hand-in-hand plan after-school support policy of elementary school in Miaoli County. *Xuexiao Xingzheng*, 83, 177-197. (in Chinese)】
- 羅素貞 (2011)。臺北市立圖書館弱勢兒童閱讀服務。臺北市立圖書館館訊, 28(3), 53-64。https://doi.org/10.29773/BTPL.201103.0005 【Luo, Shu-Zhen (2011). Taipei

- Public Library reading services for underprivileged children. *Bulletin of the Taipei Public Library*, 28(3), 53-64. <https://doi.org/10.29773/BTPL.201103.0005> (in Chinese)】
- 蘇秀枝 (2005)。國小學童課後托育、補習才藝與學業成就、行為適應之關係—以台中縣大里市為例。朝陽人文社會學刊，3(1)，173-223。【Su, Hsiu-chih (2005). The relations between participation in after-school programs and talent classes and children's grades and behavioral adjustment in Tali City in Central Taiwan. *Chaoyang Journal of Humanities and Social Sciences*, 3(1), 173-223. (in Chinese)】
- Aabø, S., Audunson, R., & Vårheim, A. (2010). How do public libraries function as meeting places? *Library & Information Science Research*, 32(1), 16-26. <https://doi.org/10.1016/j.lisr.2009.07.008>
- American Library Association. (2019). *Core values of librarianship*. <http://www.ala.org/advocacy/intfreedom/corevalues>
- Bailey, J. P. (1999). Quantifying library quality: A homework center report card. *American Libraries*, 30(8), 59-62.
- Bevin, D., & Goulding, A. (1999). Homework clubs in public libraries. *New Library World*, 100(2), 49-59. <https://doi.org/10.1108/03074809910263774>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Bundy, A. (2006). Supporting students: The educational contribution of Australia's public libraries. *Australasian Public Libraries and Information Services*, 19(3), 126-135.
- Cassell, K. A., & Walther, J. H. (2006). Supporting children and teens after school: A library approach. *The Bottom Line*, 19(1), 16-21. <https://doi.org/10.1108/08880450610643025>
- Conger, R. D., Wallace, L. E., Sun, Y., Simons, R. L., McLoyd, V. C., & Brody, G. H. (2002). Economic pressure in African American families: A replication and extension of the family stress model. *Developmental Psychology*, 38(2), 179-193.
- Dodge, K. A., Pettit, G. S., & Bates, J. E. (1994). Socialization mediators of the relation between socioeconomic status and child conduct problems. *Child Development*, 65(2), 649-665. <https://doi.org/10.2307/1131407>
- Dowd, F. S. (1989). The public library & the latchkey problem: A survey. *School Library Journal*, 35(11), 19-24.
- Free Library of Philadelphia. (2020). *LEAP afterschool program*. <https://libwww.freelibrary.org/programs/leap/>
- Hartman, M. (2011). Out of school and in the library: Connecting with resources in the out of school time (OST) field. *Young Adult Library Services*, 9(4), 10-12.
- Huffman, C., & Rua, R. J. (2008). Measuring the effectiveness of homework centers in libraries. *Children & Libraries*, 6(3), 25-29.
- International Federation of Library Associations and Institutions. (2012). *IFLA code of ethics for librarians and other information workers (full version)*. <https://www.ifla.org/publications/ifla-code-of-ethics-for-librarians-and-other-information-workers-full-version/>
- Lacour, M., & Tissington, L. D. (2011). The effects of poverty on academic achievement. *Educational Research and Reviews*, 6(7), 522-527.

- Long, S. A. (2000). Message from ALA President Sarah Ann Long: Two afterschool programs ASPIRE to take the initiative. *American Libraries*, 31(3), 7.
- Mcdermott, I. E. (2002). Internet express: Homework help on the web. *Searcher*, 10(4), 10-18.
- Mediavilla, C. (2001). Why library homework centers extend society's safety net. *American Libraries*, 32(11), 40-42.
- Mediavilla, C. (2003). Homework helpers. *School Library Journal*, 49(3), 56-59.
- Mediavilla, C. (2018). Staffing your homework-help center: Pairing the right minds with student learners. *American Libraries*, 49(7-8), 34-37.
- Minkel, W. (2002). When homework is good politics. *School Library Journal*, 48(4), 39.
- Pateman, J., & Vincent, J. (2010). *Public libraries and social justice*. Ashgate.
- Rioux, K. (2010). Metatheory in library and information science: A nascent social justice approach. *Journal of Education for Library and Information Science*, 51(1), 9-17.
- Rua, R. J. (2008). After-school success stories: In six years of trial and error, a grassroots effort to save a school district has blossomed into a thriving after-school program helping thousands of at-risk students. *American Libraries*, 39(10), 46-48.
- Shaffer, G. (2006). Stop the presses! A hip new after-school program has turned kids into modern-day Jimmy Olsens and Lois Lanes. *School Library Journal*, 52(7), 38-39.
- Social Responsibilities Round Table. (2019). *About SRRT*. <http://www.ala.org/rt/srrt/about-srrt>
- Train, B., & Elkin, J. (2000). Homework clubs: A model for the qualitative evaluation of public library initiatives. *New Review of Children's Literature and Librarianship*, 6(1), 177-192. <https://doi.org/10.1080/13614540009510638>



# Exploring Article Process Charge of Open Access Journals from the Perspectives of Publication Characteristics and Citation Impact Indicators: A Case Study in the Medical Field<sup>ψ</sup>

Chia-Yu Lin<sup>a</sup> Wen-Yau Cathy Lin<sup>b\*</sup>

## Abstract

*The complex APC operation mechanism of open access (OA) journal is generating discussion in the academic community. To understand the prices of APC, this study explored whether APC pricing in medicine is related to journal-related characteristics and journal influence-related indicators. In all, 47 medicine-related fields in JCR-SCIE 2017 were collected, among which 3,420 journals had published OA articles and 2,037 OA journals were analyzed, which removal of duplicate titles and picking out fixed APC prices. Results showed that APCs and journal age were not significantly correlated; journals established earlier may have charged lower APCs than those established later. Overall, the majority of OA journals charged approximately US\$3,000. APCs and publication frequency were positively correlated, APCs increased with publication frequency. APCs and the number of articles published were negatively correlated, perhaps because hybrid OA journals published fewer articles but charged higher APCs. Regarding journal influence, IF and II values were significantly correlated with APCs. With respect to rankings, for journals ranked in Q1 and Q4, correlations between ranking and APC were significant. This study suggests that future may use questionnaires or conduct in-depth interviews to gain insight into why authors submitted articles to OA journals, their willingness to submit the articles and how prices they were willing to pay in APCs.*

**Keywords:** Open access, Article processing charge, Journal publishing characteristics, Impact factor, Journal ranking, Immediacy index

<sup>ψ</sup> This article is based on the first author Chia-Yu Lin's master thesis "A Study of the Open Access Journal Article Process Charge in Medical: The Perspectives of Journal Publishing Characteristics and Citation Impact", and the original research idea is inspired by her advisor Wen-Yau Cathy Lin.

<sup>a</sup> Master, Department of Information & Library Science, Tamkang University, New Taipei, Taiwan

<sup>b</sup> Professor, Department of Information and Library Science, Tamkang University, New Taipei City, Taiwan

\* To whom all correspondence should be addressed. E-mail: wylin@mail.tku.edu.tw

## **SUMMARY**

### **Introduction**

OA journals adopt complex operating methods, and some charge high publication fees, generating discussion in the academic community. Contrary to the publication methods adopted by toll access journals, OA journals require that authors or their organizations pay the publication costs. This payment is referred to as the “article processing charge” (APC), and it imposes a heavy burden on authors (Nelson & Eggett, 2017; Sotudeh & Ghasempour, 2018). Although the future operation of the APC mechanism of OA journals is still unknown, studies have revealed that APCs are the main source of profit for OA journals (Davis & Walters, 2011; Laakso & Björk, 2012).

Regarding standards for setting APCs, Solomon and Björk (2012b) commented that APCs vary by field, journal publishing country, number of articles published, and journal influence, where more influential journals (i.e., those with higher rankings) have higher APCs. Dorta-Gonzalez et al. (2017) stated that, currently, full OA journals generally rank low in influence, and those that rank comparatively higher in influence charge higher APCs. Those researchers who cannot afford them agonize that their results will not be published in an influential OA journal. Concerns have been raised about the unequal distribution of scientific resources caused by such discrimination (Papin-Ramcharan & Dawe, 2006; Siler et al., 2018; Solomon & Björk, 2012a).

This study used journal publication characteristics and influence to investigate whether APC pricing in medicine is related to OA journal-related characteristics (i.e., the year the journal was established, the frequency that the journal publishes papers and the number of articles published) and journal influence-related indicators. The study questions were as follows:

1. In the field of medicine, what are the relationships between journal-related characteristics and APCs?
2. In the field of medicine, what are the relationships between APCs and the journal's impact factor (IF), Journal Ranking and immediacy index (II) values?

### **Research Methods**

This study used bibliometrics to collect medical domain-related full OA journals in JCR-SCIE 2017 and used the SCI-EXPANDED function of WoS to filter out hybrid journals that published OA articles but were not listed as OA journals in the JCR.

In all, 3,825 journals in 47 medicine-related fields in JCR-SCIE 2017 were counted, among which 3,420 had published OA articles. However, journals



can be listed under multiple categories and be counted multiple times; after the removal of duplicates, 2,665 OA journals were identified. Considering that APC charging method is very diverse and complex, and there will be different pricing amounts according to different conditions. Therefore, only journals with fixed APC are for analysis. In total 2,037 journals were calculated.

Concerning APCs, most OA journals have multiple pricing rules because OA articles have different authorization terms and publication durations. Nonetheless, to allow for comparisons, a fixed APC price was used when calculating the APC charged by a journal. Because the data were collected during the first half of 2019, the APCs were displayed in US\$ based on the average exchange rate of this period.

## Results

### Correlations Between APCs and OA Journal-Related Characteristics

#### 1. Year Journal Was Established

Pearson correlation coefficient analysis was conducted on 2,036 journals (one journal without publication year information was removed), and the coefficient between the year journal was established to 2020 is .041 ( $p = .066$ ). For most categories, the APC was concentrated US\$3,000.

#### 2. Publication Frequency

Pearson correlation coefficient analysis of publication frequency and APC was .239 ( $p < .001$ ), indicating significant correlation and that APC may be affected by publication frequency. The median publication frequency among the categories was six times a year. TRANSPLANTATION, which had both the highest median APCs and median publication frequency.

#### 3. Number of OA Articles Published

In the Pearson correlation analysis of the number of OA articles published and APC was  $-.150$  ( $p < .001$ ), indicating that the number of OA articles published was negatively correlated with APC. In other words, the APC increased as the number of OA articles decreased. This result may have been caused by the different OA journal types.

Concerning the medians of number of OA articles and APCs for all journal categories. TROPICAL MEDICINE, had the highest median number of OA articles (120.5 articles) and the lowest median APC.

### Correlations Between Journal Influence-Related Indicators and APCs

#### 1. Correlations Between Journal IF and APCs

Journals without IF data (five journals) were removed, and Pearson correlation analysis was conducted on 2,032 journals. The coefficient for journal IF value and APC was .168 ( $p < .001$ ), which was significant.

Because some journals had extremely high IF values, the journals were divided into two groups. For the  $> 10$  group (57 journals), the correlation



coefficient was .025 ( $p = .854$ ), which was nonsignificant. By contrast, for the  $< 10$  group (1,975 journals), the correlation coefficient was .312 ( $p < .001$ ), which was significant. APC increased with IF, however, when the IF value exceeded 10, no reasonable explanation could be provided for the APC growth.

**2. Correlations Between Journal Ranking and APC**  
All journals (i.e., 2,619, which included journals counted more than once) the correlation coefficient was  $-.349$  ( $p < .001$ ), which was significant and indicated that journal ranking and APC were negatively correlated. In other words, higher APCs droved by better journal ranking.

Quartile analyses indicated that for journals ranked in Q1 and Q4, the correlations between journal ranking and APC were significant (achieving a significance level of  $p < .01$ ). By contrast, for Q2 and Q3 journals, no significant correlations were observed. Table 1 presents the correlations between journal ranking and APC.

**Table 1 Correlations Between Journal Ranking and APC**

Ranking distribution	Number of journals published	Correlation
Overall	2,619	$-.349^{**}$ (.000)
Q1 (1-25%)	671	$-.119^{**}$ (.002)
Q2 (26-50%)	725	$-.038$ (.309)
Q3 (51-75%)	719	$-.017$ (.651)
Q4 (76-100%)	504	$-.175^{**}$ (.000)

Note: Journals included those counted more than once; numbers in parentheses are  $p$  values.  
 $^{**}$  correlation level of  $p < .01$  (two-tailed).

**3. Correlations Between II and APCs**  
Pearson correlation analysis for journal II value and APC was .227 ( $p < .001$ ), achieving significance. Because some journals had high II extreme values, to properly identify the correlations between journal II values and APCs charged, the journals were divided into two groups. For the  $> 1$  group (446 journals), the correlation coefficient was .142 ( $p = .003$ ); the  $< 1$  group (1,591 journals), the correlation coefficient was .402 ( $p < .001$ ), both have significant.

The II value  $< 1$  group had a stronger II value–APC correlation. By contrast, when the journal II value was greater than one, the correlation between journal II value and APC charged (i.e., the rule of APC growth) became less visible.

**Discussion and Suggestion**

Through the two aspects of journal publication characteristics and influence to explore the correlation with APC price. The results of the study found that publication characteristics and APC only showed a low correlation, which further showed that journal publisher may still regard journal influence as the main factor when setting APC.

We suggest that future studies investigate library, publisher, and sponsor unit personnel's awareness of, adaptive strategies for, and views on APC mechanisms to help academic researchers achieve a more in-depth understanding of the publishing mechanism of APC.

### **ROMANIZED & TRANSLATED REFERENCES FOR ORIGINAL TEXT**

- 林家鈺、林雯瑤 (2021)。醫學領域開放取用期刊文章處理費之研究。圖書資訊學刊, 19(2), 173-203。https://doi.org/10.6182/jlis.202112\_19(2).173 【Lin, Chia-Yu, & Lin, Wen-Yau Cathy (2021). Open access journal article process charges in the field of medicine. *Journal of Library and Information Studies*, 19(2), 173-203. https://doi.org/10.6182/jlis.202112\_19(2).173 (in Chinese)】
- 邱炯友 (2006)。學術傳播與期刊出版。遠流。【Chiu, Jeong-Yeou (2006). *Scholarly communication and journal publishing*. Yuan-Liou. (in Chinese)】
- Anderson, R. (2004). Author disincentives and open access. *Serials Review*, 30(4), 288-291. https://doi.org/10.1016/j.serrev.2004.09.001
- Antelman, K. (2004). Do open-access articles have a greater research impact? *College & Research Libraries*, 65(5), 372-382. https://doi.org/10.5860/crl.65.5.372
- Asai, S. (2021). An analysis of revising article processing charges for open access journals between 2018 and 2020. *Learned Publishing*, 34(2), 137-143. https://doi.org/10.1002/leap.1334
- Baro, E. E., & Eze, M. E. (2017). Perceptions, preferences of scholarly publishing in open access routes: A survey of academic librarians in Nigeria. *Information and Learning Science*, 118(3-4), 152-169. https://doi.org/10.1108/ILS-03-2017-0015
- Björk, B.-C., & Solomon, D. (2012). Open access versus subscription journals: A comparison of scientific impact. *BMC Medicine*, 10. https://doi.org/10.1186/1741-7015-10-73
- Björk, B.-C., & Solomon, D. (2014). How research funders can finance APCs in full OA and hybrid journals. *Learned Publishing*, 27(2), 93-103. https://doi.org/10.1087/20140203
- Budzinski, O., Grebel, T., Wolling, J., & Zhang, X. (2020). Drivers of article processing charges in open access. *Scientometrics*, 124(3), 2185-2206. https://doi.org/10.1007/s11192-020-03578-3
- Copiello, S. (2020). Business as usual with article processing charges in the transition towards OA publishing: A case study based on Elsevier. *Publications*, 8(1), 3. https://doi.org/10.3390/publications8010003
- Davis, P. M., & Walters, W. H. (2011). The impact of free access to the scientific literature: A review of recent research. *Journal of the Medical Library Association*, 99(3), 208-217. https://doi.org/10.3163/1536-5050.99.3.008
- Dorta-González, P., González-Betancor, S. M., & Dorta-González, M. I. (2017). Reconsidering the gold open access citation advantage postulate in a multidisciplinary context: An analysis of the subject categories in the Web of Science database 2009-2014. *Scientometrics*, 112(2), 877-901. https://doi.org/10.1007/s11192-017-2422-y
- Eysenbach, G. (2006). Citation advantage of open access articles. *PLoS Biology*, 4(5), e157. https://doi.org/10.1371/journal.pbio.0040157

- Fukuzawa, N. (2017). Characteristics of papers published in journals: An analysis of open access journals, country of publication, and languages used. *Scientometrics*, 112(2), 1007-1023. <https://doi.org/10.1007/s11192-017-2414-y>
- Gadd, E., Fry, J., & Creaser, C. (2018). The influence of journal publisher characteristics on open access policy trends. *Scientometrics*, 115(3), 1371-1393. <https://doi.org/10.1007/s11192-018-2716-8>
- Hampson, C., & Stregger, E. (2017). Measuring cost per use of library-funded open access article processing charges: Examination and implications of one method. *Journal of Librarianship and Scholarly Communication*, 5(1), p.eP2182. <https://doi.org/10.7710/2162-3309.2182>
- Harnad, S., & Brody, T. (2004). Comparing the impact of open access (OA) vs. non-OA articles in the same journals. *D-Lib Magazine*, 10(6). <https://doi.org/10.1045/june2004-harnad>
- Jahn, N., & Tullney, M. (2016). A study of institutional spending on open access publication fees in Germany. *PeerJ*, 4, e2323. <https://doi.org/10.7717/peerj.2323>
- Kingsley, D. A. (2014). Paying for publication: Issues and challenges for research support services. *Australian Academic and Research Libraries*, 45(4), 262-281. <https://doi.org/10.1080/00048623.2014.945135>
- Laakso, M., & Björk, B.-C. (2012). Anatomy of open access publishing: A study of longitudinal development and internal structure. *BMC Medicine*, 10(1), 124. <https://doi.org/10.1186/1741-7015-10-124>
- McVeigh, M. E. (2004). *Open access journals in the ISI citation databases: Analysis of impact factors and citation patterns: A citation study from Thomson Scientific*. <https://pdf4pro.com/amp/view/open-access-journals-in-the-thomson-isi-citation-databases-4f8b6b.html>
- Moed, H. F. (2007). The effect of “Open Access” on citation impact: An analysis of ArXiv’s condensed matter section. *Journal of the American Society for Information Science and Technology*, 58(13), 2047-2054. <https://doi.org/10.1002/asi.20663>
- Monson, J., Highby, W., & Rathe, B. (2014). Library involvement in faculty publication funds. *College and Undergraduate Libraries*, 21(3-4), 308-329. <https://doi.org/10.1080/10691316.2014.933088>
- Mukherjee, B. (2014). Green and gold open access in India. *Learned Publishing*, 27(1), 21-32. <https://doi.org/10.1087/20140104>
- Nelson, G. M., & Eggett, D. L. (2017). Citations, mandates, and money: Author motivations to publish in chemistry hybrid open access journals. *Journal of the Association for Information Science and Technology*, 68(10), 2501-2510. <https://doi.org/10.1002/asi.23897>
- Norris, M., Oppenheim, C., Rowland, F. (2008) The citation advantage of open-access articles. *Journal of the American Society for Information Science and Technology*, 59(12), 1963-1972. <https://doi.org/10.1002/asi.20898>
- Papin-Ramcharan, J., & Dawe, R. A. (2006). The other side of the coin for open access publishing - A developing country view. *Libri*, 56(1), 16-27. <https://doi.org/10.1515/LIBR.2006.16>
- Pinfield, S., Salter, J., & Bath, P. A. (2016). The “total cost of publication” in a hybrid open-access environment: Institutional approaches to funding journal article-processing

- charges in combination with subscriptions. *Journal of the Association for Information Science and Technology*, 67(7), 1751-1766. <https://doi.org/10.1002/asi.23446>
- Pinfield, S., Salter, J., & Bath, P. A. (2017). A “Gold-centric” implementation of open access: Hybrid journals, the “total cost of publication,” and policy development in the UK and beyond. *Journal of the Association for Information Science and Technology*, 68(9), 2248-2263. <https://doi.org/10.1002/asi.23742>
- Schimmer, R., Geschuhn, K. K., & Vogler, A. (2015). Disrupting the subscription journals’ business model for the necessary large-scale transformation to open access: A Max Planck Digital Library open access policy white paper. *ScienceOpen Research*. <https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AJRG23.v1>
- Siler, K., Haustein, S., Smith, E., Larivière, V., & Alperin, J. P. (2018). Authorial and institutional stratification in open access publishing: The case of global health research. *PeerJ*, 6, e4269. <https://doi.org/10.7717/peerj.4269>
- Smith, E., Haustein, S., Mongeon, P., Shu, F., Ridde, V., & Larivière, V. (2017). Knowledge sharing in global health research - the impact, uptake and cost of open access to scholarly literature. *Health Research Policy and Systems*, 15(1). <https://doi.org/10.1186/s12961-017-0235-3>
- Solomon, D., & Björk, B.-C. (2012a). Publication fees in open access publishing: Sources of funding and factors influencing choice of journal. *Journal of the American Society for Information Science and Technology*, 63(1), 98-107. <https://doi.org/10.1002/asi.21660>
- Solomon, D. J., & Björk, B.-C. (2012b). A study of open access journals using article processing charges. *Journal of the American Society for Information Science and Technology*, 63(8), 1485-1495. <https://doi.org/10.1002/asi.22673>
- Solomon, D., & Björk, B.-C. (2016). Article processing charges for open access publication-the situation for research intensive universities in the USA and Canada. *PeerJ*, 4, e2264. <https://doi.org/10.7717/peerj.2264>
- Solomon, D., Laakso, M., & Björk, B.-C. (2013). A longitudinal comparison of citation rates and growth among open access journals. *Journal of Informetrics*, 7(3), 642-650. <https://doi.org/10.1016/j.joi.2013.03.008>
- Sotudeh, H., & Ghasempour, Z. (2018). The world’s approach toward publishing in Springer and Elsevier’s APC-funded open access journals. *College & Research Libraries*, 79(2), 257-278. <https://doi.org/10.5860/crl.79.2.257>
- Sotudeh, H., Ghasempour, Z., & Yaghtin, M. (2015). The citation advantage of author-pays model: The case of Springer and Elsevier OA journals. *Scientometrics*, 104(2), 581-608. <https://doi.org/10.1007/s11192-015-1607-5>
- Tang, M., Bever, J. D., & Yu, F. -H. (2017). Open access increases citations of papers in ecology. *Ecosphere*, 8(7). <https://doi.org/10.1002/ecs2.1887>
- Walters, W. H., & Linvill, A. C. (2011). Characteristics of open access journals in six subject areas. *College & Research Libraries*, 72(4), 372-392. <https://doi.org/10.5860/crl-132>
- Wang, L. L., Liu, X. Z., & Fang, H. (2015). Investigation of the degree to which articles supported by research grants are published in open access health and life sciences journals. *Scientometrics*, 104(2), 511-528. <https://doi.org/10.1007/s11192-015-1624-4>

- West, J. D., Bergstrom, T., & Bergstrom, C. T. (2014). Cost effectiveness of open access publications. *Economic Inquiry*, 52(4), 1315-1321. <https://doi.org/10.1111/ecin.12117>
- Xia, J., Myers, R. L., & Wilhoite, S. K. (2011). Multiple open access availability and citation impact. *Journal of Information Science*, 37(1), 19-28. <https://doi.org/10.1177/0165551510389358>
- Yan, E., & Li, K. (2018). Which domains do open-access journals do best in? A 5-year longitudinal study. *Journal of the Association for Information Science and Technology*, 69(6), 844-856. <https://doi.org/10.1002/asi.24002>
- Yuen, J., Muquit, S., & Whitfield, P. C. (2019). Correlation between cost of publication and journal impact. Comprehensive cross-sectional study of exclusively open-access surgical journals. *Journal of Surgical Education*, 76(1), 107-119. <https://doi.org/10.1016/j.jsurg.2018.06.029>

JoEMLS English Summary



# Estimation of Topic Similarity and Its Application to Measuring Stability of Topic Modeling

Sung-Chien Lin

## Abstract

*Topic modeling stability is a measurement of the extent to which models produced by the same modeling approach for the same corpus and with the same initial conditions have similar topics. Since the method used for calculating similarity between topics is considered the basis for measuring topic modeling stability and topic alignment is a key step in the measurement, the present study first calculated the proportion of identical paired topics among the optimal combinations of paired topics generated using different topic similarity calculation methods, and then observed the distribution of similarity scores of paired topics for each method. Finally, this study performed an analysis of the effects of the number of topics on topic modeling stability. The topic modeling method used in this study is commonly used LDA topic modeling, and the corpus used to establish topic models including about 30,000 posts was collected from the PTT Bulletin Board System (BBS) Book message board. The results indicated that there is a high proportion of identical paired topics among the different methods of measuring similarity, although the similarity scores of paired topics for each method had different distributions due to the different kinds and amounts of information of word distribution in each topic they used. The results also revealed that with the increase of the number of topics, the stability noticeably decreased.*

**Keywords:** Topic modeling, latent Dirichlet allocation (LDA), Stability measurement, Topic similarity estimation, Topic alignment

## SUMMARY

### Introduction

Topic modeling can reveal topic structures contained in a corpus and aid in the rapid and effective analyses of large amounts of text. Currently, latent Dirichlet allocation (LDA; Blei et al., 2003) is regarded as the most popular topic

---

Assistant Professor, Department of Information and Communications, ShihHsin University, Taipei, Taiwan  
E-mail: scl@mail.shu.edu.tw

modeling technique among researchers and is widely used for problems involving text analyses (Lancichinetti et al., 2015). However, in practice, even with the same parameters and corpus, the models produced with this technique somewhat differ from each other, calling into question the reliability of the analysis results (Maier et al., 2018). This problem casts doubt on the usefulness of LDA topic modeling (Belford et al., 2018; Chuang et al., 2015).

Topic modeling stability is a measurement of the extent to which models produced by the same modeling approach for the same corpus and with the same initial conditions have similar topics. Several methods can be used to measure topic modeling stability. For instance, in the present study, the framework used for measuring topic modeling stability (De Waal & Barnard, 2008; Greene et al., 2014) involved producing multiple topic models through repeated modeling with the same corpus and number of topics and then performing topic alignment between any two topic models by using the Hungarian algorithm to determine the optimal combination of topic pairs. In this combination, the mean similarity of the topic pairs was the agreement score of the two models, whereas the mean of the agreement scores was the measurement of the topic modeling stability.

According to this measurement framework, the method used for calculating similarity between topics is considered the basis for measuring topic modeling stability. Belford et al. (2018) and Greene et al. (2014) used Jaccard's score (JAC) to calculate topic similarity; however, their approach considered only a small portion of information in the word distribution of each topic. Therefore, in the present study, the following six methods for measuring topic similarity were used and compared: Jensen-Shannon divergence (JSD), normalized discounted cumulative gain (NDCG), cosine measure (COS), JAC, rank-biased overlap score (RBO), and Kendall's  $\tau$  coefficient (KEN). Topic alignment is a key step in this measurement framework. If two different methods for measuring topic similarity yield highly similar optimal combinations of topic pairs, the two methods may have similar stability measurement outcomes. The distribution of the similarity score of paired topics can also indicate which methods are more likely to identify the topics that appear in most models after topic alignment.

This study performed the following analysis tasks:

Task 1: Conduct an analysis of the proportion of identical paired topics among the optimal combinations of paired topics generated using different topic similarity calculation methods.

Task 2: Perform an analysis of the distribution of similarity scores of paired topics for each method.

Overall, the study conducted by Greene et al. (2014) is regarded as one of the few studies analyzing the effects of the number of topics on topic modeling



stability. However, the corpus used in that study had few topics, which were already clearly defined. Therefore, a corpus with a greater number of topics was used in the present study.

Task 3: Perform an analysis of the effects of the number of topics on topic modeling stability.

## Research Methods

Word segmentation, part-of-speech tagging, and stop word removal were performed on 32,895 posts collected from the PTT Bulletin Board System (BBS) Book message board. Posts containing at least five words were selected to form a corpus for analyzing topic modeling stability. The final corpus included 20,287 posts and 1,579,116 words. The topic modeling inputs consisted of this corpus and a dictionary. For each different number of topics ( $K = 5, 10, 15, \dots, 100$ ), a total of 20 models were created with fixed prior parameters  $\alpha$  and  $\beta$ .

Next, the six methods of measurement mentioned earlier were used with any two topic models to calculate the similarity between each topic pair. The results of each method were then adjusted to be between 0 and 1. The greater the similarity between any two topics was, the greater the score was. The similarity scores of all pairs of topics between every two models were then entered into the Hungarian algorithm to align the topics and obtain an optimal combination of topic pairs. Analysis Tasks 1 and 2 were then performed.

Finally, the agreement score between every two topic models was obtained by averaging the optimal topic pair similarity scores. Analysis Task 3 was then performed using the mean agreement score between each pair of topic models as the stability measurement.

## Research Result

### Task 1

This task involved assessing whether different methods of measuring similarity had the same effect when measuring stability based on the proportion of identical topic pairs in the optimal combinations of topic pairs. The results obtained indicated a high proportion of identical paired topics among the different methods of calculating topic similarity. The proportion of identical paired topics among the six methods reached 76.99%, and the total proportion even increased to 94.09% in four or more methods. For any two methods, the proportion of identical topic pairs was 84% or higher, suggesting that any two methods had similar stability outcomes. However, slight differences were observed between the methods that involved the use of all word distribution data, such as JSD, NDCG, COS, and KEN, and the methods that involved only a few keywords, such as JAC and RBO.

## **Task 2**

If a method for calculating topic similarity can yield a high similarity score between postalignment paired topics, then this means that this method can differentiate between similar topics within different models and thereby identify stable topics in each model. In this study, rather high similarity scores were observed among most of the paired topics when JSD, NDCG, and COS were used, which are methods that involve the use of the occurrence probability of all words in each topic, showing that these methods can easily identify stable topics in models. JAC and RBO are methods that involve the use of a set of keywords to represent topics. In this study, these two methods yielded similarity scores that were scattered across a wide range. In addition, approximately 5% of the similarity scores were 0, because the corresponding paired topics had completely different keywords. The KEN method considers every word to have a consistent order of occurrences among paired topics. However, each topic contains several irrelevant and low-probability words, which may cause similar topics to exhibit dissimilar orders and hence lower the similarity scores.

## **Task 3**

This task entailed measuring the stability of topic models with different numbers of topics. The results revealed that with the increase of the number of topics, the stability noticeably decreased. This may be because with the increase of the number of topics, the topic ranges in the model became narrower, and the distribution of words in the topic became more prone to change. This may have resulted in an increasing number of topics being unable to align with similar topics in another model, thereby lowering the stability.

## **Suggestions and Future Research**

In this study, topic alignment was performed using the Hungarian algorithm, and the agreement score between models was calculated on the basis of the similarity scores between paired topics. Future researchers may refer to Maier et al. (2018) and use the proportion of possible pairs as an indicator of model stability to develop a method of measurement that is suited to direct interpretation.

During text analyses with topic modeling, the number of topics is considered a key parameter that determines the scope, accuracy, and interpretability of the model. Several studies have employed perplexity or topic coherence as an indicator of topic model quality to determine the optimal number of topics, and some have even involved manual reviews (Maier et al., 2018). Therefore, we suggest integrating stability with other quality indicators to determine the optimal number of topics.

Finally and most importantly, the methods used to improve topic modeling stability should be further developed. Increasing the level of stability can help increase the possible similarity in topics among all models produced under the same input conditions and thereby enhance the reliability of the text analysis results. Among the current studies investigating this topic are those of Chuang et al. (2015), Lancichinetti et al. (2015), Koltcov et al. (2016), Agrawal et al. (2018), Maier et al. (2018), and Mantyla et al. (2018).

### **ROMANIZED & TRANSLATED REFERENCES FOR ORIGINAL TEXT**

- Agrawal, A., Fu, W., & Menzies, T. (2018). What is wrong with topic modeling? And how to fix it using search-based software engineering. *Information and Software Technology*, 98, 74-88. <https://doi.org/10.1016/j.infsof.2018.02.005>
- Ballester, O., & Penner, O. (2022). Robustness, replicability and scalability in topic modelling. *Journal of Informetrics*, 16(1), 101224. <https://doi.org/10.1016/j.joi.2021.101224>
- Belford, M., Mac Namee, B., & Greene, D. (2018). Stability of topic modeling via matrix factorization. *Expert Systems with Applications: An International Journal*, 91, 159-169. <https://doi.org/10.1016/j.eswa.2017.08.047>
- Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. In W. W. Cohen & A. Moore (Eds.), *Proceedings of the 23rd international conference on machine learning* (pp. 113-120). ACM. <https://doi.org/10.1145/1143844.1143859>
- Blei, D. M., & Lafferty, J. D. (2007). A correlated topic model of science. *The Annals of Applied Statistics*, 1(1), 17-35. <https://doi.org/10.1214/07-AOAS114>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Chuang, J., Roberts, M. E., Stewart, B. M., Weiss, R., Tingley, D., Grimmer, J., & Heer, J. (2015). TopicCheck: Interactive alignment for assessing topic model stability. In R. Mihalcea, J. Chai, & A. Sarkar (Eds.), *Proceedings of the 2015 conference of the North American Chapter of the Association for Computational Linguistics: Human language technologies* (pp. 175-184). Association for Computational Linguistics. <https://doi.org/10.3115/v1/N15-1018>
- De Waal, A., & Barnard, E. (2008). Evaluating topic models with stability. In F. Nicolls (Ed.), *Proceedings of the 19th annual symposium of the Pattern Recognition Association of South Africa* (pp. 79-84). Pattern Recognition Association of South Africa.
- Elgesem, D., Steskal, L., & Diakopoulos, N. (2015). Structure and content of the discourse on climate change in the blogosphere: The big picture. *Environmental Communication*, 9(2), 169-188. <https://doi.org/10.1080/17524032.2014.983536>
- Greene, D., O'Callaghan, D., & Cunningham, P. (2014). How many topics? Stability analysis for topic models. In T. Calders, F. Esposito, E. Hüllermeier, & R. Meo (Eds.), *ECML PKDD 2014: Machine learning and knowledge discovery in databases* (pp. 498-513). Springer. [https://doi.org/10.1007/978-3-662-44848-9\\_32](https://doi.org/10.1007/978-3-662-44848-9_32)

- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(Suppl. 1), 5228-5235. <https://doi.org/10.1073/pnas.0307752101>
- Hofmann, T. (1999). Probabilistic latent semantic analysis. In K. B. Laskey & H. Prade (Eds.), *UAI'99: Proceedings of the fifteenth conference on Uncertainty in artificial intelligence* (pp. 289-296). Morgan Kaufmann.
- Jacobi, C., Van Atteveldt, W., & Welbers, K. (2016). Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, 4(1), 89-106. <https://doi.org/10.1080/21670811.2015.1093271>
- Järvelin, K., & Kekäläinen, J. (2002). Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems*, 20(4), 422-446. <https://doi.org/10.1145/582415.582418>
- Kim, D., & Oh, A. (2011). Topic chains for understanding a news corpus. In A. Gelbukh (Ed.), *CICLing 2011: Computational linguistics and intelligent text processing* (pp. 163-176). Springer. [https://doi.org/10.1007/978-3-642-19437-5\\_13](https://doi.org/10.1007/978-3-642-19437-5_13)
- Koltcov, S., Nikolenko, S. I., Koltsova, O., Filippov, V., & Bodrunova, S. (2016). Stable topic modeling with local density regularization. In F. Bagnoli et al. (Eds.), *INSCI 2016: Internet science* (pp. 176-188) Springer. [https://doi.org/10.1007/978-3-319-45982-0\\_16](https://doi.org/10.1007/978-3-319-45982-0_16)
- Kuhn, H. W. (1955). The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2), 83-97. <https://doi.org/10.1002/nav.3800020109>
- Lancichinetti, A., Sirer, M. I., Wang, J. X., Acuna, D., Kording, K., & Amaral, L. A. N. (2015). High-reproducibility and high-accuracy method for automated topic classification. *Physical Review X*, 5(1), 011007. <https://doi.org/10.1103/PhysRevX.5.011007>
- Li, P.-H., Fu, T.-J., & Ma, W.-Y. (2020). Why attention? Analyze BiLSTM deficiency and its remedies in the case of NER. *AAAI Conference on Artificial Intelligence*, 34(05), 8236-8244. <https://doi.org/10.1609/aaai.v34i05.6338>
- Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., Schmid-Petri, H., & Adam, S. (2018). Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. *Communication Methods and Measures*, 12(2-3), 93-118. <https://doi.org/10.1080/19312458.2018.1430754>
- Mantyla, M. V., Claes, M. & Farooq, U. (2018). Measuring LDA topic stability from clusters of replicated runs. In M. Oivo (Chair), *ESEM'18: Proceedings of the 12th ACM/IEEE international symposium on empirical software engineering and measurement* (Article No. 4). ACM. <https://doi.org/10.1145/3239235.3267435>
- Nichols, L. G. (2014). A topic model approach to measuring interdisciplinarity at the National Science Foundation. *Scientometrics*, 100(3), 741-754. <https://doi.org/10.1007/s11192-014-1319-2>
- Niekler, A., & Jähnichen, P. (2012). Matching results of latent Dirichlet allocation for text. In N. Rußwinkel, U. Drewitz, & H. Van Rijn (Eds.), *Proceedings of the 11th international conference on cognitive modeling* (pp. 317-322). Universitätsverlag der TU.

- Panichella, A., Dit, B., Oliveto, R., Di Penta, M., Poshynanyk, D., & De Lucia, A. (2013). How to effectively use topic models for software engineering tasks? An approach based on genetic algorithms. In D. Notkin, B. H. C. Cheng, & K. Pohl (Eds.), *Proceedings of the 35th international conference on software engineering: ICSE 2013* (pp. 522-531). IEEE. <https://doi.org/10.1109/ICSE.2013.6606598>
- Quinn, K. M., Monroe, B. L., Colaresi, M., Crespin, M. H., & Radev, D. R. (2010). How to analyze political attention with minimal assumptions and costs. *American Journal of Political Science*, 54(1), 209-228. <https://doi.org/10.1111/j.1540-5907.2009.00427.x>
- Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. In X. Cheng & H. Li (Chairs), *WSDM '15: Proceedings of the eighth ACM international conference on web search and data mining* (pp. 399-408). ACM. <https://doi.org/10.1145/2684822.2685324>
- Strehl, A., & Ghosh, J. (2002). Cluster ensembles - A knowledge reuse framework for combining multiple partitions. *Journal of Machine Learning Research*, 3, 583-617.
- Sun, X., Liu, X., Li, B., Duan, Y., Yang, H., & Hu, J. (2016). Exploring topic models in software engineering data analysis: A survey. In Y. Chen (Ed.), *Proceedings of the 17th IEEE/ACIS international conference on software engineering, artificial intelligence, networking and parallel/distributed computing (SNPD)* (pp. 357-362). IEEE. <https://doi.org/10.1109/SNPD.2016.7515925>
- Teh, Y. W., Jordan, M. I., Beal, M. J., & Blei, D. M. (2006). Hierarchical Dirichlet processes. *Journal of the American Statistical Association*, 101(476), 1566-1581.
- Tirunillai, S., & Tellis, G. J. (2014). Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent Dirichlet allocation. *Journal of Marketing Research*, 51(4), 463-479. <https://doi.org/10.1509/jmr.12.0106>
- Toubia, O., Iyengar, G., Bunnell, R., & Lemaire, A. (2019). Extracting features of entertainment products: A guided latent Dirichlet allocation approach informed by the psychology of media consumption. *Journal of Marketing Research*, 56(1), 18-36. <https://doi.org/10.1177/0022243718820559>
- Wang, Q., Cao, Z., Xu, J., & Li, H. (2012). Group matrix factorization for scalable topic modeling. In W. Hersch (Chair), *SIGIR '12: Proceedings of the 35th international ACM SIGIR conference on research and development in information retrieval* (pp. 375-384). ACM. <https://doi.org/10.1145/2348283.2348335>
- Webber, W., Moffat, A., & Zobel, J. (2010). A similarity measure for indefinite rankings. *ACM Transactions on Information Systems*, 28(4), Article 20. <https://doi.org/10.1145/1852102.1852106>
- Yang, Y., Pan, S., Song, Y., Lu, J., & Topkara, M. (2016). Improving topic model stability for effective document exploration. In G. Brewka (Ed.), *IJCAI'16: Proceedings of the twenty-fifth international joint conference on artificial intelligence* (pp. 4223-4227). AAAI.

# JoEMLS 註釋 (Notes) 暨參考文獻 (References)

## 羅馬化英譯說明

2015年1月31日修訂

1. 本刊針對部分國外西文專業資料庫之引文索引建檔與中文辨讀之需求，凡屬中文稿件之英文摘錄末，特別增列中文羅馬化拼音之「註釋」(或「參考文獻」)一式。
2. 作者(含團體作者)、機構名稱(出版者)、地名(出版地)：依事實與習慣為英譯，如無法查證時，中國大陸地區作者以漢語拼音處理，台灣以威妥瑪拼音(Wade-Giles system)處理。
3. 出版品、篇名：採用(登載於原刊名、篇名等之正式英譯)照錄原則；若原刊文無英譯，則由本刊依漢語拼音音譯著錄之。

e.g. 南京大學學報 *Journal of Nanjing University*

e.g. 中國科學引文數據庫 *Chinese Science Citation Database*

e.g. 玉山國家公園解說志工工作滿足之研究 *Yushan National Park jieshuo zhigong gongzuo manzu zhi yanjiu*

e.g. 教育資料與圖書館學 *Journal of Educational Media and Library Sciences*

4. 混用狀況：地名、機構、人名與其他事實描述，交錯共同構成篇名之一部分時，為避免冗長拼音難以辨讀，可將該名詞中之「地名、機構、人名」依事實與習慣英譯，其餘字詞則由本刊補以漢語拼音處理。

e.g. 「中國科學院與湯姆森科技資訊集團聯手推出中國科學引文索引」

“Chinese Academy of Sciences yu Thomson Scientific Lianshou Tuichu Chinese Science Citation Database”

5. 本刊文章註釋(Notes)或參考文獻(References)羅馬化英譯規則，仍遵循Chicago(Turabian)或APA之精神及原則，進行必要且相對應之編排處理。此羅馬化作業屬權宜措施，不可取代原有正式之引文規範。

6. 羅馬化範例：

範例1－註釋(Notes)

林信成、陳瑩潔、游忠諺，「Wiki協作系統應用於數位典藏之內容加值與知識匯集」，教育資料與圖書館學 43卷，3期(2006)：285-307。【Sinn-Cheng Lin, Ying-Chieh Chen, and Chung-Yen Yu, “Application of Wiki Collaboration System for Value Adding and Knowledge Aggregation in a Digital Archive Project,” *Journal of Educational Media & Library Sciences* 43, no. 3 (2006): 285-307. (in Chinese)】

範例2－參考文獻(References)

林雯瑤、邱炯友(2012)。教育資料與圖書館學四十年之書目計量分析。教育資料與圖書館學，49(3)，297-314。【Lin, Wen-Yau Cathy, & Chiu, Jeong-Yeou (2012) A bibliometric study of the *Journal of Educational Media & Library Sciences*, 1970-2010. *Journal of Educational Media & Library Sciences*, 49(3), 297-314. (in Chinese)】

## About Romanized & Translated Notes/References for Original Text

The main purpose of Romanized and Translated Notes (or References) at the end of English Summary is to assist Western database indexers in identifying and indexing Chinese citations. This Romanization system for transliterating Chinese cannot be a substitute for those original notes or references listed with the Chinese manuscript. The effect of Chinese Romanization for citation remains to be seen.



## Notes for Contributors

1. In accordance with academic regulations and peer review principles, *Journal of Educational Media & Library Sciences* (*JoEMLS*) welcome submissions of manuscripts mainly on topics related to library science, information science and technology, and publishing. The other library related fields such as instructional technology and information communication are also accepted. *JoEMLS* welcome all of manuscripts, but does not publish full translations of non-self-published works.
2. *JoEMLS* adopts open access that does not charge contributors and authors an article processing charge (APC). This journal is a tri-annual journal, published in March, July, and November and published by Tamkang University Press and co-published with the Department of Information & Library Science (DILS) and Chueh Sheng Memorial Library.
3. It is required that all or part of the manuscript submitted to *JoEMLS* has not been published and will not be simultaneously submitted to or published elsewhere. In addition, contributors must ensure that they do not infringe on the copyright of others or compromise academic ethics.
4. Contributions are accepted on the strict understanding that the author is responsible for the accuracy of all contents of the published materials. Publication does not necessarily imply that these are the opinions of the Editorial Board or Editors, nor does the Board or Editors accept any liability for the accuracy of such comment, report, and other technical and factual information. The source should be indicated when the manuscript cites the literature. It is the author's responsibility to obtain written permissions to quote or reproduce material that has appeared in another publication. This includes both copyright and ownership rights, e.g., photographs, illustrations, and data.
5. The authors of any submissions to *JoEMLS* hereby agree that if any submission being accepted by the Journal, then *JoEMLS*, Tamkang University Library, and the Department of Information & Library Science (DILS) shall be authorized to duplicate, publish on the Internet, and publish by any other means for the purpose of non-profit and education use.
6. The authors of any submissions to *JoEMLS* hereby agree that if any submission being accepted by the Journal, then *JoEMLS* shall be authorized to grant a non-exclusive license to the National Central Library for collecting such a submission into the Remote Electronic Access/Delivery System (READncl System), or grant sublicense to other database providers for collecting such a submission into their databases, and to duplicate, publish on the Internet, downloaded, and printed by authorized users of those providers. In addition, the format of submissions may be changed in order to meet the requirements of each database.
7. The copyright of the *JoEMLS* articles belongs to the authors. In addition to the above, readers of *JoEMLS* may use the articles under the Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) model.
8. Manuscripts must be in either Chinese or English. Manuscript lengths requirements:
  - (1) Full & regular research articles in IMRAD format should be between 6,000 and 12,000 words for English (less than 2,000 words for Chinese) in length.
  - (2) Brief communications are of approximately 3,000 words for English (more than 4,000 words for Chinese).
  - (3) Review articles, Observation reports which tend to be review articles and Book review should be more than 4,000 words for English (more than 8,000 words for Chinese).
  - (4) Letters to the Editor should not exceed 1,500 words for English and Chinese in length and may be: comments or criticisms of articles recently published in *JoEMLS*; and preliminary announcements of an original work of importance warranting immediate publications.
9. In order to respect the differences in humanities and sociology research, the Chicago-Turabian Style or the American Psychological Association Author-date format is used in the field of library and information science, please choose one of these two styles to follow. To comply with manuscript, please be clear and detailed in the citation format of notes or references, and the relevant citation format is available on request or refer to the *JoEMLS* website.
10. All Chinese submissions should include the title, abstract (within 300 words), keywords (within 6 words each) in both English and Chinese, no more than 16 tables/figures and 5 appendices. A brief autobiographical note should be provided in both English and Chinese, including full name, post & title, affiliation, e-mail address, and full international contact details. For non-Chinese submissions, provide the above information in English only.
11. If the manuscript is co-authored by more than one author, the First Author should be the equivalent of the Principal Author. The Principal Author must clearly specify the corresponding author and co-authors in a proper sequence. Submission of manuscripts previously published in conference proceedings or revision based on thesis should be clearly indicated in the front page of manuscripts.
12. *JoEMLS* adopts a double-blind peer review policy, revision should be returned to the editor within the deadline for further peer review process. Revision submitted after the period could be rejected or treated as a new manuscript by the Journal.
13. To improve the English readability of the manuscript, a 1,200- to 1,500-word English Summary with appropriate citations and a proof of academic English professional editing must be submitted when the manuscript is accepted for the manuscript in Chinese. A brief English Summary is a supplement to Chinese article. Authors who contribute to the *JoEMLS* in Chinese language would need to supply English Summaries themselves. Such English Summary will carry a disclaimer: "This English Summary and InSight Point are provided by the author(s), and the author(s) have certified or verified that the translation faithfully represents the Chinese version of their own in the journal. It is for convenience of the English users and can be used for reference and citation."
14. The *JoEMLS* editor reserves the right to revise and edit all accepted manuscripts for clarification, and to provide Romanized and Translated Notes (or References) at the end of English Summary, to assist Western database indexers in identifying and indexing Chinese citations. This Romanization system for transliterating Chinese cannot be a substitute for those original notes or references listed with the Chinese manuscript.
15. For the Book Review column, *JoEMLS* is looking for book recommendations as well as individuals who are willing to review the titles. Please contact the *JoEMLS* editor (joemls@mail2.tku.edu.tw) for more information.
16. Corresponding author will receive one free copy of *JoEMLS*. However, authors can find online full-text of PDF format via Open Access mechanism on the websites of *JoEMLS*.
17. Submissions should go through the ScholarOne submission online system (<https://mc.manuscriptcentral.com/joemls>); however, articles submitted to *JoEMLS* (joemls@mail2.tku.edu.tw) as email attachments in one of the following preferred formats, Word or Rich Text Format, are acceptable.

### 紙本訂購資訊 (Subscription for Printed Edition)

Address changes, subscriptions and purchase of back issues, exchanges should be addressed to: Journal of Educational Media & Library Sciences, Department of Information and Library Science, Tamkang University.

Address: 151, Ying-chuan Rd., Tamsui, Taipei 25137, Taiwan

Tel.: +886 2 2621 5656 ext.2382

Fax: +886 2 2620 9931

E-mail: joemls@mail2.tku.edu.tw

A crossed cheque should be made payable to "TAMKANG UNIVERSITY".

一年新臺幣1,500元, 單本新臺幣500元 (臺灣地區)

58卷3期(含)之前的單本訂價新臺幣400元、年度訂費新臺幣1,200元。

Annual subscription (payable in advance) US\$100 (Outside Taiwan)

國外航空郵費另加(Additional charge for airmail outside Taiwan)

US\$15.00 (per year) for America, Europe, Australia & Africa

US\$8.00 (per year) for Japan, Korea, Thailand & the Philippines

US\$6.00 (per year) for Hong Kong & Macao

訂閱本刊, 請以匯款郵局(局號2441285・帳號0388761, 戶名: 教育資料與圖書館學或劃線支票, 戶名抬頭請填寫《教育資料與圖書館學》匯寄訂費, 謝謝。

本刊網頁: <http://joemls.dils.tku.edu.tw/>





Tamkang University Press  
Taiwan

JoEMLS English Summary



ISSN 2309-9100  
ISSN-L 1013-090X



9 771013 090005