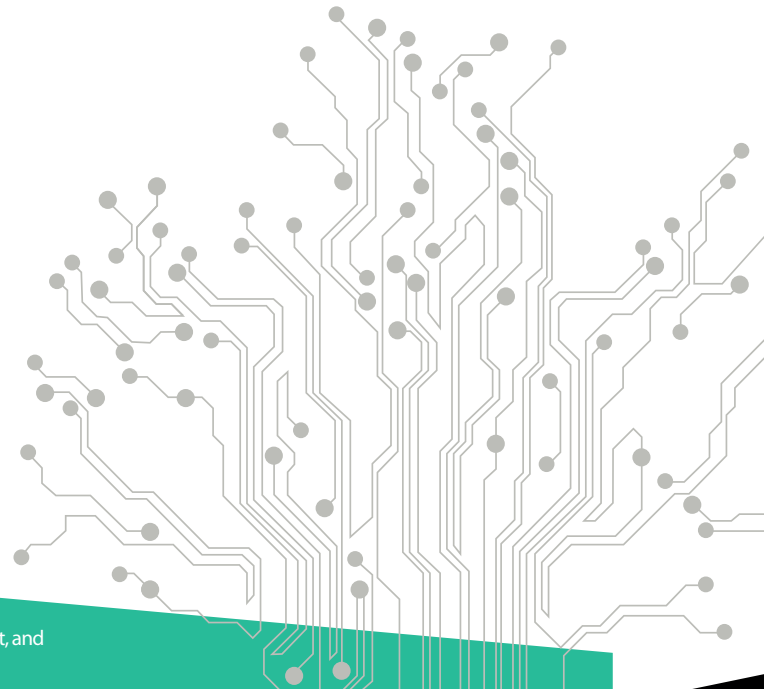


Journal of Information Science Theory and Practice



- 06**
Empirical Comparison of Word Similarity Measures Based on Co-Occurrence, Context, and a Vector Space Model
- 18**
Cognitive Biases and Their Effects on Information Behaviour of Graduate Students in Their Research Projects
- 32**
Factors Influencing Student Satisfaction with the Fieldwork Course in LIS in South Korea
- 45**
Evaluating the Information Access Skills of Students of a College of Education in Ghana
- 55**
Key Traffic Metrics as a Basis to Measure Library Performance
- 68**
Development of Evaluation Perspective and Criteria for the DataON Platform

JISTaP 

General Information

Aims and Scope

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

Indexed/Covered by



Publisher

Korea Institute of Science and Technology Information
66, Hoegi-ro, Dongdaemun-gu, Seoul, Republic of Korea
(T) +82-2-3299-6102
(F) +82-2-3299-6067
E-mail: jistap@kisti.re.kr
URL: <http://www.jistap.org>

Managing Editor: Ji-Young Kim, Eungi Kim

Copy Editor: Ken Eckert

Design & Printing Company: SEUNGLIM D&C

4F, 15, Mareunnae-ro, Jung-gu, Seoul, Republic of Korea
(T) +82-2-2271-2581~2
(F) +82-2-2268-2927
E-mail: sdnc@sdnc.co.kr

Open Access and Creative Commons License Statement

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

Co-Editors-in-Chief

Gary Marchionini
University of North Carolina, USA

Dong-Geun Oh
Keimyung University, Korea

Associate Editors

Kiduk Yang
Kyungpook National University, Korea

Taesul Seo
Korea Institute of Science and Technology Information, Korea

Managing Editors

Ji-Young Kim
Korea Institute of Science and Technology Information, Korea

Eungi Kim
Keimyung University, Korea

Editorial Board

Dan Albertson
University at Buffalo, SUNY, USA

Daniel Martínez Ávila
San Paulo State University, Brazil

Beeraka Ramesh Babu
University of Madras, India

Pia Borlund
University of Copenhagen, Denmark

France Bouthillier
McGill University, Canada

Kathleen Burnett
Florida State University, USA

Lemen Chao
Renmin University of China, China

Ina Fourie
University of Pretoria, South Africa

Boryung Ju
Louisiana State University, USA

Thomas Mandl
Universiät Hildesheim, Germany

Lokman I. Meho
American University of Beirut, Lebanon

Jin Cheon Na
Nanyang Technological University, Singapore

Dan O'Connor
Rutgers University, USA

Helen Partidge
University of Southern Queensland, Australia

M. P. Satija
Guru Nanak Dev University, India

Ou Shiyan
Nanjing University, China

Paul Solomon
University of South Carolina, USA

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

Consulting Editors

Wayne Buente
University of Hawaii, USA

Sujin Butdisuwan
Mahasarakham University,
Thailand

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

M. Krishnamurthy
DRTC, Indian Statistical Institute, India

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

P. Rajendran
SRM University, India

Tsutomu Shihota
St. Andrews University, Japan

Guancan Yang
Renmin University of China, China

Table of Contents

JISTaP

Vol. 8 No. 2 June 30, 2020
Journal of Information Science Theory and Practice • <http://www.jistap.org>

Articles		06
	Empirical Comparison of Word Similarity Measures Based on Co-Occurrence, Context, and a Vector Space Model - Natsuki Kadowaki, Kazuaki Kishida	06
	Cognitive Biases and Their Effects on Information Behaviour of Graduate Students in Their Research Projects - Sara Behimehr, Hamid R. Jamali	18
	Factors Influencing Student Satisfaction with the Fieldwork Course in LIS in South Korea - Sung-Jong Cha, Kyung-Jae Bae	32
	Evaluating the Information Access Skills of Students of a College of Education in Ghana - Franklina Adjoa Yebowaah, Christopher M. Owusu-Ansah	45
	Key Traffic Metrics as a Basis to Measure Library Performance - Olga M. Udartseva	55
	Development of Evaluation Perspective and Criteria for the DataON Platform - Suntae Kim	68
Call for Paper		79
Information for Authors		80

Empirical Comparison of Word Similarity Measures Based on Co-Occurrence, Context, and a Vector Space Model

Natsuki Kadowaki

Graduate School of Library and Information Science,
Keio University, Tokyo, Japan
E-mail: kadowaki.72@keio.jp

Kazuaki Kishida*

Faculty of Letters, Keio University, Tokyo, Japan
E-mail: kz_kishida@keio.jp

ABSTRACT

Word similarity is often measured to enhance system performance in the information retrieval field and other related areas. This paper reports on an experimental comparison of values for word similarity measures that were computed based on 50 intentionally selected words from a Reuters corpus. There were three targets, including (1) co-occurrence-based similarity measures (for which a co-occurrence frequency is counted as the number of documents or sentences), (2) context-based distributional similarity measures obtained from a latent Dirichlet allocation (LDA), nonnegative matrix factorization (NMF), and Word2Vec algorithm, and (3) similarity measures computed from the tf-idf weights of each word according to a vector space model (VSM). Here, a Pearson correlation coefficient for a pair of VSM-based similarity measures and co-occurrence-based similarity measures according to the number of documents was highest. Group-average agglomerative hierarchical clustering was also applied to similarity matrices computed by individual measures. An evaluation of the cluster sets according to an answer set revealed that VSM- and LDA-based similarity measures performed best.

Keywords: word similarity, word clustering, topic model, word embedding

Open Access

Received date: January 23, 2020
Accepted date: April 1, 2020

*Corresponding Author: Kazuaki Kishida

Professor
Faculty of Letters, Keio University, 2-15-45 Mita, Minato-ku,
Tokyo 108-8345, Japan
E-mail: kz_kishida@keio.jp

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

1. INTRODUCTION

Word similarity or semantic similarity between words is often determined to improve the effectiveness of some applications in the field of information retrieval (IR) and other related areas, such as text categorization. For example, suppose that a very short query is used to search a database and thus returns insufficient or irrelevant results. This is because the query did not contain words that accurately represented the user's needs. However, a new query with added words that are similar to the original query words should obtain better results, including more relevant documents that only contain the newly added words. This is referred to as a query expansion technique (Zazo, Figuerola, Berrocal, & Rodríguez, 2005). Word similarity measurements also play important roles in bibliometric studies that attempt to articulate an "intellectual structure" inherent to a set of scientific documents (i.e., a co-word analysis) (e.g., Khasseh, Soheili, Moghaddam, & Chelak, 2017; Ravikumar, Agrahari, & Singh, 2015).

Although human generated thesauri such as WordNet are sometimes used to find similar words, such a task is usually accomplished through automatic corpus processing. In this context, word similarity can be determined in the three following sources: 1) co-occurrence frequencies of two words, 2) the degree to which context words appear around two corresponding target words, and 3) word vectors consisting of weights in individual documents. First, co-occurrence frequencies are typically counted as the number of sentences or documents in which both of two target words appear. Second, a context word information method entails that a feature vector for each word is algorithmically estimated. Here, the Word2Vec algorithm (Mikolov, Yih, & Zweig, 2013) has been widely applied to estimate feature vectors. In comparing these methods, Liebeskind, Dagan, and Schler (2018) called the first "a first-order, co-occurrence-based approach," but referred to the second as "a second-order, distributional similarity approach" (p. 1446). Third, some studies have constructed word vectors by juxtaposing the tf-idf weights of each word in individual documents according to a vector space model (VSM) in the IR field. Here, cosine values of the word vectors are computed as similarity measures.

Researchers have variously used the above three methods to conduct word similarity measures depending on their specific applications and needs. However, it is unclear which similarity measure is better. This is because few studies have systematically compared results between the three. For example, although Dagan, Lee, and Pereira (1999) tried to examine performance of some similarity measures through an experiment on word sense

disambiguation, it did not cover all three types of measures mentioned above. In particular, because Word2Vec is a new algorithm, its effect on computation of word similarity has not yet been known enough, and so it would be worthwhile to compare word similarity computed by Word2Vec with those by traditional methods such as co-occurrence-based and VSM-based approaches. As a result, empirical comparison among the three types of word similarity measure using a common dataset would bring us a new insight on the measures.

Thus, this study empirically investigated results from the above three-word similarity measures through a comparative experiment that implemented a portion of the Reuters Corpus Volume 1 (RCV1) as a test set. More specifically, 50 words were carefully selected as a sample. Similarity matrices were then calculated for the sample using each of the three methods. Here, the purpose was to directly compare similarity values through a Pearson correlation coefficient and evaluate the clustering results obtained by applying a standard hierarchical clustering algorithm to the matrices. This provided new insight on the effectiveness of word similarity measures as a metric for constructing word clusters for use in many applications.

The rest of this paper proceeds as follows: Section 2 reviews the three abovementioned types of word similarity measures and discusses previous related studies, while Section 3 describes this study's experimental method of empirically comparing each measure, and Section 4 discusses the results.

2. COMPUTING WORD SIMILARITY

This section reviews three types of word similarity measures and discusses related studies.

2.1. Co-Occurrence-Based Similarity

The number of documents in which word w_j appears (i.e., document frequency) is denoted by n_j . If n_{jk} indicates the number of documents including both w_j and w_k , then the degree of similarity between them (which is written as s_{jk}) can be computed as follows:

$$s_{jk} = \frac{n_{jk}}{\sqrt{n_j n_k}} \quad (1)$$

which is a cosine measure. It is also possible to calculate the Dice or Jaccard coefficient from the following statistics: n_{jk} , n_j , and n_k . When there are large differences in the document frequencies

between two words, then $n_{jk}/\min(n_j, n_k)$ may be more useful; this is called the overlapped coefficient. For example, if $n_{jk}=10$, $n_j=1,000$, and $n_k=10$, then the overlap coefficient is 1.0, thus indicating that w_k always co-occurs with w_j . However, the cosine coefficient becomes small (i.e., 0.1) because it is affected by the large document frequency of w_j .

The co-occurrence-based similarity method is advantageous for its simplicity in obtaining the data needed for calculation. For example, it is easy to determine three numbers (i.e., n_{jk} , n_j , and n_k) using an IR system (e.g., a database searching service). This is a primary reason that co-word analyses conducted for bibliometric studies usually employ the co-occurrence-based similarity method.

In the IR field, co-occurrence-based similarity is sometimes used for query expansion. Basically, words (or phrases) with high degrees of similarity to an original query word are automatically added to the query under the assumption that documents including words similar to those in a user query are also relevant to the query. When computing similarity, word co-occurrences are usually counted within a set of particular sentences rather than whole documents (e.g., Jing & Croft, 1994; Xu & Croft, 1996). Further, Mandala, Tokunaga, and Tanaka (1999) adopted a variable-length window size when calculating word co-occurrences.

Unfortunately, previous research has found that query expansions achieved through co-occurrence-based similarity have almost no effect on improving search performance in the IR field (e.g., Peat & Willett, 1991). This is partly because newly added words tend to appear in relevant documents that were already retrieved through the original query words. For example, suppose that w_j appears in an original query; even if w_k is added to the query based on its co-occurrence-based similarity to w_j , it is still difficult for w_k to help detect relevant documents in which original w_j does not appear. This is because high co-occurrence-based similarity entails that the two words will co-occur in many documents.

2.2. Context-Based Distributional Similarity

A basic assumption of distributional similarity is that two words are semantically similar if they share common context words. For example, consider the two following sentences:

- (A) Last night, I observed Mercury with my telescope.
- (B) During the night, we observed Jupiter by using Martin's telescope.

It is possible to detect a high similarity between “Mercury” and “Jupiter” by focusing on the common context words of “night,” “observed,” and “telescope.”

Indeed, Chen, Fankhauser, Thiel, and Kamps (2005) measured word similarity by counting context words to automatically construct a thesaurus, while Terra and Clarke (2003) examined word similarity as computed through a co-occurrence-based method involving each target word and a common context word. These studies are an example of measuring word similarity based on simply counting context words. Meanwhile, Lin, Sun, Wu, and Xiong (2016) attempted to represent a word vector from a set of its context words, which was then applied to a clustering of tweets. Likewise, Liebeskind et al. (2018) used a vector of context words to automatically construct a Hebrew thesaurus in which cosine and Dice coefficients were employed. The context word vector is useful for computing a word similarity as discussed later. Finally, Pekar and Staab (2003) measured the distance between two nouns based on their collocation with verbs (a similar idea).

A word similarity measure based on context words can easily be estimated as a cosine coefficient between row or column vectors in a word-by-word matrix, as follows:

$$M = [n_{jk}], j, k = 1, \dots, M,$$

where M is the total number of different words in a target corpus. However, the row (or column) vectors are generally sparse and high dimensional. This sometimes leads to a computational problem that was solved in a study by Schütze and Pedersen (1997), where word clustering and singular-value decomposition (SVD) were used to reduce the issue of high dimensionality (see below for details); each word was finally represented as a 20-dimensional real-valued “thesaurus vector,” which was then used to express documents in an IR algorithm.

The thesaurus vector is a type of word feature vector that allows us to obtain word similarity measures by computing cosine coefficients between vectors, as follows:

$$s_{jk} = \frac{x_j^T x_k}{\|x_j\| \|x_k\|} \quad (2)$$

where x_j and x_k are feature vectors of word w_j and w_k , respectively. Elements of the feature vector are usually real numbers, which provide a distributional representation of each word. This kind of word similarity is often referred to as a distributional similarity.

Because each element of the word feature vector corresponds to a latent topic (or sense) inherent to a given corpus, any mismatches between context words that are caused by surface descriptive variants when detecting common context words is expected to be compensated for through the use of feature

vectors. Take examples (A) and (B) above; if “binoculars” is used as a context word in another sentence, then it is not matched with “telescope” in these sentences even though both words refer to optical devices that are often employed in similar situations. When one or more elements in the word feature vector imply such similar device types, an overlap in sentence contexts would be more successfully reflected by a similarity value that is computed by the feature vectors.

An occurrence frequency of word w_j in document d_i is denoted by x_{ij} ($i = 1, \dots, N; j = 1, \dots, M$), where N indicates the total number of documents in a given corpus. $N \times M$ matrix $\mathbf{X} = [x_{ij}]$ can be decomposed so that $\mathbf{X}^T = \mathbf{A}\mathbf{V}\mathbf{C}^T$, which is an SVD. A diagonal element of \mathbf{V} corresponds to a latent topic (or sense); it is then possible to interpret the j -th row of \mathbf{A} as a feature vector of word w_j (i.e., a thesaurus vector in Schütze & Pedersen, 1997). If extracting only the L largest diagonal elements in \mathbf{V} ($L < \min(N, M)$), then documents are represented by only “major” latent topics. This is an important procedure in latent semantic analysis (LSA) (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990), which was actually applied to measure word similarities when constructing thesauri from corpora (e.g., Lagutina, Larionov, Petryakov, Lagutina, & Paramonov, 2018; Mohsen, Al-Ayyoub, Hmeidi, & Al-Aiad, 2018). When \mathbf{V} is omitted, the decomposition is then as follows:

$$\mathbf{X}^T \cong \mathbf{A}\mathbf{C}^T \quad (3)$$

which is termed a nonnegative matrix factorization (NMF) if \mathbf{A} and \mathbf{C} are nonnegative matrices.

Hofmann (1999) proposed probabilistic LSA (PLSA) that estimates a sequence of probabilities for each word,

$$p(w_j | z_1), p(w_j | z_2), \dots, p(w_j | z_L),$$

from a corpus in which z_j denotes a latent topic ($j = 1, \dots, M; h = 1, \dots, L$). The sequence of real numbers can also be treated as a feature vector of each word. It is then possible to employ other models (e.g., the latent Dirichlet allocation [LDA] developed by Blei, Ng, and Jordan, 2003) to obtain a word feature vector consisting of probability $p(w_j | z_h)$.

Borrowing terms from Waltz and Pollack (1985) that discussed the human understanding of the world from a cognitive science perspective, each element of a word feature vector corresponds to a “microfeature” of a concept that is represented by the word; its value indicates a level of “activation” of the microfeature. Gallant et al. (1992) were inspired by Waltz and Pollack (1985) in their construction of 300-dimensional

vectors to represent words in their own IR system. This may be regarded as an attempt that was relevant to the early stages of the neural network approach, although elements of the vector were determined by simply using data related to word frequencies in a document set.

The distributional representation of words appears in a multilayer model of neural networks when applying the model to textual data. This means that each word in an input text is embedded as a real-valued vector in the network. Word embedding plays a key role in implementing an artificial intelligence system based on a deep learning model when input data are textual. Usually, the distributional representation is computed by applying an algorithm to a large-scale corpus (e.g., Wikipedia) independently from training data that are inherently prepared for the learning process. Word2Vec is a particularly well-known algorithm in this context.

An assumption of Word2Vec is that words appearing with a similar context are similar. As such, any feature vectors estimated through Word2Vec can be used to compute context-based distributional word similarities. For example, the vector can be directly used as \mathbf{x}_j in Equation 2. Both Shunmugam and Archana (2016) and Poostchi and Piccardi (2018) incorporated word feature vectors obtained through Word2Vec into a k-means algorithm for word clustering.

Other algorithms for word embedding have been also developed. For instance, Pennington, Socher, and Manning (2014) proposed an algorithm based on global vectors (GloVe) constructed from global corpus statistics, whereas a shallower window is used to count the co-occurrences of a target word and its context word in Word2Vec. Similarly, Zhao, Liu, Li and Du (2016) explored a word embedding algorithm designed to consider context words within a full document range rather than a given sentence. There are also other word-embedding algorithms that enable users to compute distributional similarities between words.

2.3. Similarity Based on a Vector Space Model

According to the VSM developed by G. Salton’s research group, a word is often represented as a vector of which element is a weight of the word in each document. For example, element v_{ij} is defined as a tf-idf weight, as follows:

$$v_{ij} = x_{ij} \log \frac{N}{n_j} \quad (4)$$

which constructs N -dimensional feature vector $\mathbf{x}_j = [v_{1j}, v_{2j}, \dots, v_{Nj}]^T$ for word w_j ($j = 1, \dots, M$). The cosine coefficient between vectors (i.e., Equation 2) has often been used to create an M

$\times M$ similarity matrix to achieve query expansion in the IR field. The resulting matrix is termed a similarity thesaurus (Qiu & Frei, 1993). If the similarity matrix is denoted by W , then M -dimensional query vector q is modified so that $\tilde{q} = Wq$ (note that matrix W can be also generated through a co-occurrence similarity method).

Zazo et al. (2005) proved that query expansion achieved through the similarity thesaurus had a positive effect when search queries were short, while Mohsen et al. (2018) used a similarity thesaurus to expand queries in the Arabic language. Similarity thesauri have also been used to improve machine learning performance. For example, Xu and Yu (2010) used a similarity thesaurus to detect spam e-mails through a neural network model, while Li, Yang, and Park (2012) used one to enhance the effectiveness of text categorization.

Studies have also explored different similarity measures between word vectors achieved through documents in which the word appears. For example, Jo (2016) represented a word using a set of documents in which it was included. Here, the similarity between two words was measured based on the similarity between the two document sets.

3. EXPERIMENTAL PROCEDURE FOR COMPARING WORD SIMILARITY MEASURES

This study conducted an experiment to empirically identify the characteristics of the three types of word similarity measures reviewed in Section 2, including the a) co-occurrence-based, b) context-based distributional, and c) VSM-based vector type used to construct similarity thesauri. More specifically, the real values of these similarity measures were computed from a set of documents extracted from the RCV1 for direct comparison.

3.1. Data

Word similarities were computed using a total of 6,374 records to which a single topical code was assigned in a set of news articles published between August 20-31, 1996. This was used as a test set to explore the document clustering algorithm in an experiment by Kishida (2011). However, this study placed no special implications on the restringing of articles to which a single topical code was assigned. Rather, this was only done because the articles had been intensively checked by an author prior to this experiment.

Nouns, verbs, adjectives, adverbs, cardinal numbers, and foreign words were extracted from the headlines and main texts of each news article using version 3.9.2 of the Stanford POS tagger (Toutanova, Klein, Manning, & Singer, 2003).

3.2. Comparative Analysis Methods

Word similarity measures were empirically compared via two metrics of proximity or difference, as follows:

- (A) Pearson correlation coefficient values, which were calculated between two matrices of similarity measures
- (B) normalized mutual information (nMI) scores, which were calculated between two sets of document groups that were generated from individual similarity measures through a clustering algorithm

The Pearson correlation coefficient was standardly computed after vectorizing target matrices, meaning that the coefficient was directly calculated from corresponding pairs of $M(M-1)/2$ elements of two matrices. While this was a direct comparison, an indirect comparison was achieved by using the nMI scores between clustering results. An examination of clustering results is useful for applications in which word clusters have important functions. Although several ways of normalization can be used to define nMI (Kishida, 2014), this experiment normalized MI scores according to maximum entropy values.

Each word was stemmed through the Porter algorithm. A resulting list of word stems was then created in descending order of document frequency (i.e., n_j). Among those appearing in 100 or more news articles, the authors then intentionally selected 50 words that unambiguously represented a concept belonging to only one of the five following topics: (a) the economy, (b) politics, (c) crime, (d) war, and (e) sports. These categories were determined after carefully examining the top-ranked word stem list; the authors did not find any topic to which 10 or more words belong other than the five topics in the dataset. Table 1 shows the 50 total word stems across all five topics.

It was expected that a comparative analysis would be easier to conduct when only considering unambiguous words even

Table 1. All 50 word stems selected for comparison

Economy	Politics	Crime	War	Sports
bond	communist	arrest	armi	champion
cash	congress	crimin	attack	cricket
debt	democrat	jail	bomb	game
dollar	diplomat	kidnap	fight	leagu
export	govern	law	guerrilla	player
import	minist	legal	militari	soccer
invest	parliament	murder	rebel	sport
market	polit	prison	soldier	team
monet	politician	prosecutor	troop	tenni
trade	republ	victim	war	tournament

though it was also important to examine ambiguous words (i.e., those with multiple meanings). Although some words were equally related to two topic categories (e.g., “bomb” may be related to both “crime” and “war”), the authors considered that the 50 stems were adequate for this experiment. Note that the five groups shown in Table 1 could be used as a ground truth, thus allowing an external evaluation of all clustering results. This is another benefit of specifically selecting a set of unambiguous words as a target group.

3.3. Computing Word Similarity Measures

Table 2 shows the word similarity measures examined in this study’s experiment. The co-occurrence-based similarity was computed through Equation 1 in two cases where the numbers of sentences and documents were used for n_{jk} , n_j , and n_k , respectively. On the other hand, the context-based distributional similarity was calculated through the NMF, LDA, and Word2Vec frameworks, which generated word feature vectors \mathbf{x}_j ($j = 1, \dots, M$) in Equation 2, respectively. The number of dimensions of the feature vectors was commonly set to 100, thus corresponding to the number of latent topics in both NMF and LDA (i.e., $L = 100$). Word feature vectors were also constructed by using Equation 4, from which VSM-based word similarity was calculated by Equation 2. Note that the NMF and LDA algorithms were executed after removing words (stems) that only appeared in one article. They were also deleted from word feature vectors in VSM to maintain the same condition.

This experiment only used a cosine measure in Equation 1 or 2, although it was also possible to calculate the Dice, Jaccard, or overlapped coefficients. Further, information-theoretic measures (e.g., the point-wise mutual information, Kullback-Leibler

divergence, and Jensen-Shannon divergence) are often employed when measuring word similarities used for natural language processing (NLP) (Dagan et al., 1999). However, these were outside the scope of this study, which also excluded nonsymmetric measures explored in the NLP field (Kotlerman, Dagan, Szpektor, & Zhitomirsky-Geffet, 2010). As such, this experiment solely focused on the cosine measure widely applied throughout the IR and bibliometrics (scientometrics) fields.

After computing s_{jk} in Equation 1 or 2 for the 50 stems shown in Table 1 ($j, k = 1, \dots, 50$) according to the individual definitions of six similarity measures in Table 2, a Pearson correlation coefficient was calculated for each pair. Because $M = 50$, the sample size in calculation of the coefficient was 1,225 ($= 50 \times 49 \div 2$). Next, a group-average agglomerative hierarchical clustering (AHC) was executed for each set of similarity values s_{jk} after converting it to a distance metric so that $1.0 - s_{jk}$. This experiment only used the group-average method because it clearly outperformed a complete linkage method during a preliminary analysis. Classical multidimensional scaling (MDS) was also partly used to observe visual proximities between the words derived from each word similarity measure.

3.4. Experimental System

For the computational process, the Word2Vec algorithm was executed for the test set using an Apache Spark module (Word2Vec class). The `hclust` and `cmdscale` functions of R packages (version 3.6.1) were then applied for the AHC and MDS, respectively. Other computer processing modules were constructed using the Java language. Probability $p(w_j | z_k)$ in the LDA model was estimated via Gibbs sampling (Griffiths & Steyvers, 2004). More specifically, in the r -th iteration of

Table 2. Word similarity measures

Word similarity s_{jk}	Acronym
(A) Co-occurrence-based similarity	
(i) Using the number of sentences in Equation 1	CoocS
(ii) Using the number of documents in Equation 1	CoocD
(B) Context-based distributional similarity	
(i) Using a row of \mathbf{A} in NMF (see Equation 3) as feature vector \mathbf{x}_j in Equation 2	NMF
(ii) Using sequence $p(w_j z_1), \dots, p(w_j z_L)$ obtained by LDA as feature vector \mathbf{x}_j in Equation 2	LDA
(iii) Using word embedding through the Word2Vec algorithm as feature vector \mathbf{x}_j in Equation 2	W2V
(C) VSM-based similarity for similarity thesauri	
(i) Using N dimensional vector of which element is a tf-idf weight in Equation 4 as feature vector \mathbf{x}_j in Equation 2	VSM

L refers to the number of columns in \mathbf{A} , which corresponds to the number of latent topics in LDA, while N denotes the number of documents. NMF, nonnegative matrix factorization; LDA, latent Dirichlet allocation; VSM, vector space model.

Table 3. Predetermined parameters

Word similarity		Predetermined parameters
Co-occurrence-based	CoocS, CoocD	None
Distributional	NMF	Iterations: 100
	LDA	Hyperparameters: $\alpha = 0.1, \beta = 0.01$ Iterations: 2,100 (burn-in period: 100)
	W2V	Max iterations: 2,000, Window-size: 5
VSM-based	VSM	$N = 6,374$

CoocS, co-occurrence-based similarity according to number of sentences; CoocD, co-occurrence-based similarity according to number of documents; NMF, nonnegative matrix factorization; LDA, latent Dirichlet allocation; W2V, Word2Vec; VSM, vector space model.

the sampling, $p_r(w_j | z_k)$ was computed as a percentage of w_j in tokens to which the k -th latent topic was allocated. After R iterations, $p(w_j | z_k)$ was estimated as an average so that $p(w_j | z_k) = R^{-1} \sum_{r=1}^R p_r(w_j | z_k)$ except for iterations during the burn-in period. Finally, NMF was obtained via an algorithm developed by Lee and Seung (1999) under the condition that a row of A was normalized by its norm. The predetermined parameters for each process are shown in Table 3.

4. EXPERIMENTAL RESULTS OF COMPARING WORD SIMILARITY MEASURES

A total of 26,594 stems of nouns, verbs, adjectives, adverbs, cardinal numbers, and foreign words were obtained from 6,374 news articles after removing those that only appeared in one article ($N = 6,374$; $M = 26,594$). Thus, a total of 1,237,831 tokens were included in our test set (i.e., collection length), meaning that the average document length was 194.2. This set of sentences and documents was then used to calculate word similarity measures (Table 2) for the 50 selected words (stems) (Table 1).

4.1. Comparison Based on a Pearson Correlation Coefficient

Table 4 shows the Pearson correlation coefficient values between the six word-similarity measures computed from each set of 1,225 similarity pairs between all 50 stems, as described in Section 3.3. The closest similarity measures were the co-occurrence-based similarity according to number of documents (CoocD) and VSM-based similarity (VSM), of which the value was 0.939. On the other hand, the value was lowest (0.550) between the co-occurrence-based similarity according to number of sentences (CoocS) and NMF-based similarity (NMF).

Table 4. Pearson correlation coefficient between word similarity measures

	CoocS	CoocD	NMF	LDA	W2V
CoocD	0.849				
NMF	0.550	0.681			
LDA	0.689	0.766	0.607		
W2V	0.592	0.645	0.562	0.664	
VSM	0.909	0.939	0.645	0.795	0.644

CoocS, co-occurrence-based similarity according to number of sentences; CoocD, co-occurrence-based similarity according to number of documents; NMF, nonnegative matrix factorization; LDA, latent Dirichlet allocation; W2V, Word2Vec; VSM, vector space model.

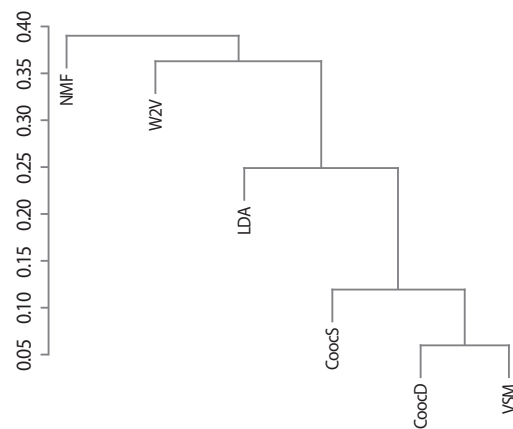


Fig. 1. Agglomerative hierarchical clustering results according to a Pearson correlation coefficient. NMF, nonnegative matrix factorization; W2V, Word2Vec; LDA, latent Dirichlet allocation; CoocS, co-occurrence-based similarity according to number of sentences; CoocD, co-occurrence-based similarity according to number of documents; VSM, vector space model.

Fig. 1 shows the result of group-average AHC executed for the correlation matrix shown in Table 4 after each correlation value was simply converted to a distance metric (i.e., = 1.0–correlation). The relatedness among similarity measures is clearly demonstrated through the dendrogram in Fig. 1, which indicates that the co-occurrence-based similarity (CoocS and CoocD) and VSM-based similarity measure formed a group. These are traditional measures that have been used for a very long time in the IR field. On the other hand, the LDA-based similarity measure was relatively near the group, while the Word2Vec- and NMF-based similarity measures differed from those.

4.2. Comparing Clustering Results

As an example, two MDS plots for the data from the VSM-based and Word2Vec-based similarity measures are shown in

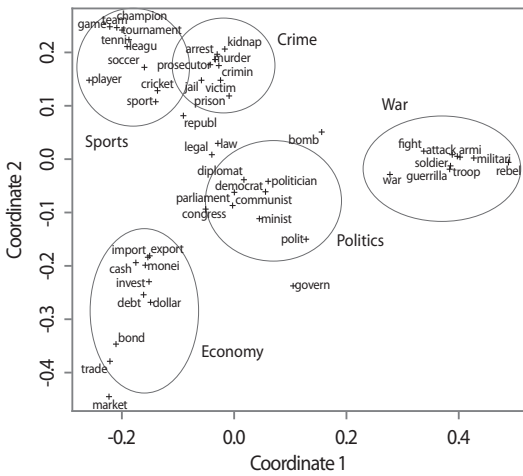


Fig. 2. Multidimensional scaling plot (vector space model-based similarity measure).

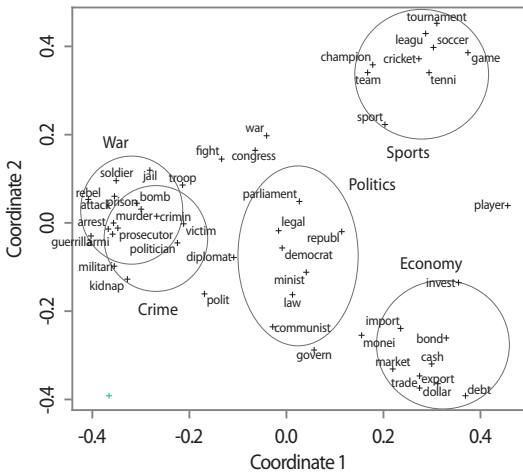


Fig. 3. Multidimensional scaling plot (Word2Vec-based similarity measure).

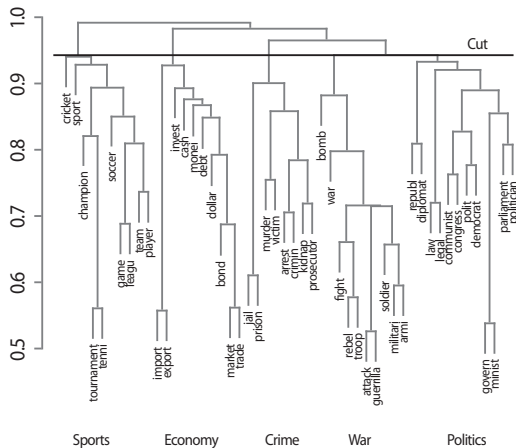


Fig. 4. Cut of the dendrogram according to the vector space model-based similarity.

Figs. 2 and 3, respectively. Although the crime and war word groups overlapped in the MDS plot according to the Word2Vec-based similarity measure (Fig. 3), there appeared to be no large differences between the two maps.

A typical AHC clustering result is shown in Fig. 4 through a dendrogram that was drawn using data from the VSM-based similarity measure, in which five clusters were generated through a cut operation (i.e., the cutree function of the R package). Clustering was successful because the words in each subtree (except for the words “law” and “legal”) corresponded to one of the groups shown in Table 1.

Table 5 shows the nMI scores among the word-group sets (i.e., cluster sets) generated by cutting the dendrograms obtained from data of individual similarity measures when the number of clusters was set to be five (i.e., $H = 5$ in which H means the number of word clusters). As shown, a set of the five groups in Table 1 were also included as an “Answer” that could be employed as a ground truth provided by human annotators. According to this ground truth, the most successful clustering result was obtained from the VSM-based similarity measure, followed by the co-occurrence and LDA-based measures (CoocS, CoocD, and LDA). However, the NMF-based measure provided the lowest nMI score with the “Answer” in this experiment.

The exact number of clusters is unknown in many situations involving document clustering. This experiment therefore attempted to increase the numbers of clusters from five to 10 (i.e., $H = 5,6,7,8,9,10$) although 50 words were intentionally selected from five total topics. Fig. 5 shows the nMI scores between cluster sets and the “Answer” according to the number of clusters. Here, it is evident that the VSM- and LDA-based similarity measures produced better overall clustering results. Note that the number of clusters in the “Answer” was always fixed to five when calculating the nMI scores shown in Fig. 5.

Table 5. Normalized mutual information scores between word-group sets according to word similarity measures ($H = 5$)

	CoocS	CoocD	NMF	LDA	W2V	VSM
CoocD	0.930					
NMF	0.630	0.644				
LDA	0.930	1.000	0.644			
W2V	0.753	0.691	0.583	0.691		
VSM	0.764	0.764	0.597	0.764	0.699	
Answer	0.787	0.787	0.590	0.787	0.715	0.957

H means the number of word clusters. CoocS, co-occurrence-based similarity according to number of sentences; CoocD, co-occurrence-based similarity according to number of documents; NMF, nonnegative matrix factorization; LDA, latent Dirichlet allocation; W2V, Word2Vec; VSM, vector space model.

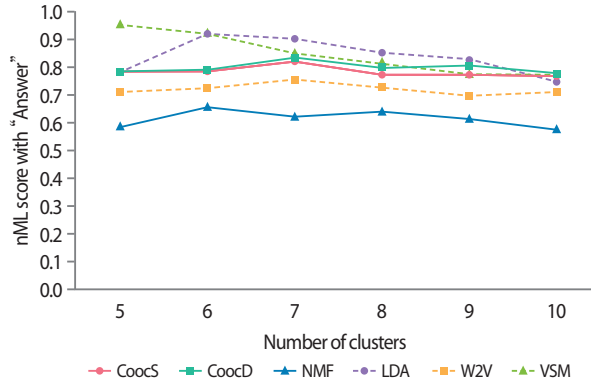


Fig. 5. Normalized mutual information (nMI) scores with "Answer" cluster sets ($H = 5, 6, 7, 8, 9, 10$). H means the number of word clusters. CoocS, co-occurrence-based similarity according to number of sentences; CoocD, co-occurrence-based similarity according to number of documents; NMF, nonnegative matrix factorization; LDA, latent Dirichlet allocation; W2V, Word2Vec; VSM, vector space model.

Table 6. Average normalized mutual information scores between word group sets according to word similarity measures ($H = 5$ to 10)

	CoocS	CoocD	NMF	LDA	W2V	VSM
CoocD	0.853					
NMF	0.664	0.676				
LDA	0.833	0.849	0.650			
W2V	0.713	0.683	0.615	0.743		
VSM	0.788	0.821	0.663	0.895	0.714	
Answer	0.786	0.802	0.620	0.843	0.726	0.851

H means the number of word clusters. CoocS, co-occurrence-based similarity according to number of sentences; CoocD, co-occurrence-based similarity according to number of documents; NMF, nonnegative matrix factorization; LDA, latent Dirichlet allocation; W2V, Word2Vec; VSM, vector space model.

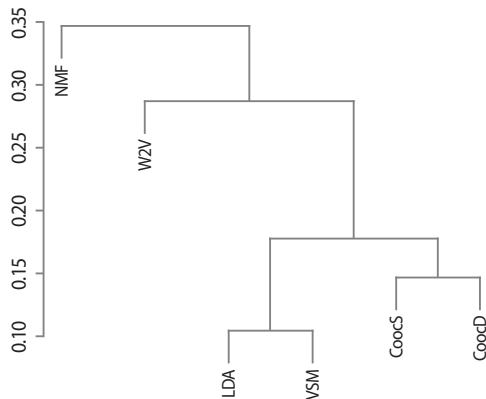


Fig. 6. Results of agglomerative hierarchical clustering according to average normalized mutual information scores. NMF, negative matrix factorization; W2V, Word2Vec; LDA, latent Dirichlet allocation; VSM, vector space model; CoocS, co-occurrence-based similarity according to number of sentences; CoocD, co-occurrence-based similarity according to number of documents.

The averages of nMI scores from $H = 5$ to 10 are shown in Table 6. That is, while the nMI scores in Table 5 only show cases in which $H = 5$, Table 6 provides the results of averaging the six nMI scores for each pair of cluster sets. Fig. 6 also shows a dendrogram that was generated using a set of the average nMI scores as a similarity matrix (except for the "Answer"). The dendrogram in Fig. 6 is similar to that shown in Fig. 1, which was drawn according to Pearson correlation coefficient values. The co-occurrence-, LDA-, and VSM-based similarity measures form a cluster that is remotely located from the Word2Vec- and NMF-based measures. Specifically, a pair of LDA- and VSM-based similarity measures was strongly related due to the generation of similar cluster sets. Likewise, the two co-occurrence-based measures of CoocS and CoocD were closely located in the dendrogram. This was predictable because the only difference between them was found in the ranges of textual data when counting co-occurrence frequencies.

5. DISCUSSION

As shown in Figs. 1 and 6, the co-occurrence- (CoocS and CoocD), LDA-, and VSM-based similarity measures represented relatively similar relationships between words, while the Word2Vec and NMF algorithms provided different similarities. Because the multiplication of tf-idf values (v_{ij}) of a word between two documents becomes zero if the word does not appear in one document, it is easy to conjecture a resemblance between the VSM- and co-occurrence-based similarity measures in view of the number of documents (CoocD).

If all tf values within a given corpus are commonly 1 (i.e., $x_{ij} = 1$), then the inner product of the VSM-based similarity measure in the numerator of Equation 2 is as follows:

$$\mathbf{x}_j^T \mathbf{x}_k = \sum_{i=1}^N x_{ij} \log \frac{N}{n_j} x_{ik} \log \frac{N}{n_k} = n_{jk} \times \log \frac{N}{n_j} \times \log \frac{N}{n_k} = n_{jk} r_j r_k \quad (5)$$

where $r_j = \log \frac{N}{n_j}$ and $r_k = \log \frac{N}{n_k}$. Because $x_{ij} = 1$, the cosine measure is computed as follows:

$$s_{jk} = n_{jk} \times \frac{r_j}{\sqrt{n_j r_j^2}} \times \frac{r_k}{\sqrt{n_k r_k^2}} = \frac{n_{jk}}{\sqrt{n_j} \sqrt{n_k}},$$

which is equal to the co-occurrence-based similarity measure. If x_{ij} changes to $x_{ij} + 1$ in document d_i , then the difference of the inner product in Equation 5 is as follows:

$$\Delta(\mathbf{x}_j^T \mathbf{x}_k) = \log \frac{N}{n_j} \times x_{1k} \log \frac{N}{n_k} \quad (6)$$

Equation 6 suggests that the inner product increases depending on the tf of the other word and idf values of two words. In fact, the amount of change in s_{jk} is more complicated because a value of the denominator of Equation 2 (i.e., $\|\mathbf{x}_j\| \|\mathbf{x}_k\|$) also varies with $x_{1j} \rightarrow x_{1j} + 1$.

Further, the definitions of CoocS and CoocD become equivalent if all documents each consist of single sentences (i.e., short texts). In a corpus of such short documents, the tf of each word would be near 1, for which there would be only a small difference between the VSM- and co-occurrence-based similarity measures. Although the news articles with headlines and main texts used in this study were not short (average document length was 194.2), the values of the VSM- and co-occurrence-based similarity measures were relatively similar (Fig. 1).

Regarding the clustering results, the VSM- and LDA-based similarity measures generated similar cluster sets (Fig. 6); these were also near the answer set (Table 6). The true answer of clustering words is highly dependent on the target corpus, meaning that the grouping of 50 words shown in Table 1 is not always true when used as the “Answer.” For example, there may be a corpus in which “cash” appears in only documents categorized as “sports.” Clustering results also change according to the clustering algorithm. The better performance shown from the VSM- and LDA-based similarity measures must be interpreted in consideration of these limitations.

News articles tend to contain a definite target topic (e.g., the economy or sports). This means that documents (not words) were clearly partitioned into topic groups. The LDA model included a document-based probability ($p(z_k | d_i)$), while VSM-based similarity was also measured from document vectors. Such document-linked architecture may have contributed to the better performance found with the LDA- and VSM-based similarity measures in the experiment using the news article set. Further, the Word2Vec algorithm only checked for co-occurrences within a small window in each text (the window size was set to five in this experiment). As such, the benefits of using topically cleared news articles could not be incorporated, thus possibly resulting in the lower performance found in this experiment.

6. CONCLUSION

This paper reported on the results of an experiment designed to examine word similarity measures using a portion of

the RCV1. We thereby found similar values between a co-occurrence-based similarity measure and one based on tf-idf weights (i.e., a VSM-based measure). We also compared cluster sets generated by the average-group AHC algorithm from individual similarity matrices, thereby finding better clustering results through the VSM- and LDA-based similarity measures. On the other hand, the Word2Vec- and NMF-based similarity measures differed from the other tested measures. From a practical viewpoint, this experiment suggested that word similarity measures computed by LDA and VSM are expected to enhance effectiveness of query expansion or related applications because it is considered that they can identify ‘true’ similar words more correctly. While VSM is known to generate such effective similarities, it is interesting that LDA also works well.

As discussed above, this study had some limitations. For one, we only used a set of news articles to compute similarity measures. Two, only an AHC algorithm was applied to the similarity matrices thus obtained. Three, symmetric similarity measures other than the cosine and nonsymmetric similarity measures were outside the scope of this study’s experiment. As such, future studies should conduct experiments from both the theoretical and empirical viewpoints to gain a deeper understanding of the tested word similarity measures.

REFERENCES

- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Chen, L., Fankhauser, P., Thiel, U., & Kamps, T. (2005, October 31–November 5). Statistical relationship determination in automatic thesaurus construction. In O. Herzog, H.-J. Schek, N. Fuhr, A. Chowdhury, & W. Teiken (Eds.), *Proceedings of the 14th ACM International Conference on Information and Knowledge Management (CIKM '05)* (pp. 267-268). Association for Computing Machinery.
- Dagan, I., Lee, L., & Pereira, F. C. N. (1999). Similarity-based models of word cooccurrence probabilities. *Machine Learning*, 34(1-3), 43-69.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407.
- Gallant, S., Hecht-Nielsen, R., Caid, W., Qing, K., Carleton, J., & Sudbeck, D. (1992, November 4-6). HNC’s MatchPlus System. In D. K. Harman (Ed.), *The First Text REtrieval*

- Conference (TREC-1)* (pp. 107-111). National Institute of Standards and Technology.
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America*, 101(suppl 1), 5228-5235.
- Hofmann, T. (1999, August 15-19). Probabilistic latent semantic indexing. In F. Gey, M. A. Hearst, & R. Tong (Eds.), *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '99)* (pp. 50-57). Association for Computing Machinery.
- Jing, Y., & Croft, W. B. (1994, October 11-13). An association thesaurus for information retrieval. In J.-L. F. Brentano & F. Seitz (Eds.), *RIAO '94: Intelligent Multimedia Information Retrieval Systems and Management – Vol. 1* (pp. 146-160). Le Centre de Hautes Etudes Internationales d'Informatique Documentaire.
- Jo, T. (2016, July 25-28). String vector based AHC as approach to word clustering. In R. Stahlbock & G. M. Weiss (Eds.), *Proceedings of the International Conference on Data Mining DMN'16* (pp. 133-138). Lancaster Centre for Forecasting.
- Khasseh, A. A., Soheili F., Moghaddam, H. S., & Chelak, A. M. (2017). Intellectual structure of knowledge in iMetrics: A co-word analysis. *Information Processing & Management*, 53(3), 705-720.
- Kishida, K. (2011). Double-pass clustering technique for multilingual document collections. *Journal of Information Science*, 37(3), 304-321.
- Kishida, K. (2014). Empirical comparison of external evaluation measures for document clustering by using synthetic data. *IPSJ SIG Technical Report*, 2014-IFAT-113, 1-7.
- Kotlerman, L., Dagan, I., Szpektor, I., & Zhitomirsky-Geffet, M. (2010). Directional distributional similarity for lexical inference. *Natural Language Engineering*, 16(4), 359-389.
- Lagutina, K., Larionov, V., Petryakov, V., Lagutina, N., & Paramonov, I. (2018, November 13-16). Sentiment classification of Russian texts using automatically generated thesaurus. In S. Balandin, T. S. Cinotti, F. Viola, & T. Tyutina (Eds.), *Proceedings of the 23rd Conference of Open Innovations Association FRUCT* (pp. 217-222). FRUCT Oy.
- Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401, 788-791.
- Li, C. H., Yang, J. C., & Park, S. C. (2012). Text categorization algorithms using semantic approaches, corpus-based thesaurus and WordNet. *Expert Systems with Applications*, 39(1), 765-772.
- Liesbeskind, C., Dagan, I., & Schler, J. (2018, May 7-12). Automatic thesaurus construction for modern Hebrew. In N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis, & T. Tokunaga (Eds.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (pp. 1446-1451). European Language Resources Association.
- Lin, H., Sun, B., Wu, J. & Xiong, H. (2016, June 24-26). Topic detection from short text: A term-based consensus clustering method. In B. Yang (Ed.), *2016 13th International Conference on Service Systems and Service Management (ICSSSM 2016)* (pp. 1-6). IEEE.
- Mandala, R., Tokunaga, T., & Tanaka, H. (1999, August 15-19). Combining multiple evidence from different types of thesaurus for query expansion. In F. Gey, M. A. Hearst, & R. Tong (Eds.), *SIGIR '99: Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 191-197). Association for Computing Machinery.
- Mikolov, T., Yih, W. T., & Zweig, G. (2013, June 9-14). Linguistic regularities in continuous space word representations. In L. Vanderwende, H. Daumé III, & K. Kirchhoff (Eds.), *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 746-751). Association for Computational Linguistics.
- Mohsen, G., Al-Ayyoub, M., Hmeidi, I., & Al-Aiad, A. (2018, April 3-5). On the automatic construction of an Arabic thesaurus. *2018 9th International Conference on Information and Communication Systems (ICICS)* (pp. 243-247). IEEE.
- Peat, H. J., & Willett, P. (1991). The limitations of term co-occurrence data for query expansion in document retrieval systems. *Journal of the American Society for Information Science*, 42(5), 378-383.
- Pekar, V., & Staab, S. (2003, April 12-17). Word classification based on combined measures of distributional and semantic similarity. In A. Copestake & J. Hajič (Eds.), *EACL '03: Proceedings of the Tenth Conference on European Chapter of the Association for Computational Linguistics – Vol. 2* (pp. 147-150). Association for Computational Linguistics.
- Pennington, J., Socher, R., & Manning, C. D. (2014, October 25-29). GloVe: Global vectors for word representation. In A. Moschitti, B. Pang, & W. Daelemans (Eds.), *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532-1543). Association for Computational Linguistics.

- Poostchi, H., & Piccardi, M. (2018, December 10-12). Cluster labeling by word embeddings and WordNet's hypernymy. In S. M. Kim & X. Zhang (Eds.), *Proceedings of the Australasian Language Technology Association Workshop 2018* (pp. 66-70). Association for Computational Linguistics.
- Qiu, Y., & Frei, H. -P. (1993, June 27-July 1). Concept based query expansion. In R. Korfhage, E. M. Rasmussen, & P. Willett (Eds.), *SIGIR '93: Proceedings of the 16th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 160-169). Association for Computing Machinery.
- Ravikumar, S., Agrahari, A., & Singh, S. N. (2015). Mapping the intellectual structure of scientometrics: A co-word analysis of the journal *Scientometrics* (2005-2010). *Scientometrics*, 102(1), 929-955.
- Schütze, H., & Pedersen, J. O. (1997). A cooccurrence-based thesaurus and two applications to information retrieval. *Information Processing & Management*, 33(3), 307-318.
- Shunmugam, D. A., & Archana, P. (2016). An empirical investigation of word clustering techniques for natural language understanding. *International Journal of Engineering Science and Computing*, 6(10), 2637-2646.
- Terra, E. L., & Clarke, C. L. A. (2003, May 27-June 1). Frequency estimates for statistical word similarity measures. In M. Hearst & M. Ostendorf (Eds.), *NAACL '03: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology – Vol. 1* (pp. 165-172). Association for Computational Linguistics.
- Toutanova, K., Klein, D., Manning, C. D., & Singer, Y. (2003, May 27-June 1). Feature-rich part-of-speech tagging with a cyclic dependency network. In M. Hearst & M. Ostendorf (Eds.), *NAACL '03: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology – Vol. 1* (pp. 173-180). Association for Computational Linguistics.
- Waltz, D. L., & Pollack, J. B. (1985). Massively parallel parsing: A strongly interactive model of natural language interpretation. *Cognitive Science*, 9(1), 51-74.
- Xu, H., & Yu, B. (2010). Automatic thesaurus construction for spam filtering using revised back propagation neural network. *Expert Systems with Applications*, 37(1), 18-23.
- Xu, J., & Croft, W. B. (1996, August 18-22). Query expansion using local and global document analysis. In H.-P. Frei, D. Harman, P. Schäuble, & R. Wilkinson (Eds.), *SIGIR '96: Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 4-11). Association for Computing Machinery.
- Zazo, A. F., Figuerola, C. G., Berrocal, J. L. A., & Rodríguez, E. (2005). Reformulation of queries using similarity thesauri. *Information Processing & Management*, 41(5), 1163-1173.
- Zhao, Z., Liu, T., Li, B., & Du, X. (2016, August 29-September 2). Cluster-driven model for improved word and text embedding. In G. A. Kaminka, M. Fox, P. Bouquet, E. Hüllermeier, V. Dignum, & (Eds.), *ECAI'16: Proceedings of the Twenty-second European Conference on Artificial Intelligence* (pp. 99-106). IOS Press.

Cognitive Biases and Their Effects on Information Behaviour of Graduate Students in Their Research Projects

Sara Behimehr

Department of Library and Information Science,
Kharazmi University, Tehran, Iran
E-mail: behimehr.sara@gmail.com

Hamid R. Jamali*

School of Information Studies,
Charles Sturt University, Wagga Wagga, NSW, Australia
E-mail: h.jamali@gmail.com

ABSTRACT

Cognitive biases can influence human information behaviour and decisions made in information behaviour and use. This study aims to identify the biases involved in some aspects of information behaviour and the role they play in information behaviour and use. Twenty-five semi-structured face-to-face interviews were conducted in an exploratory qualitative study with graduate (MA and PhD) students who were at the stage of their dissertation/thesis research. Eisenberg & Berkowitz Big6™ Skills for Information Literacy was adopted as a framework for interviews and the analysis was done using grounded theory coding method. The findings revealed the presence of twenty-eight biases in different stages of information behaviour, including availability bias (affects the preference for information seeking strategies), attentional bias (leads to biased attention to some information), anchoring effect (persuades users to anchor in special parts of information), confirmation bias (increases the tendency to use information that supports one's beliefs), and choice-supportive bias (results in confidence in information seeking processes). All stages of information seeking were influenced by some biases. Biases might result in a lack of clarity in defining the information needs, failure in looking for the right information, misinterpretation of information, and might also influence the way information is presented.

Keywords: cognitive biases, information behaviour, information use, research, students

Open Access

Received date: October 30, 2019
Accepted date: May 10, 2020

*Corresponding Author: Hamid R. Jamali
Senior Lecturer
School of Information Studies, Charles Sturt University,
Locked Bag 588, Wagga Wagga, NSW 2678, Australia
E-mail: h.jamali@gmail.com

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

1. INTRODUCTION

Information behaviour involves processing information and making decisions, which are cognitive processes, and therefore it is interesting to study them from a cognitive approach (Fidel, 2012, p. 129). One of the cognitive phenomena that might be involved in information behaviour is cognitive biases. We already know that people might not follow rational or normative models of decision making and that biases can have a negative impact on decision outcomes (Kahneman, Slovic, & Tversky, 1982).

The term 'cognitive bias' was introduced in the 1970s in order to "describe people's systematic but purportedly flawed patterns of responses to judgment and decision problems" (Wilke & Mata, 2012, p. 531). Cognitive biases occur because human cognition has limited abilities to properly attend to and process all the information that is available (Kruglanski & Ajzen, 1983). In the past, researchers have discovered that humans rely on some simplifying strategies, or rules of thumb (called heuristics), for decision making. As Bazerman and Moore (2012, p. 6) stated, heuristics are standard rules that implicitly direct our judgment. They are mechanisms for coping with the complex environment surrounding our decisions. Although they are generally helpful, using them sometimes results in serious errors. In a decision-making context, bias is not being prejudicial and unwilling to have an open mind; rather, it refers to "an unconscious inclination toward a particular outcome or belief that can affect how humans search for and process information" (Schmutte & Duncan, 2014, p. 69). Cognitive biases are "similar to optical illusions in that the error remains compelling even when one is fully aware of its nature. Awareness of the bias, by itself, does not produce a more accurate perception. Cognitive biases, therefore, are, exceedingly difficult to overcome" (Heuer, 2007, p. 112). So far researchers have identified a large number of cognitive biases, and the list compiled by Benson (2016) includes about 180 such biases which are categorised into four main groups, including too much information, not enough meaning, a need to act fast, and what should we remember? These are four main issues that affect the way our brain processes information. For instance, 'too much information' or 'what we should remember' mean that our brain uses tricks to pick the bits of information that are likely to be useful now or in the future.

Knowing which cognitive biases are involved and how they might play a part in seeking, retrieving, gathering, interpreting, and using information can shed light on some of the lesser-known dimensions of information behaviour. While cognitive biases have received some research attention in areas such as

health and finance, we do not know enough about them in the general context of information behaviour. When there is no vision of the potential impact of cognitive biases on information behaviour, improving information services, access, and use can be challenging. This study aims to contribute to our understanding of the role of cognitive biases in information behaviour. The study contributes to the thin library and information science literature on cognitive biases as one of the few qualitative studies that explore cognitive biases at different stages of information seeking. To this aim, we investigate cognitive biases in a specific user group (graduate students) and a specific context (their project), so the findings can be better contextualised. More specifically the research seeks to answer two questions:

- Which of the cognitive biases occur in the information behaviour of graduate students in their projects?
- How do cognitive biases influence students in each stage of their information behaviour?
- In the rest of the paper, the term 'biases' refers to 'cognitive biases.'

2. LITERATURE REVIEW

Cognitive approach is a significant approach to human information behaviour, and some of its aspects have been covered in works such as Belkin (1990), Ellis (1989), Kuhlthau (1993), and Ingwersen (1992, 2001). For instance, the relationship between different cognitive styles (learning styles) and search behaviour was studied by Ford, Wood, and Walsh (1994). Cognitive styles (along with demographic factors) are considered in the category of individual differences in studies of information behaviour (Ingwersen & Järvelin, 2005).

One aspect of the cognitive viewpoint is the issue of uncertainty in information behaviour. Humans can rely on some heuristics principles and biases to be able to decide under uncertainty (Tversky & Kahneman, 1974). Some studies such as Ingwersen (1996) have dealt with the role of uncertainty in information behaviour, and Kuhlthau (1993) included uncertainty as a principle effect in her information behaviour model. Wilson, Ford, Ellis, Foster, and Spink (2002) showed that the concept of uncertainty can be operationalised in different stages of information seeking processes so that users can express the degree of uncertainty they have.

Despite some studies on certain cognitive aspects in information behaviour, such as the abovementioned studies, cognitive biases have not received much research attention in information science, and more specifically in information

behaviour. However, there have been several studies on biases in some other fields, including accounting and auditing (e.g., Griffith, Hammersley, & Kadous, 2013), business and behavioural finance (e.g., Kariofyllas, Philippas, & Siriopoulos, 2017), information systems (e.g., Arnott, 2006), management (e.g., Pick & Merchant, 2010), health (e.g., Hussain & Oestreicher, 2018), and psychology (e.g., Miloff, Savva, & Carlbring, 2015).

There have been two reviews of bias studies in information systems. One review (Fleischmann, Amirpur, Benlian, & Hess, 2014) covered 84 studies (1992-2012) and concluded that the research in this area was sparse and disconnected and that there was considerable potential for further research. The review showed that most of the studies were in the area of information system usage (70), information system management (27), and software development (11). Overall, 120 biases were examined in the studies, with framing and anchoring being the most commonly examined biases. The review also revealed that researchers used a range of methods for their studies including experiments, surveys, case studies, and interviews. Another review with a focus on software engineering (Mohanani, Salman, Turhan, Rodriguez, & Ralph, 2018) covered 65 papers (1990-2016) that investigated 37 biases with anchoring, confirmation, overconfidence, and availability being the most examined ones.

In health information, a retrospective analysis on search and decision behaviours of 75 clinicians showed that reading the same documents did not result in the same answer by clinicians. The researchers hypothesised that clinicians experience anchoring effect, order effect, exposure effect, and reinforcement effect while searching and these biases might influence their decisions (Lau & Coiera, 2007). A study on undergraduate students showed that applying debiasing strategies to anchoring and order biases influenced their ability to answer health-related questions accurately, as well as the strategies used to conduct searches and retrieve information (Lau & Coiera, 2009). Interviewing with lay individuals revealed that incorrect or imprecise domain knowledge led people to look for health information on irrelevant sites, often seeking out data to confirm their incorrect initial hypotheses due to observed selective perception and confirmation biases (Keselman, Browne, & Kaufman, 2008). Another study (Schweiger, Oeberst, & Cress, 2014) also confirmed the presence of confirmation bias in health-related information searching and that presenting users with tag clouds, including popular tags that challenged their bias, had the potential to counter biased information processing.

Context and tools used for searching information might play a role in biases. Past studies have shown that web searchers are

subject to bias from search engines. Search engines, for instance, strongly favoured a particular, usually positive, perspective irrespective of the truth (White, 2013). Search context can make users more susceptible to confirmation bias (Kayhan, 2015). An experiment showed that when disconfirming evidence was identified using a different word or phrase, the search engine would generate a result set consisting mostly of confirming evidence, which in turn would lead to downloading confirming evidence only, and thus making biased decisions (Kayhan, 2015). However, manipulation of tools, such as presenting comprehensible information in Google's knowledge graph box, could help counter bias information processing (Ludolph, Allam, & Schulz, 2016).

Past studies have also proposed and tested a range of debiasing strategies such as modification of the user interface of search systems (Lau & Coiera, 2009), computer-mediated counter-argument (Huang, Hsu, & Ku, 2012), using tag clouds with tags challenging biases (Schweiger et al., 2014), and adding information to Google's knowledge graph box (Ludolph et al., 2016). Most of the debiasing strategies applied in information searching studies are of cognitive (e.g., "consider the opposite") or a technological type (e.g., provision of external tools to improve the decision environment) (Ludolph et al., 2016), and some positive results have been reported for most of them. However, reviews of past studies on biases (Fleischmann et al., 2014; Mohanani et al., 2018) indicate there is still a great need for research on mitigation techniques.

As the above review indicates, cognitive biases have been examined to a certain extent in searching for health information. However, its roles in information behaviour in other contexts and among different user groups are largely unknown. This study contributes to bridging this gap by focusing on some aspects of information behaviour among graduate students.

3. METHOD

A qualitative approach was used due to the exploratory nature of the study. The lack of research on biases in information behaviour makes it difficult to form hypotheses. Therefore, instead of methods such as experimental methods, a qualitative method was used to make it possible to explore and cover a range of biases and various stages of information behaviour. Moreover, qualitative methods and interviews have been used in the past for the study of existence and the effects of biases in other areas, such as software engineering (Mohanani et al., 2018) and information systems (Fleischmann et al., 2014). Participants were 25 graduate students of Kharazmi University (Iran) who were chosen using

purposive sampling. The students had to be in the dissertation stage, which is when they had finished their taught subjects and were doing research for their dissertation or were writing it up. Recruitment notes were distributed through bulletin boards on campus and the participants were chosen from those who initially expressed their interest in a way to increase diversity in terms of gender, discipline, and research stage. Participants consisted of 14 women and 11 men students from a range of disciplines including library and information science, mathematics, literature, geography, business administration, international relations, law, accounting, economics, management, and geology. Twenty of the participants were PhD students and five were Masters students. They were between 25 and 38 years old with an average age of 32.

Semi-structured interviews were used for data collection. To guide the interviews, Big6™ Skills for Information Literacy (Lowe & Eisenberg, 2005) was adopted as a framework. This model was developed by Eisenberg and Berkowitz (1990). The model includes a process encompassing six stages from task definition to evaluation (listed below). Although this model is mostly dedicated to information problem solving rather than specifically to information seeking processes or behaviour, it was adopted because its stages roughly cover the stages in a research process that students might go through for doing a dissertation/thesis. The steps or stages are not too broad, nor are they too narrow. Also, as Lowe and Eisenberg stated, it is a flexible process that includes necessary elements for solving problems and completing tasks, and it has the potential for the study of human information behaviour. Moreover, it is not always a linear process and can be applied to any information situation, whether an academic or everyday information problem or need (2005, p. 66). The six stages of information problem solving are presented in Fig. 1 (Lowe & Eisenberg, 2005, p.65).

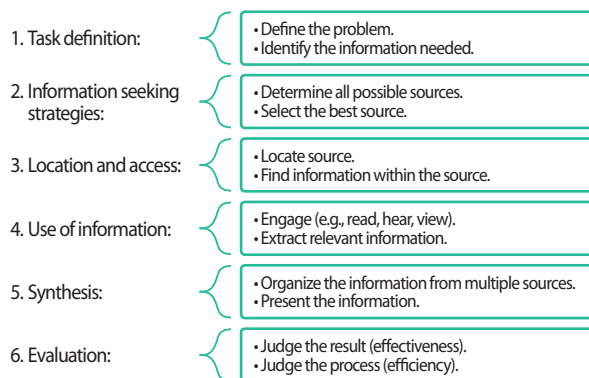


Fig. 1. Six stages of the Big6 model. Reproduced from Lowe and Eisenberg, Information Today 2015.

An interview protocol was developed based on the stages of the Big6 model. A few pilot interviews were conducted to test the interview protocol and as a result, some modifications were made in the interview protocol including its structure. Before interviews, participants were asked to think of real actions they had taken for their research and answer the questions based on their lived experiences. Interviews consisted of a series of questions (overall between 16 to 21 questions) grouped based on the six stages in the Big6 model, plus a few overall questions at the end about their decision making during their research. Questions included items related to interaction with information in the context of students' projects, successful and unsuccessful search and information access experiences, students' actions, and decision making throughout the six stages of the model. For instance, for the locating and accessing stage, we asked them how they located and got access to the information resources they had identified and how they made decisions in that regard. Or, for the use of information, we asked them what they did with the information they found, what type of information they considered to use, and if they ignored or avoided any information and why. Interviews were held on campus at a time and location convenient to the participants. They were recorded and transcribed for analysis. Participants received an information sheet and signed a consent form. Interviews took on average about 50 minutes. The data collection was continued until the researchers felt the point of saturation had been reached. Fictional first names have been used in the interview quotations in the results section so the gender of the participants is clear to readers.

Grounded theory coding method was used for data analysis with two main stages of coding the interviews, including substantive coding and theoretical coding, based on Glaser's approach (Glaser & Strauss, 1967). After transcribing the interviews, each interview was broken down into smaller units of analysis and coding was done in order to reach the main concepts. The data analysis process is nonlinear, and some other sub steps like data comparison on as many dimensions as possible were involved. Based on the aims of the study, the main focus of the researchers during the coding was on identifying the biases and their roles in the students' information behaviour. Identification of biases was based on definitions of biases and matching the examples of the behaviour of students with the definitions of biases.¹ Also, bias identification was based on the whole of the interviews and on all the descriptions which were presented by participants and all conceptualisations of the

¹ To determine a cognitive bias, in many cognitive biases' sources, like Dobelli (2013), an example of a situation or behaviour has been presented.

data, although in this article only some sentences from different interviews have been mentioned as evidence. As an example of the process of the data analysis, an excerpt from one of the interviews is presented below:

... once I've found the information I'm looking for, I start reading and taking notes. For using the information I check if the information is useful for my work. If there is some piece of information in the work that is important for my thesis and somehow supports the idea, model or hypothesis in my thesis I would highlight it and it catches my interest and I make sure I make note of it...

For the excerpt above, initially codes such as 'a decision about the use of information,' 'a tendency to favor supporting information,' and 'the effect of the user's mentality on selecting information' were assigned. During the next stage, it was determined that the incident was related to the fourth stage of Big6 (use of information) and that was assigned as a code. Also, the codes were changed to 'inclination towards information that is aligned with own views.' At the final stage, the comparison of codes with the definition of the biases resulted in coding the data as the incident of confirmation bias during stage four of information seeking.

To ensure the credibility of the findings, a few actions were taken including doing data analysis twice with a two-month interval to compare the results, and they were largely in accord with each other. The comparison showed the analysis was consistent. Member checking and external audit were also used. For member checking, the analysis of the interview transcripts was sent to the participants about one or two weeks after the interviews and they were asked to check if the interpretations of the interviews were sound. For the external audit, a second researcher was asked to go through the data and the outcome of the analysis to check if the findings and interpretations were supported by the evidence in the data.

4. RESULTS

The interviews revealed the presence of 28 biases in the information behaviour of the participants. Table 1 shows the list of these biases and their effect on information behaviour. For definitions of biases please see the Appendix. The number of participants (3rd column) shows how many participants experienced that bias. Frequency shows the total frequency of the bias in all of the interviews, as each bias might have been present in more than one stage of the information behaviour of a participant. The stages column shows the stages of information

behaviour (based on Big6 stages) in which biases occurred. The final column is the effects that an information bias could have on the information behaviour. These biases can be categorised into four groups according to the problems that biases help us address, based on the categorisation proposed by Benson (2016). They include:

- Information overload (biases No. 1-12);
- Lack of meaning (biases No. 13-19);
- The need to act fast (biases No. 20-27); and
- How to know what needs to be remembered for later (biases No. 28, 5, & 14).

To better understand the data presented in Table 1 and the effects of biases on information behaviour, each stage based on Lowe and Eisenberg's (2005) model is discussed below. It should be noted that the biases discussed under each stage were not restricted only to that stage. The relationships between biases and stages of information seeking are merely based on the evidence of their presence in each stage that were identified during the analysis (see the coding example above).

4.1. Stage 1: Task Definition

When participants are choosing the subject of their research and defining the topic of their information needs, due to ambiguity aversion they might avoid choosing topics that do not have a rich past literature or avoid research locations with which they are not familiar. For instance, Mary said "I needed to choose a location for my research and I decided to choose this city because I was familiar with that city and I lived there for a few years so I knew its various quarters and areas. If it was another city about which I knew little, what would I do?"

Besides avoiding challenging situations, participants may favour things that bring peace of mind (selective perception). Philip said "I wanted my research topic to be something about which some information already existed. My supervisor suggested a few topics and I checked them and picked the one that looked more familiar to me based on what I already knew and could use. This was because I already knew a few things about the topic and had an overall knowledge of it."

If participants are overconfident and not fully aware of the limits of their knowledge and skills, they may end up choosing a topic that they do not understand properly. Participants referred to such situations in several interviews; for instance, "the topic was very interesting and professional and I spent a lot of time on it. Although I knew it was a difficult one I started it and after a few months, I realised it was beyond my knowledge and skills" (Mary Anne).

Recency bias sounds like a natural habit in our information

Table 1. Cognitive biases, their frequency, and effects on information behaviour

No.	Cognitive bias	No. of participants	Frequency	Stages	Effect
1	Availability bias	14	16	2, 3	Affects the preference and selection of information-seeking strategies and choice of information centres.
2	Attentional bias	13	13	2	Influences the specific aspects of information sources to which users pay attention to select the sources.
3	Illusion truth effect	6	8	4, 5	Increases the tendency to obtain, present, and share information that is perceived as straightforward and unambiguous.
4	Picture superiority effect	5	8	4, 5	Increases the tendency to obtain illustrated/graphical information, and use such information for presenting and sharing.
5	Negativity bias	13	13	4	Makes users inclined towards obtaining negative information.
6	Anchoring bias	12	15	1, 4	Persuades users to anchor in the first chosen subject or a special part of the information.
7	Conservatism bias	9	11	1, 2, 4	Makes users avoid new and different topics, or inclined to ignore recent and new information, or emphasise their previous knowledge and ignore recent information.
8	Framing effect	11	12	4	Increases the tendency to present the information in a preferred and customary format.
9	Confirmation bias	15	19	4, 5	Increases the tendency to use the information that is supportive of one's beliefs and perspectives, or to synthesise information in a way to confirm one's beliefs.
10	Choice-supportive bias	14	14	6	Results in confidence about and positive evaluation of the soundness of information-seeking processes.
11	Selective perception	9	14	1, 3, 4	Increases the tendency to choose topics, resources or information centres that seem more aligned with one's expectations.
12	Observer-expectancy effect	7	7	5	Increases the tendency to synthesise and present biased information to confirm or support one's prior beliefs.
13	Recency bias	6	9	1, 4, 5	Results in choosing subjects or information to which users have been exposed and interacted with more recently; and in presenting the information that has been obtained more recently.
14	Stereotypical bias	12	12	2	Make users prefer and choose some sources according to stereotypes.
15	Bandwagon effect	12	16	1, 2, 3, 5	Increases the likelihood of choosing subjects, sources or information centres that are popular and favoured by other people; and increases the tendency to present information that is attractive to more people.
16	Curse of knowledge	10	11	1, 2	Restricts users to their prior knowledge and results in an inability to explore new & different aspects of an issue, or restricts them to a limited number of keywords derived from their prior knowledge.
17	Illusion of transparency	10	10	5	Results in a misconception of transparency of the presented information.
18	Pro-innovation bias	6	8	2, 5	Urges users to use a variety of new technologies to obtain and present information.
19	Planning fallacy	14	14	6	Results in the search process not progressing as planned or scheduled.
20	Self-serving bias	12	12	6	Results in the tendency to blame external factors for the failures of information seeking.
21	False consensus effect	13	13	6	Leads to an overestimation of the extent to which others agree with one's opinion.
22	Overconfidence effect	11	14	1, 2, 6	Results in choosing a subject that is beyond one's abilities and skills, in an overestimation of the soundness of chosen information-seeking strategies, and in an unrealistic positive evaluation of one's information seeking.
23	Status quo bias	12	16	2, 3, 5	Makes users reluctant to change their usual information-seeking strategies; influences the preference for information sources and centres users are already familiar with, and increases the tendency to present the information in the format one is already accustomed to.
24	Reactance	11	11	2	Increases the urge to obtain the information on which there is access restriction.
25	Belief bias	4	5	2, 5	Results in making decisions when searching for or synthesising information based on what sounds more trustworthy and reliable to one.
26	Information bias	11	11	2	Makes users seek and collect too much information.
27	Ambiguity aversion bias	10	14	1, 2	Makes users choose subjects that seem less ambiguous or avoid unfamiliar sources and centres.
28	Serial position effect	10	14	5	Increases the attention to the order that information is presented (first results when obtaining, and more recently obtained information when presenting).

behaviour, as we tend to keep in mind recent events. As a result, the books we have recently read or a seminar we have recently attended are information sources that come to our mind when we are thinking about choosing topics. Information resources might be preferred for various reasons including being popular in a community. Bandwagon effect indicates that participants might favour information that is widely favoured by other people. For instance, Philip stated that “of course the topic I am working on is of interest to many people, i.e. many researchers in our field are interested in this topic. In the last few years, many people have done some work on the topic and in a way it is a hot topic in our field and that’s why I chose it.” The interviews showed that the curse of knowledge (Birch, Brosseau-Liard, Haddock, & Ghrear, 2017) also played a role at this stage of information behaviour. Based on their prior knowledge and prior perceptions participants tended to ignore some aspects of the problem and limited themselves to what they already knew.

The other bias in this stage was conservation bias, which means people tend to keep their current perceptions and they do not react to new information. When choosing research topics, participants might not choose what seems to be radical and take action based on their preconceptions. Anchoring bias was also identified at this stage, which indicates participants could not leave what they considered in the beginning and move on. In other words, they anchor on a specific topic that has been formed in their mind for some reason although it may not have the potential for a suitable research topic. Kasey stated that “one thing that I could not get rid of was that I frequently went back to the same topic. I had a specific thing on my mind, which my supervisor suggested and I thought this is the best and I got stuck on it but I could not reach any conclusion on that matter. I wasted a lot of time on that.”

4.2. Stage 2: Information Seeking Strategies

Participants were not inclined to use information sources that were new to them and with which they had no prior experience due to ambiguity aversion. For example, participants insist on their keywords and avoid new ones. Participants might also be overconfident in their methods and strategies for seeking information. Due to the bandwagon effect, participants were in favour of information sources that are prevalent and common or were recommended by their peers. Jessica stated, “it has happened for me and my friends that we use a set of keywords for searching and we frequently see the same results. The keywords are right and I found them based on the information I have and based on my knowledge, however, I learned that there are other things and I need to expand my mind.”

Due to attention bias, participants pay more attention to

certain aspects of information (e.g., author’s reputation or affiliation) that affected their choices. When participants are deciding about their search strategies, they do so based on their past experiences in order to choose the best strategies. They consider past successful strategies as those that are more likely to lead to success (availability bias) (Pompian, 2006). An example is: “the most important thing and the first thing I do is to ask one or two of my friends that I know are knowledgeable and experienced to tell me where I can find such information and I frequently found this helpful” (David).

Based on status quo bias it is expected that participants will stick to their search strategies and are reluctant to change their habits. Those who experience this bias use the same websites, information sources, resources, and libraries that they have used in the past: “I always go to a website that I have been using since 2011 when I was a student because I know I can find what I need there” (Ruth).

The participants’ strong preference for using online resources based on the belief that the era of print resources is over indicates belief bias. Another example of belief bias is the prevalent belief among participants that English resources are superior to all of the resources in the Persian language. Besides our dependence on our beliefs, stereotyping (Hinton, 2000) also influences participants’ behaviour. The superiority of English resources can also be considered as a stereotype among Iranian students. For instance, “I don’t read Persian articles and it happens rarely that I consult Persian resources. I only use Persian resources occasionally when I need to say what has been done nationally. However, mostly my preference is English for various reasons, including that foreign authors are more knowledgeable” (Kim).

Information bias refers to the tendency to collect too much information. In today’s world with an overabundance of information, some participants drown in information by continuing extensive searching for information. Ruth said:

“I try to look at all resources as much as possible. My personal desire is to go all the way and find every piece of information that exists on the topic. I constantly think there might be more and I should carry on and I look for more and more information. Maybe the next article gives me a different perspective and that’s why I like looking for information. It is self-evident that more information is always better.”

The other identified bias is pro-innovation bias (Rogers, 2010), which leads participants to enthusiastically embrace new technologies and techniques for searching and finding resources. Some of the participants were pro-innovation and tended to use new technologies to obtain information. They consider new technologies to have many advantages and few deficiencies.

Reactance plays a part when participants are prevented from accessing certain information resources. As past studies (Jamali & Shahbaztabar, 2017) have shown, in such situations where participants are discouraged or banned from accessing some information, they become keener on accessing such resources and will try harder (reaction to censorship). Mary very firmly stated that “when a person or a group decides that you should not have or read certain information, one becomes more enthusiastic about obtaining that information. No one has the right to decide what I should or should not know.”

Another bias at this stage is the serial position effect, which results in paying more attention to resources that are higher in the list of retrieved resources. Barbara said that “when searching, I think I pay more attention to the first results, not only in Google but in any database, the top results receive more attention and you download those items” (Barbara).

4.3. Stage 3: Location and Access

Similar to the two previous stages, bandwagon effect impacts on participants at this stage, for instance in preference for using online information resources that are recommended by and are popular amongst their peers. Jason stated that “I first search databases that are well known in our field, those that have been around for many years and everybody in our field uses those. One of our lecturers gave us a list of prestigious journals that was pretty exhaustive and we search in those journals and check their websites.”

At this stage, participants use the location for finding and accessing information that they consider more likely to provide access to the information based on their past experiences. This availability bias was evident in several interviews. For instance, “I need to check the websites and look for resources, for instance, I need to use Google Scholar or ScienceDirect. I first use Google Scholar because I’ve used it before and I know that I can find the articles I need so I use it again to find new articles” (Daniel). They also tend to choose information centres or sources with which they are familiar and accustomed to and they avoid any change due to status quo bias. Kasey said, “I think once I have found the resources I wanted I’d go to the special library of our discipline or the special website to get them. I have used them before and I know how to get the results fast. This is better than trying to find a new way.”

Participants choose choices that are aligned with their expectations (selective perception). Peter said “I never go to our university’s library. It is very dull, its staff are impatient. Everything is old and worn out. I like to go to large libraries where you feel comfortable, unlike our university library to which I wouldn’t go even if I know it has the information I need.”

4.4. Stage 4: Use of Information

When encountering and using information, participants may favour some information pieces over others. Several biases may play a role in such situations. Participants might favour information that is aligned with their perceptions and beliefs (confirmation bias). Ethan mentioned that “in the content, specific things might catch my attention. For instance, I am reading something and I see that something, documented or not, is what I was looking for and it confirms my idea and it is something that I agree with, therefore I make sure I use it in my work.” Participants may also prefer information that is simple and unambiguous and easy to understand (illusion of truth effect) (Hirshleifer, 2001). Victor stated that “I disregarded the resources that had complex concepts. But as I said I knew that they were relevant due to the references to them. Overall if the resource has concepts that I find difficult to understand and I can’t figure it out after a few times of reading or if they require a lot of time, I will disregard them.” Visual or graphical information is also often preferred (picture superiority effect).

It is evident that participants may receive information that is negative in relation to their research topic, things such as sad statistics about something and dissatisfaction in people. Participants stated that they paid more attention to such negative information compared to positive information and prefer to extract such information. Julia said: “When I think, I can see that the first thing in a text that gets my attention is the negative news. The negatives are seen more than the positives, negative statistics and news are more unfortunate and sorrowful and should receive more attention.”

Besides the above biases, people put emphasis on specific features of resources and choose the resources based on those that could be due to selective perception or anchoring. An example of selective perception is when participants prefer resources in the Persian language because they want to avoid the challenge of reading English resources, as they do not feel as confident in the English language. For anchoring, Nathan stated that “for this work, I first read a few articles and I understood the overall topic based on those articles. Because I didn’t know much beforehand I moved forward based on the theories I learned from those articles and I put the basis of my work on those theories and I used those as keywords for searching.” Moreover, participants anchor on the quantity of the work required by their university for a dissertation or thesis (e.g., number of pages) and that quantity drives their work. Also, based on conservation bias, participants disregard new information and react better to information that is aligned with their prior perceptions. Joanne said, “I have seen multiple times in the research seminar class that when discussing a topic, some other participants put too

much emphasis on one particular resource or work and don't want to listen to someone else or change their opinion, no matter how many arguments you put forward."

4.5. Stage 5: Synthesis

When presenting information, participants were influenced by biases such as confirmation bias, illusion of truth effect, picture superiority effect, and bandwagon effect. As a result of confirmation bias, participants tend to present information that confirms their perceptions. Ethan said, "in my opinion, content that I have extracted in the previous stages that are things that are aligned with my opinion, and sound interesting to me and confirm my ideas, I present those in my chapters." The illusion of truth effect indicates that participants prefer to synthesise, organize, and present the information in a simple way. Participants also prefer to use pictures (picture superiority): "when I am reading, pictures help me more than text so when I am writing I make more use of pictures. If there is no picture, readers have to read and visualise in their mind, which might not be a simple thing to do and might result in mistakes. So I prefer to use pictures when I put myself in readers' shoes" (Diana). In bandwagon effect (VandenBos, 2007), participants tend to present information that is interesting to more people: "we need to talk about things in our work that are hot topics of the day and are the buzzwords" (Noah).

Framing effect influenced both the format and structure of the content that participants present. Participants might prefer to present information in certain formats and appearances, and/or highlight or play down parts of the content. For example, Daniel said, "sometimes one could see that some information resources present information in certain ways to achieve the desired outcome. They tell you the story the way they want to. I think this is natural and I do it too. I might emphasise part of the content and play down some other parts that might undermine the conclusion of my work. You could present someone's life story in different ways that result in forming a different opinion in the audience."

Recency bias and serial position effect explain situations where participants rely on recently obtained information in presenting information. At this stage, these two biases may overlap a bit. Belief bias means participants will easily accept an argument if they find it easy to believe and they may not look at the evidence. For instance, if they believe that the information they present is useful they put a lot of effort into presenting such information.

While participants believed that in many cases the information they presented was clear, their audience thought otherwise and asked for clarifications. This implies the illusion

of transparency (Brown & Stopa, 2007), when students think they are very clear about what they present but the reactions and feedback they receive indicate that they have not been successful in conveying their message. One student notes, "It has happened several times when I gave a seminar or in my confirmation presentation, that other students ask questions that I think I have already clearly answered in my text or presentation. But it seems it hasn't been sufficient, therefore, I am always afraid that in my viva I face criticism that this doesn't mean what you mean to say..." (Kasey).

4.6. Stage 6: Evaluation

If the person has an unrealistically positive evaluation of the information-seeking process and sees all of the information relevant to his or her information needs, then overconfidence effect might be involved. Diana stated that "I wasn't after new search strategies and I didn't want to learn about those, I was pretty sure that the right search strategies were what I already knew and I already used. I wasn't concerned whether my approach or strategy was correct or incorrect." Sometimes because of choice-supportive bias (Lind, Visentini, Mäntylä, & Del Missier, 2017), participants approve the validity of the process and its outcome, so they support their own decisions and positively evaluate their actions and behaviour. An example is: "I think overall the process that I went through was good. I mean I think I couldn't do more. My method is fine, and things I've done, places I searched, and the ways I used to access the information are all correct and I reach desired outcomes and this has been the case so far" (Leah).

According to self-serving bias, people often attribute problems with the information-seeking process (and their failure) to external factors (e.g., the education system, time). On the other hand, they attribute their successes to their merits and internal factors. This theme was quite common in the participants' comments in interviews. Based on false consensus effect, participants think that all other students are in agreement with them about the process of looking for information and they see their beliefs as self-evident. For instance, participants frequently commented in the interviews that everyone agrees about the superiority of online resources to print ones, or that 'Googling' is the best and first choice for looking for information. During the information-seeking process and at the stage of evaluation, participants realise that they have not progressed according to their plan and that the process has taken longer than they thought it would. This is due to the planning fallacy (Ehrlinger, Readinger, & Kim, 2016), as many of the issues and challenges are not considered in such planning. Chloe experienced this planning issue and stated that "unfortunately I am in a term

when I should have finished half of my thesis based on a standard four-year doctoral period. When I started I thought I would finish in four years and had a lot of plans and always tried to stick to my plans. But now I am behind in most of my plans. Of course, I must say there are factors that are not under my control such as our supervisors and the university's situation." When participants were asked if they were satisfied with their information seeking and about their strengths and weaknesses and so on, we see the effect of biases, as most of them would answer that they had done the right things and that they attribute success to their strengths and attribute problems to external factors. They also think that their peers behave in the same way and that they are all in agreement about this.

5. DISCUSSION AND CONCLUSIONS

This study showed the prevalence and diversity of biases in some aspects of information behaviour of graduate students. The findings show how biases as psychological and cognitive factors might influence our information behaviour. Biases can play roles in all aspects of information behaviour from defining the details of information needs, deciding about and selecting information sources, extracting and using the information, organising and sharing obtained information, and in evaluating the process of looking for information.

Similar to findings of this research, several past studies including Lau and Coiera (2007), Keselman et al. (2008), and Blakesley (2016) have highlighted the influence of biases on information behaviour. Although some of the past studies focused on specific biases and some investigated the issue only in the context of searching in the digital environment, we can see that biases such as anchoring, confirmation, and status quo biases that are prevalent among participants in this study were also identified in previous studies. Some of the biases such as availability bias (stages 2 & 3), conservation bias (stages 1, 2, & 4), selective perception (stages 1, 3, & 4), recency bias (stages 1, 4, & 5), bandwagon effect (stages 1, 2, 3, & 5), overconfidence effect (stages 1, 2, & 6), and status quo bias (stages 2, 3, & 5) were present in several stages of information behaviour. Due to the continuity and non-linear nature of the processes of looking for information, it seems normal for biases to play a role in different stages of information behaviour. Some of the biases (e.g., confirmation bias) that affect participants when finding information also affect them when presenting information.

The findings of this study show that the roles that biases play in information behaviour have some consequences. Biases might make users favour some choices and be biased against

other choices. In the stage of defining information needs, participants might pay more attention to some issues and ignore or oversee other issues because they are under the influence of biases such as selective perception, bandwagon effect, and ambiguity aversion. Biases also might make participants favour or avoid some choices when selecting and choosing information resources (physical or digital) and referring to information centres (physical or virtual). Biases might also result in search failure: for instance, by sticking to the keywords they already know which may not be the best keywords. They might focus on one dimension of the topic they are familiar with and ignore the other dimensions. Another example of failure is when availability bias encourages participants to favour information sources and resources that they deem more suitable to meet their needs based on their past experiences.

Attentional bias (Baron, 2008), belief bias, framing effect, and stereotyping are among biases that may lead to failure in finding relevant information. This is because participants focus their attention on certain aspects of information and act on that basis. For instance, a user might use authors' authority or impact factors as proxies for quality rather than considering the quality of the content, or a Persian-speaking student might favour English sources with the assumption that they are superior in terms of the quality of content.

The findings showed that consequences of some of the biases might become evident in extracting, using, synthesizing, and sharing information. Likewise, Lau and Coiera (2007) in their study showed that different users presented different answers to the same questions in similar experimental scenarios while they all received the same information. The conclusion was that they experienced anchoring, exposure, and order biases while searching for information. Similarly, in the current study the participants seemed to have experienced anchoring, negativity (Ito, Larsen, Smith, & Cacioppo, 1998), confirmation, conservation, the illusion of truth, picture superiority (Ma, 2016), and framing biases when extracting and presenting information. The study also found that some of the biases can have negative consequences in the stages of synthesising (organising and presenting) information. Status quo bias and framing effect might make users present the information in a way they are used to and seems desirable to themselves, which may not be the best or correct way of presenting the information. Or, observer-expectancy effect can result in bias in the presentation of information. The study by Lomangino (2016) showed the role of confirmation bias in reporting research findings.

Finally, the consequence of biases in the evaluation of effectiveness and efficiency of the information-seeking process

is that users might blame external factors in their failure due to overconfidence effect and other similar biases. As a result, users will not seek to improve their skills because they do not blame their knowledge and skills for the lack of success. This might be also related to the attribution style of users, as those who attribute their success or failure to external factors (compared to those who attribute them to internal factors) experience a lower level of satisfaction and do not aspire to improve their capabilities (Behzadi & Sanatjoo, 2019).

We must note that the presence of biases does not mean that users are always under the influence of such biases. Their role and consequence of their involvement in our behaviour vary based on contextual factors. The results of their influence are not always the same either. Biases and the use of mental shortcuts do not always yield negative outcomes and sometimes they serve as simple ways to achieve the desired outcome. We also need to bear in mind that biases are not the only factor in information behaviour and many factors play a role. Biases might also be related to some of the known concepts of information behaviour such as information avoidance (which might be related to confirmation bias, selective perception, and conservation bias); this has been discussed by Behimehr and Jamali (2020) and requires further investigation.

While this study did not involve any debiasing techniques and concepts, the findings have some implications for debiasing. Designing and implementing debiasing techniques, which can entail changes in user interfaces (Lau & Coiera, 2009) requires knowing which cognitive bias users might experience in each stage of their information seeking. Therefore, studies such as this one are needed for system designers if debiasing techniques are to be implemented in information systems.

This study had some limitations as it focused on a small group of students in the context of a specific task (theses/dissertations). The other issue is that identification of biases which are latent phenomena is not straightforward and one needs to be cautious in attributing biases to certain behaviour of individuals, and this might require benchmarking against objective baselines (Kahneman & Tversky, 1979). Although we have used the evidence in the interviews as an indication of the possible presence of biases, it is far from perfect. The evidence in this study is suggestive rather than in any sense conclusive. Moreover, the study had an exploratory nature and the aim was to explore if and how biases might play a role in information behaviour. Fleischmann et al. (2014, p. 10) argued that although biases are latent phenomena, qualitative and argumentative methods such as interviews are not inappropriate for studying them. We hope that the results of this study can stimulate further much needed studies on the subject.

Future studies should examine biases in different contexts (e.g., the digital environment) and for different user groups. Some of the participants in this study were library and information science students, which could be counted as a limitation of the study. Their research method coursework is similar in nature to that of other disciplines in social sciences. However, they are more exposed to information literacy education. The impact of information literacy education on cognitive biases could be the subject of a future study. The role of culture could also be considered in biases in users from different countries. Possible relationships between cognitive biases and cognitive styles should also be studied. Also, there is a need for more studies on mitigating methods and strategies for the negative impact of biases. In the context of information behaviour and information literacy, increasing awareness of cognitive biases, teaching critical thinking skills, and designing information systems so they expose users to counter bias information might help.

REFERENCES

- Arnott, D. (2006). Cognitive biases and decision support systems development: A design science approach. *Information Systems Journal*, 16(1), 55-78.
- Baron, J. (2008). *Thinking and deciding (4th ed)*. Cambridge: Cambridge University Press.
- Bazerman, M. H., & Moore, D. A. (2012). *Judgment in managerial decision making*. Hoboken: Wiley.
- Behimehr, S., & Jamali, H. R. (2020). Relations between cognitive biases and some concepts of information behavior. *Data and Information Management*, 4(2), 1-10.
- Behzadi, H., & Sanatjoo, A. (2019). Attributional style of emotions and its relationship with users' search behavior. *Journal of Information Science*, 45(1), 105-116.
- Belkin, N. J. (1990). The cognitive viewpoint in information science. *Journal of Information Science*, 16(1), 11-15.
- Benson, B. (2016). Cognitive bias cheat sheet. *Better Humans*. Retrieved November 22, 2017 from <https://betterhumans.coach.me/cognitive-bias-cheat-sheet-55a472476b18>.
- Birch, S., Brosseau-Liard, P. E., Haddock, T., & Ghrear, S. E. (2017). A 'curse of knowledge' in the absence of knowledge? People misattribute fluency when judging how common knowledge is among their peers. *Cognition*, 166, 447-458.
- Blakesley, E. (2016). Cognitive bias and the discovery layer. *The Journal of Academic Librarianship*, 42(3), 191.
- Brown, M. A., & Stopa, L. (2007). The spotlight effect and the illusion of transparency in social anxiety. *Journal of Anxiety Disorders*, 21(6), 804-819.

- Dobelli, R. (2013). *The art of thinking clearly*. New York: Farrar, Straus and Giroux (FSG).
- Ehrlinger, J., Readinger, W. O., & Kim, B. (2016). Decision-making and cognitive biases. Retrieved December 17, 2017 from https://www.researchgate.net/publication/301662722_Decision_Making_and_Cognitive_Biases.
- Eisenberg, M. B., & Berkowitz, R. E. (1990). *Information problem-solving: The Big Six Skills approach to library and information skills instruction*. Norwood: Ablex.
- Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of Documentation*, 45(3), 171-212.
- Fidel, R. (2012). *Human information interaction: An ecological approach to information behavior*. London: The MIT Press.
- Fleischmann, M., Amirpur, M., Benlian, A., & Hess, T. (2014, June 9-11). *Cognitive biases in information systems research: A scientometric analysis*. Paper presented at the European Conference on Information Systems (ECIS), Tel Aviv, Israel. Retrieved February 20, 2017 from: <http://aisel.aisnet.org/ecis2014/proceedings/track02/5>.
- Ford, N., Wood, F., & Walsh, C. (1994). Cognitive styles and searching. *Online and CD-ROM Review*, 18(2), 79-86.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.
- Griffith, E. E., Hammersley, S., & Kadous, K. (2013). Auditing complex estimates: Understanding the process used and problems encountered. *SSRN Electronic Journal*, Retrieved October 11, 2017 from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1857175.
- Heuer, R. J. (2007). *Psychology of intelligence analysis*. New York: Central Intelligence Agency.
- Hinton, P. R. (2000). *Stereotypes, cognition and culture*. Hove, East Sussex: Psychology Press.
- Hirshleifer, D. (2001). Investor psychology and asset pricing. *The Journal of Finance*, 56(4), 1533-1597.
- Huang, H., Hsu, J. S., & Ku, C. (2012). Understanding the role of computer-mediated counter-argument in countering confirmation bias. *Decision Support Systems*, 53(3), 438-447.
- Hussain, A., & Oestreicher, J. (2018). Clinical decision-making: Heuristics and cognitive biases for the ophthalmologist. *Survey of Ophthalmology*, 63(1), 119-124.
- Ingwersen, P. (1992). *Information retrieval interaction*. London: Taylor Graham.
- Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation*, 52(1), 3-50.
- Ingwersen, P. (2001). Cognitive information retrieval. *Annual Review of Information Science and Technology*, 34, 3-52.
- Ingwersen, P., & Järvelin, J. (2005). Information retrieval in context: IRiX. *ACM SIGIR Forum*, 39(2), 31-39.
- Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, 75(4), 887-900.
- Jamali, H. R., & Shahbazzabar, P. (2017). The effects of internet filtering on users' information-seeking behaviour and emotions. *Aslib Journal of Information Management*, 69(4), 408-425.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 47(2), 263-292.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kariofyllas, S., Philippas, D., & Siriopoulos, C. (2017). Cognitive biases in investors' behaviour under stress: Evidence from the London Stock Exchange. *International Review of Financial Analysis*, 54(2), 54-62.
- Kayhan, V. (2015, December 13-16). *Confirmation bias: Roles of search engines and search contexts*. Paper presented at the Thirty Sixth International Conferences on Information System (CICIS), Fort Worth, TX. Retrieved August 25, 2017 from: <https://pdfs.semanticscholar.org/2b04/1eb44fd8031596ed8e73124801c5bcf550b1.pdf>.
- Keselman, A., Browne, A. C., & Kaufman, D. R. (2008). Consumer health information seeking as hypothesis testing. *Journal of the American Medical Informatics Association: JAMIA*, 15(4), 484-495.
- Kruglanski, A. W., & Ajzen, I. (1983). Bias and error in human judgment. *European Journal of Social Psychology*, 13(1), 1-44.
- Kuhlthau, C. C. (1993). A principle of uncertainty for information seeking. *Journal of Documentation*, 49(4), 339-335.
- Lau, A. Y., & Coiera, E. W. (2007). Do people experience cognitive biases while searching for information? *Journal of the American Medical Informatics Association: JAMIA*, 14(5), 599-608.
- Lau, A. Y. S., & Coiera, E. W. (2009). Can cognitive biases during consumer health information searches be reduced to improve decision making? *Journal of the American Medical Informatics Association*, 16(1), 54-65.
- Lind, M., Visentini, M., Mäntylä, T., & Del Missier, F. (2017). Choice-supportive misremembering: A new taxonomy and review. *Frontiers in Psychology*, 8, Article 2062.
- Lomangino K. M. (2016). Countering cognitive bias: Tips for recognizing the impact of potential bias on research.

- Journal of the Academy of Nutrition and Dietetics*, 116(2), 204-207.
- Lowe, C. A., & Eisenberg, M. B. (2005). Big6 skills for information literacy. In K. E. Fisher, S. Erdelez, & L. McKechnie (Eds.), *Theories of information behavior* (pp. 63-69). Medford: Information Today.
- Ludolph, R., Allam, A., & Schulz, P. J. (2016). Manipulating Google's knowledge graph box to counter biased information processing during an online search on vaccination: Application of a technological debiasing strategy. *Journal of Medical Internet Research*, 18(6), e137.
- Ma, Y. (2016). Can more pictures bring more readership?: An examination of the "Picture superiority effect" in the news consumption process. *Social and Behavioral Sciences*, 236, 34-38.
- Miloff, A., Savva, A., & Carlbring, P. (2015). Cognitive bias measurement and social anxiety disorder: Correlating self-report data and attentional bias. *Internet Interventions*, 2(3), 227-234.
- Mohanani, R., Salman, I., Turhan, B., Rodriguez, P., & Ralph, P. (2018). *Cognitive biases in software engineering: A systematic mapping study*. *IEEE Transactions on Software Engineering*. doi: 10.1109/TSE.2018.2877759.
- Pick, K., & Merchant, A. K. (2010). *Blind spots, biases and other pathologies in the boardroom*. New York: Business Expert Press.
- Pompain, M. M. (2006). *Behavioral finance and wealth management: How to build optimal portfolios that account for investor biases*. Hoboken: Wiley.
- Rogers, E. M. (2010). *Diffusion of innovations*. New York: Free Press.
- Schmutte, J., & Duncan, J. R. (2014). Making independence decisions under the code of professional conduct: Understanding and controlling common cognitive biases. *The CPA Journal*, 68-71.
- Schweiger, S., Oeberst, A., & Cress, U. (2014). Confirmation bias in web-based search: A randomized online study on the effects of expert information and social tags on information search and evaluation. *Journal of Medical Internet Research*, 16(3), e94.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131.
- VandenBos, G. R. (2007). *APA dictionary of psychology*. American Psychological Association. Retrieved December 11, 2017 from <https://dictionary.apa.org/anchoring-bias>.
- White, R. (2013, July 28-August 1). Beliefs and biases in web search. In G. J. F. Jones & P. Sheridan (Eds.), *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 3-12). ACM.
- Wilke, A., & Mata, R. (2012). Cognitive bias. In V. Ramachandran (Ed.), *Encyclopedia of human behavior* (pp. 531-535). London: Academic Press.
- Wilson, T. D., Ford, N., Ellis, D., Foster, A., & Spink, A. (2002). Information seeking and mediated searching: Part 2. Uncertainty and its correlates. *Journal of the American Society for Information Science and Technology*, 53(9), 704-715.

APPENDIX. Definition of cognitive biases

No.	Cognitive biases	Brief definition
1	Availability bias	A mental shortcut that allows people to estimate the probability of an outcome based on how prevalent or familiar that outcome appears in their lives (Pompian, 2006, p. 94).
2	Attentional bias	A failure to look for evidence against an initial possibility, or a failure to consider alternative possibilities (Baron, 2008, p. 188).
3	Illusion of truth effect	People are more inclined to accept the truth of a statement that is easy to process (Hirshleifer, 2001, p. 1542).
4	Picture superiority effect	Viewing pictures are more easily and frequently recalled than are concepts that are learned by viewing their written word form counterparts (Ma, 2016, p. 1).
5	Negativity bias	Negative information tends to influence evaluations more strongly than comparably extreme positive information (Ito, Larsen, Smith, & Cacioppo, 1998, p. 887).
6	Anchoring bias	A tendency for one's first impressions or reactions to become the reference point (or anchor) for subsequent thoughts and actions (Schmutte & Duncan, 2014, p. 69).
7	Conservatism bias	People cling to their prior views at the expense of acknowledging new information (Pompian, 2006, p. 119).
8	Framing effect	The tendency to respond to various situations differently based on the context in which a choice is presented (framed) (Pompian, 2006, p. 237).
9	Confirmation bias	The tendency to selectively search for or interpret information in a way that confirms one's preconceptions or hypotheses (Wilke & Mata, 2012, p. 532).
10	Choice-supportive bias	The tendency to retroactively ascribe positive attributes to an option one has selected and/or to demote the forgone options (Lind, Visentini, Mantyla, & Fabio, 2017).
11	Selective perception	The process in which people choose to attend to one or a few stimuli from the myriad array of stimuli presented to the senses at any one time (VandenBos, 2007).
12	Observer-expectancy effect	Any expectations, beliefs, or personal preferences of a researcher that unintentionally influence his or her recordings during an observational study (VandenBos, 2007).
13	Recency bias	A cognitive predisposition that causes people to more prominently recall and emphasise recent events and observations than those that occurred in the near or distant past (Pompian, 2006, p. 216).
14	Stereotypical bias	The notion that memory is distorted towards stereotypes (Hinton, 2000).
15	Bandwagon effect	People's tendency to align themselves with the majority opinion and do or believe things because many other people appear to be doing or believing the same (VandenBos, 2007).
16	Curse of knowledge	A tendency to be biased by one's current knowledge state when attempting to reason about a more naive perspective (Birch, Brosseau-Liard, Haddock, & Ghrear, 2017).
17	Illusion of transparency	A tendency for people to overestimate the extent to which their internal thoughts, feelings and attitudes 'leak out' and are seen by others (Brown & Stopa, 2007, p. 806).
18	Pro-innovation bias	It is the implication in diffusion research that innovation should be diffused and adopted by all member of a social system, that it should be diffused more rapidly, and that the innovation should be neither re-invented nor rejected (Rogers, 2010, p. 100).
19	Planning fallacy	A tendency for individuals to underestimate the time required to complete a task (Ehrlinger, Readinger, & Kim, 2016, p.9).
20	Self-serving bias	The tendency to interpret events in a way that assigns credit for success to oneself but denies one's responsibility for failure, which is blamed on external factors (VandenBos, 2007).
21	False consensus effect	The tendency to assume that one's own opinions, beliefs, attributes, or behaviours are more widely shared than is actually the case. A robustly demonstrated phenomenon, the false-consensus effect is often attributed to a desire to view one's thoughts and actions as appropriate, normal, and correct (VandenBos, 2007).
22	Overconfidence effect	Unwarranted faith in one's intuitive reasoning, judgments, and cognitive abilities (Pompian, 2006, p. 51).
23	Status quo bias	An emotional bias that predisposes people facing an array of choice options to elect whatever option ratifies or extends the existing condition (i.e., the "status quo") in lieu of alternative options that might bring about change (Pompian, 2006, p. 248).
24	Reactance	A model stating that in response to a perceived threat to - or loss of - a behavioural freedom, a person will experience reactance, a motivational state characterised by distress, anxiety, resistance, and the desire to restore that freedom. According to this model, when people feel coerced into a certain behaviour, they will react against the coercion, often by demonstrating an increased preference for the behaviour that is restrained, and may perform the behaviour opposite to that desired (VandenBos, 2007).
25	Belief bias	The tendency to be influenced by one's knowledge about the world in evaluating conclusions and to accept them as true because they are believable rather than because they are logically valid (VandenBos, 2007).
26	Information bias	The tendency to request unnecessary or unhelpful information, especially in times of uncertainty (Mohanani, Salman, Turhan, Rodriguez, & Ralph, 2018, p. 21).
27	Ambiguity aversion bias	People hesitate in situations of ambiguity, a tendency referred to as ambiguity aversion (Pompian, 2006, p. 129).
28	Serial position effect	The results of a free recall task are plotted on something called a serial position curve, which is normally U-shaped. The serial position curve, once constructed, tends to exhibit both a recency and a primacy effect (Pompian, 2006, p. 217).

Factors Influencing Student Satisfaction with the Fieldwork Course in LIS in South Korea

Sung-Jong Cha

Department of Library and Information,
Soongyei Women's College, Seoul, Korea
E-mail: sjcha@sewc.ac.kr

Kyung-Jae Bae*

Department of Library and Information,
Dongduk Women's University, Seoul, Korea
E-mail: kjbae@dongduk.ac.kr

ABSTRACT

This study attempts to draw the implications of student satisfaction with fieldwork courses in the library and information science (LIS) program in South Korea by determining the level of satisfaction experienced by students with practicums at a library and information service agency. LIS students with fieldwork experience were surveyed to assess their satisfaction and the factors affecting their satisfaction. Results from a multiple regression analysis revealed that from among factors of practicum settings, major-related attributes, and general attributes of respondents, "satisfaction with majors" and "interpersonal relationships" had a significant effect on satisfaction with practicums. Examining the six components of satisfaction with the course revealed that "practicum program" and "practicum setting" had a significant relationship with overall satisfaction. A selective intervention program to improve satisfaction with the course needs to be developed based on factors influencing overall satisfaction, focusing on "interpersonal relationship" and "satisfaction with majors."

Keywords: fieldwork course, student satisfaction with practicum, library and information science education, curriculum of practical training, library internship

Open Access

Received date: November 4, 2019
Accepted date: April 6, 2020

*Corresponding Author: Kyung-Jae Bae
Professor
Department of Library and Information, Dongduk Women's University,
60, Hwarang-ro 13-gil, Seongbuk-gu, Seoul 02748, Korea
E-mail: kjbae@dongduk.ac.kr

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

1. INTRODUCTION

The curriculum of practical training is integral to library and information science (LIS) programs because the research findings can be implemented onsite at information service agencies and libraries. Fieldwork courses focus on the practical aspects of the LIS curriculum. Although there has been a change in the perception of their value with the theory vs. practice debate, fieldwork is now generally recognized as a crucial factor in the LIS program among LIS scholars (Ball, 2008).

Several terms for fieldwork courses have been used in LIS: practicum, practice work, practicum work, field experience, field work, field problems, internship, professional field experience, cooperative education, and work study (Futas, 1994; Mediavilla, 2006). The following definition of Coleman (1989) will be used in this study to gain a better understanding of the terminology "Fieldwork Course":

A relatively short-term, professionally supervised work experience offered as part of the school's curriculum and taken during the academic sequence. The practicum is typically pursued in the same library or information agency for the total experience. Other than incidental expenses, the practicum student does not receive a salary.

LIS is a field of study involving a variety of information-mediated subjects while serving as the heart of librarian training. By 2018, 40 universities in South Korea had established departments of library and information science, wherein students are awarded librarian certification upon graduation. Fieldwork courses in the LIS curriculums of most universities are designed to be completed before graduation.

The LIS fieldwork course is offered as part of the university's curriculum on the library premises. Such structural characteristics of the course demand close industry-education cooperation between practicum sites and academia. Since the university curriculum of LIS training in South Korea is suffering substantial alienation from library practicum sites by focusing on theory rather than on practice, the fieldwork course has been a useful alternative to help reduce the gap between academics and practice.

Yet, despite its usefulness, its limitations based on its environment put the course in a challenging situation wherein its educational purpose is not effectively served. This is partially due to the gap between practice and education, a long-standing issue that has erupted at different times in LIS-related professions (Bird & Crumpton, 2014). There are concerns that such a gap between academics and practicum sites may cause confusion

among students about library-related knowledge, compromising the effectiveness of practical training and leading to a drop in overall satisfaction with education. Ball (2008) has indicated that previous studies on LIS fieldwork courses have suffered from lack of comprehensive and rigorous research and survey data gathered regarding fieldwork. Therefore, an accurate understanding of the current status of the fieldwork course is required by collecting student opinion data from the practicum sites.

Despite the importance of fieldwork courses in LIS, the current LIS curriculum contents and process of practical training pursued at universities of South Korea have been designed mainly for the benefit of the supplier, such as its significance or interest to professors or educational institutions, rather than focusing on student perspectives. Notwithstanding this fact, this study aims to empirically examine student satisfaction with the LIS fieldwork practicum and its current status and draw implications from fieldwork course satisfaction. Based on this, it also analyzes factors affecting LIS student satisfaction with the practicum to seek out strategies for a more effective and satisfactory fieldwork course that can be pursued by universities and practicum sites as well as students.

To achieve this goal, the following research items were developed for this study.

First, this study investigates the degree of satisfaction expressed by LIS practicum students who had undergone fieldwork course at a library or an information service agency. A survey was conducted on LIS practicum participants and analyzed based on their general attributes, major-related attributes, practicum conditions, satisfaction with various aspects of the practicum (setting, timeframe and duration, instructor, program, evaluation, and preparation), and overall perception.

Second, this study explores how practicum-related elements such as general attributes, major-related attributes, and practicum conditions influence LIS student satisfaction with the practicum. Regression analysis was conducted (1) to identify factors affecting student satisfaction with the practicum specifically based on respondents' general attributes, major-related attributes, and practicum conditions, and (2) to determine how they influence satisfaction.

Third, the six variables of practicum (setting, timeframe and duration, instructor, program, evaluation, and preparation) are examined to determine the overall satisfaction of respondents.

Fourth, this study aims to present the best practices of fieldwork courses conducted at libraries and information service agencies and to set the right direction for LIS students.

2. LITERATURE REVIEW

2.1. Variables of College Students' Satisfaction with Curriculum

As Aharony (Aharony, 2011; Pryor, 1999) observes, "Assessing student satisfaction can play a major role in evaluating the effectiveness of learning" and proving the usefulness of learning programs. Various studies to assess student satisfaction and to prove the value of curriculums have been made. Also, several factors affecting satisfaction, such as participant attitude, knowledge and performance gains, and intentions to participate in a future program, have also been used in previous studies (Pryor, 1999). Aharony and Raban (2008) examined the attitudes of three stakeholder groups, including librarians, working-level officials, and the faculty, about incorporating a course on the economics of information goods into LIS curricula. In addition to examining faculty discussions and deliberations on a new course, students and practitioners' attitudes were also taken into consideration.

The most frequently used variable, as a factor affecting student satisfaction, is associated with the quality of education programs (Hearn, 1985; Thomas & Galambos, 2004). Satisfaction with the quality of education offers an important comparison and highlights alternative consumer outcomes (Thomas & Galambos, 2004). Hearn (1985) suggested that, in general, stimulating coursework and good teaching were somewhat more important than opportunities for faculty/student interaction or perceived faculty knowledgeability.

Besides the variable involving the education program itself, that fact that personal disposition also influences learning remains a major concern (Cherry, Freund, & Duff, 2013). Cherry et al. (2013) in their research on Master of Library and Information Science students online found statistically significant differences in preferences and satisfaction across five demographic variables: age (generational cohort), employment status, urban status, commute distance, and program modality. Bean and Bradley (1986) conducted a study with students' grade point averages (GPAs) to identify the satisfaction-performance relationship and found that the relationship between GPA and satisfaction is statistically significant only for women. Pike (1991) also examined the relationship between grades and satisfaction using a latent variable model. The results suggest that satisfaction exerts a stronger influence on grades than grades exert on satisfaction.

Studies also indicate that gender differences can be a personal variable affecting students' learning experience. Marley (2007) in a literature review identified that gender-identity differences can affect student learning experiences at several different points within a distance education course, beginning with motivation

for enrolling, and continuing on through communication styles within the class, use of technology, learning style, and even dropout or failure rates.

Overall, the results of literature review show that different variables such as education programs, personal disposition, academic performance, and attitudes of education program participants are found to have a significant influence on student satisfaction with curricula.

2.2. Research on LIS Fieldwork Course in South Korea

When developing a curriculum for the first department of library science at Columbia University in 1884, Melville Dewey stressed the importance of a library practicum in the curriculum, claiming that lectures and reading alone are simply insufficient to foster the best librarians; besides these actions, also required are educational activities such as seminars, field studies about success stories of libraries, and hands-on experience acquired from libraries (Grogan, 2007). Library practicums were highly regarded in early days but little attention was paid to them from the 1920s to the late 1960s, by treating them as vocational training. Rothstein in 1968 asserted that it is necessary to reintroduce the library practicum curriculum into the education program for American librarians, to consolidate theoretical knowledge as well as application of theory to practice (Lynch, 2008). Over the next two decades, LIS scholars maintained that the library practicum should be incorporated as a meaningful subject into LIS curriculums in the United States. Hence, the curriculum of library practicums became an integral part of the LIS education course since the 1970s and has been competitively included in the LIS curriculum as a core subject since the 1990s.

In South Korea, the fieldwork course has been incorporated into the LIS curriculum as a core subject and research has been conducted mainly since the year 2000 focusing on the effects, status, and program strategy of fieldwork courses. Meanwhile, studies examining the perception and satisfaction of field course trainees have been limited.

Koo (1983), the first Korean scholar who investigated the practicum pursued in South Korea, examined practical training and its issues in library work on 16 LIS programs selected from 27 universities in 1983. He found that all participating LIS programs in his investigation implemented a library practicum, and among them, 11 programs require students to complete the practicum as a mandatory course. Subsequently, research on practicum program design has been steadily growing. Kim (2011) examined the contents of practical training for school teacher-librarians by conducting a survey on those who had completed their practicum and presented an education program

based on their expertise, and proposed a practicum training program consisting of 28 elements. Jung (2011) argued that his journal feedback program based on the principle of journal and abstract writing showed a positive effect when applied in practice on the library practicum curriculum designed for librarian candidates.

Since 2010 there has been intensive research to identify the needs of stakeholders (library working-level officials, students, and faculty) in the LIS field course. Yoo (2012) conducted a job-shadowing internship program to minimize the challenges faced by LIS undergraduate students during their fieldwork courses and to lessen the burden of officials at practicum sites, which enabled her to explore interns' perceptions of the fieldwork course's effectiveness and the problems interns faced. Lee and Kim (2012) surveyed practical training programs in 42 town-level public libraries and the library practicum curriculum available in four-year LIS degree programs; they identified the demands of librarians at practicum sites and of practicum students on the training conducted by C University LIS students in S public library, and presented the challenges faced by education institutions, practicum agencies, and students, along with strategies for improvement. Yoo (2014) re-evaluated field training as part of an internship program by diagnosing problems perceived by working-level staff during the implementation of the program at their libraries, revealing that the staff in charge of interns have a positive opinion about the internship program but fail to recognize (or tend not to take it seriously) that they actually have a major influence on their interns during fieldwork. Therefore, the staff who lead practical training play a critical role in student satisfaction. Damasco and McGurr (2008) stated that to be successful, it is vital for a library training program to have librarian candidates convinced of their library practicum as a meaningful, well-coordinated experience, which suggests that the faculty and staff who systematically prepare and design the training program play a pivotal role.

To sum up, various studies on LIS fieldwork course have begun in South Korea since 2010, but they are mainly focused on the design and needs assessment of practical training programs, leaving the current research relatively insufficient to carry out a curriculum evaluation, especially from the perspective of the demander, that is, based on the assessment of student satisfaction with the practicum.

3. RESEARCH METHODOLOGY

3.1. Research Subjects and Data Collection Methods

A survey was conducted on 170 LIS students from eight

universities, who had undergone the fieldwork course offered by a library and information service agency, to assess their satisfaction with the practicum.

The data for the self-administered survey were collected from September 29, 2018 through October 10, 2018, by means of a structured questionnaire. The study purpose was explained to the participants and their informed consent was obtained prior to the survey. After excluding respondents with missing or incomplete responses, 161 participants were selected.

3.1.1. Research tools

3.1.1.1. Factors Affecting LIS Student Satisfaction with Practicums

A new strategy is needed to improve the effectiveness of practicums based on the causal relationship between factors affecting fieldwork course effectiveness. Therefore, six assessment items were selected and factors based on previous findings from related studies were used to determine satisfaction of students who had completed their fieldwork course. Factors affecting satisfaction identified from student perceptions of a training program in information science in Canada, in research conducted by Cherry et al. (2013), were also used for this study as the foundation for factor selection. Cherry et al. (2013) investigated factors influencing the academic and intellectual environment, program content and structure, coursework, and facilities and services to determine student satisfaction. Considering these factors and consulting the variables extracted from previous findings, this study examines the effect of six factors as follows.

Practicum setting: A better practicum setting can enhance the effects of fieldwork courses because the surroundings, such as the atmosphere, staffing, library collection, and facility of the practicum site (library) are crucial.

Practicum timeframe and duration: Imposing a reasonable timeframe and duration on the fieldwork course can have a major impact on improving student satisfaction.

Practicum instructor: The qualities of a fieldwork instructor library work – such as he/she being a first-rate educator with effective leadership skills, high self-esteem and expertise, an enthusiastic attitude and integrity toward the course, and communication skills – can influence the effects of fieldwork course.

Practicum program: A well-structured practicum program has a profound impact on the effects of fieldwork courses; for example, a practicum program that delivers proper instructions and feedback on students' assignments and practicum tasks, and provides accurate requirements and scope for understanding fieldwork as well as systematic training procedures including

teaching methods, organization, and selection of training details to be learned by a trainee during the course.

Practicum evaluation: A fieldwork practicum is usually a credit course and thus objective and fair evaluation prepared and applied through an adequate evaluation system across the curriculum has a positive influence on the effects of the course.

Practicum preparation: It refers to the tasks of main entities such as leadership at the practicum site, the supervisor, and the trainee engaged in the fieldwork courses. Trainees choose a suitable practicum site and prepare themselves for the practicum through prior consultation with their supervisor about the types of practical training available for them, while the supervisor provides the full scope of instructions through an orientation for students regarding the contents and processes they will go through during the course. The leadership prepares the contents and processes to be imparted onsite in a systematic manner to provide proper guidance at an early stage of practicum. With these in place, the effects of the fieldwork course can be greatly enhanced.

3.1.1.2. Survey Questionnaire Construction

Using preliminary analysis and literature review, a questionnaire was designed with 46 items in all with subsections consisting of 4 items for general attributes, 6 items for major-related attributes, 5 items for practicum conditions, and 31 items for satisfaction with the practicum. Questions pertaining to satisfaction with the practicum again include 5 items each for satisfaction with the practicum setting, practicum timeframe and duration, and practicum instructor; 6 items for satisfaction with the practicum program, 4 for satisfaction with practicum evaluation, 5 for satisfaction with practicum preparation, and 1 for overall satisfaction. A 5-point Likert scale was used to assess the levels of student satisfaction in the items pertaining to satisfaction with the practicum, with 1 signifying “Not at all” and 5 signifying “Very Satisfied.”

3.2. Data Analysis and Reliability Evaluation

For interpreting the survey results on the LIS fieldwork course, statistics software (SPSS 22.0) was used for regression analysis, cross validation, descriptive statistics, and frequency analysis.

Cronbach’s α was used to test the reliability of the research tool (questionnaire). The results showed that the overall confidence coefficient is larger than 0.6 whereas the confidence coefficient for student satisfaction with the practicum is 0.850; therefore, the questionnaire as a research tool showed internal consistency. Table 1 shows the results of the reliability test.

Table 1. Reliability test results of the questionnaire on student satisfaction with practicum

Satisfaction with practicum item	Number of items	Cronbach’s α
Setting	5	0.875
Timeframe and duration	5	0.723
Instructor	5	0.909
Program	6	0.898
Evaluation	4	0.848
Preparation	5	0.849
Fieldwork course	30	0.850

3.3. Measurements of Satisfaction with Fieldwork Course

3.3.1. General Attributes of Respondents

The demographic characteristics of the respondents who participated in this survey on satisfaction with LIS fieldwork courses are as follows: 8.7 percent (14 respondents) are men and 91.3 percent (147 respondents) are women. The age distribution of the respondents shows that 49.7 percent (80) are aged between 20 and 22, 44.1 percent (71) are between 23 and 26, 1.2 percent (2) are between 27 and 29, 1.2 percent (2) are between 30 and 32, and 3.7 percent (6) are aged 33 and older, which indicates that the majority of survey participants are female college students in their early 20s. Furthermore, respondents with a religious affiliation number 31.7 percent (51) whereas those with no religious affiliation are at 68.3 percent (110), which is more than twice the number of those with a religious affiliation. Regarding interpersonal relationships, 42.2 percent (68) of the respondents answered “Strong,” 55.9 percent (90) answered “Modest,” and 1.9 percent (3) answered “Weak,” which shows

Table 2. General attributes of respondents

Item	Category	Frequency	%
Gender	Men	14	8.7
	Women	147	91.3
Age (yr)	20-22	80	49.7
	23-26	71	44.1
	27-29	2	1.2
	30-32	2	1.2
	33 and older	6	3.7
Religious affiliation	Yes	51	31.7
	No	110	68.3
Interpersonal relationship	Strong	68	42.2
	Modest	90	55.9
	Weak	3	1.9

that most respondents consider their interpersonal relationships to be above average. Table 2 shows further details. In this study, 'religious affiliation' was added specifically as a component of the general characteristics of the respondents. The reason why the general attributes of the respondents include 'religious affiliation' is that it is a factor affecting the psychological and human relations aspects of the individual, which can affect the satisfaction of the field practice.

The analytical results of the demographic characteristics of the respondents were found to exhibit a typical pattern among LIS students in South Korea, characterized by a high percentage of women college students. The results are expected to be sufficient to represent the population group.

3.3.2. Analysis of Major-related Attributes

For motives behind choosing a major, a majority of respondents, 46.6 percent (75), answered they selected their major based on "personality traits (aptitude)," while 19.9 percent (32) of respondents answered based on "GPAs," 20.5

percent (33) based on "advice from parents or others," and 5.0 percent (8) based on "advantage offered when looking for a job." On the point regarding student satisfaction with college life, 50.3 percent (81) of the respondents marked "Generally Satisfied" while 28.6 percent (46) marked "Slightly Satisfied." Regarding student satisfaction with majors, 42.9 percent (69) of respondents marked "Slightly Satisfied" while 34.8 percent (56) marked "Generally Satisfied." This indicates that, overall, the survey respondents were satisfied with their college life and major to a considerable extent. On the contrary, 60.2 percent (97) of the respondents were "Neutral" and 28.0 percent (45) marked "Weak" about career prospects, which indicates that a majority of the respondents had a neutral or grim perception about job prospects. Table 3 shows further details.

3.3.3. Analysis of Practicum Setting Attributes

Among the respondents, 47.2 percent (76) and 20.5 percent (33) did their practicums at large and small public libraries, respectively, with 9.3 percent (15) at the national library and 9.3 percent (15) at the university library. Choice of sites (library) largely depends on consultation with schools; 45.3 percent (73) of respondents chose their practicum sites using school-based consultation services, while the rest, 39.1 percent (63)

Table 3. Majors of respondents

Item	Category	Frequency	%
Motives behind choosing a major	Personality traits	75	46.6
	GPAs	32	19.9
	Advice from parents or others	33	20.5
	Advantage when looking for jobs	8	5.0
	Misc.	13	8.1
GPAs	Above 4.0	28	17.4
	3.5-3.9	55	34.2
	3.0-3.4	49	30.4
	2.5-2.9	26	16.1
	Below 2.4	3	1.9
	Satisfaction with college life	Very satisfied	7
Slightly satisfied		46	28.6
Generally satisfied		81	50.3
Slightly dissatisfied		23	14.3
Highly dissatisfied		4	2.5
Satisfaction with major	Very satisfied	20	12.4
	Slightly satisfied	69	42.9
	Generally satisfied	56	34.8
	Slightly dissatisfied	14	8.7
	Highly dissatisfied	2	1.2
Career prospects	Strong	19	11.8
	Neutral	97	60.2
	Weak	45	28.0

GPAs, grade point averages; Misc., miscellaneous.

Table 4. Practicum setting of respondents

Item	Category	Frequency	%
Practicum sites	Public library (large)	76	47.2
	Public library (small)	33	20.5
	University library	15	9.3
	School library	13	8.1
	National library	15	9.3
	Special library	9	5.6
Choice of practicum sites	Self-selection	63	39.1
	Through school	23	14.3
	Consultation with school	73	45.3
	Misc.	2	1.2
Residence during practicum	Living at home	146	90.7
	Lodging	13	8.1
	Living at a relative's house	1	0.6
	Misc.	1	0.6
Distance of practicum site from residence (min)	Less than 30	51	31.7
	30-59	65	40.4
	60-89	35	21.7
	More than 90	10	6.2
	Previous practicum experience	Yes	29
No	132	82.0	

Misc., miscellaneous.

and 14.3 percent (23), chose the sites by themselves or were recommended by the schools they were then attending. In addition, the commute distance (in minutes) between practicum site and residence is 30 to 59 minutes for 40.4 percent (65), less than 30 minutes for 31.7 percent (51), and 60 to 89 minutes for 21.7 percent (35). Regarding previous practicum experience, 82 percent (132) of the respondents said “No” compared to only 18 percent (29) of those who said “Yes,” which indicates that most respondents had no previous experience with library practicums prior to this field course. Table 4 shows further details.

3.3.4. Analysis of Satisfaction with Fieldwork Course by Components of Practicum

Among the six components for which respondents rated their satisfaction with the fieldwork course, overall satisfaction with the fieldwork course appears to be 3.72 across all items; satisfaction with the practicum instructor is 3.92, which is the highest among the components, and satisfaction with practicum preparation is 3.50, which is the lowest. However, satisfaction with the fieldwork course shows no significant difference among the six components. Table 5 shows further details of respondents’ satisfaction with the fieldwork course by each component.

3.3.5. Analysis of Satisfaction with Fieldwork Course Based on Respondents’ General Attributes

Men respondents showed a slightly higher satisfaction with the fieldwork course and those aged 20 to 22 exhibited the highest satisfaction (3.78), with no statistically significant differences in age and gender. In contrast, satisfaction of trainees with a religious affiliation (3.85) was found to be higher than that of those without a religious affiliation (3.65), at the significance level, and satisfaction of trainees with strong interpersonal relationships (3.90) was found to be higher than that of those with modest or weak interpersonal relationships, indicating a significant difference in student satisfaction with the practicum

Table 5. Respondents’ satisfaction with fieldwork course

Satisfaction with practicum item	N	Mean	Standard deviation
Setting	161	3.85	0.779
Timeframe and duration	161	3.70	0.640
Instructor	161	3.92	0.771
Program	161	3.63	0.794
Evaluation	161	3.72	0.683
Preparation	161	3.50	0.711
Total	161	3.72	0.730

Table 6. Respondents’ satisfaction with fieldwork course based on general attributes

Item	Category	Satisfaction with practicum
Gender	Men	3.76 ± 0.491
	Women	3.71 ± 0.618
	t	0.273
	p	0.785
Age (yr)	20-22	3.78 ± 0.655
	23-26	3.67 ± 0.539
	27-29	3.12 ± 0.165
	30-32	3.63 ± 0.377
	33 and older	3.59 ± 0.795
	F	0.917
Religious affiliation	Yes	3.85 ± 0.627
	No	3.65 ± 0.589
	t	2.015
	p	0.046
Interpersonal relationships	Strong	3.90 ± 0.566
	Modest	3.59 ± 0.612
	Weak	3.32 ± 0.135
	F	5.848
	p	0.004

Values are presented as mean±standard deviation.

depending on different levels of interpersonal relationships. Table 6 shows further details of student satisfaction with the fieldwork course based on respondents’ general attributes.

3.3.6. Analysis of Satisfaction with Fieldwork Course Based on Major-related Attributes

In the analysis of student satisfaction with the fieldwork course based on respondents’ majors, trainees who answered “Very Satisfied” (4.06) appeared to be higher in number than those who said “Slightly Satisfied,” “Generally Satisfied,” and “Dissatisfied.” With regard to career prospects, trainees who marked satisfaction as “Strong” (3.93) turned out to be higher in number than those who marked “Neutral” or “Weak,” which indicates a significant difference in student satisfaction with the practicum, at a significance level of 0.01, according to the different career prospects and satisfaction with the major. Furthermore, student satisfaction with college life shows a somewhat significant difference, at a significance level of 0.05, in satisfaction with the practicum, though not considerably higher. Table 7 shows further details of student satisfaction based on their major.

Table 7. Satisfaction with fieldwork course based on respondents' major

Item	Category	Satisfaction with practicum
Motives behind choosing a major	Personality traits	3.76 ± 0.638
	Based on GPAs	3.75 ± 0.522
	Advice from parents or others	3.66 ± 0.613
	Advantage when looking for a job	3.61 ± 0.441
	Misc.	3.59 ± 0.734
	F	0.359
	<i>p</i>	0.837
GPAs	Above 4.0	3.63 ± 0.574
	3.5-3.9	3.76 ± 0.505
	3.0-3.4	3.72 ± 0.709
	2.5-2.9	3.76 ± 0.640
	Below 2.4	3.38 ± 0.782
	F	0.464
	<i>p</i>	0.762
Satisfaction with college life	Very satisfied	3.65 ± 0.493
	Slightly satisfied	3.88 ± 0.597
	Generally satisfied	3.67 ± 0.564
	Slightly dissatisfied	3.70 ± 0.699
	Highly dissatisfied	2.92 ± 0.662
	F	2.787
	<i>p</i>	0.028
Satisfaction with major	Very satisfied	4.06 ± 0.594
	Slightly satisfied	3.78 ± 0.571
	Generally satisfied	3.61 ± 0.581
	Slightly dissatisfied	3.50 ± 0.583
	Highly dissatisfied	2.48 ± 0.024
	F	5.260
	<i>p</i>	0.001
Career prospects	Strong	3.93 ± 0.532
	Neutral	3.78 ± 0.595
	Weak	3.47 ± 0.601
	F	5.648
	<i>p</i>	0.004

Values are presented as mean±standard deviation. GPAs, grade point averages; Misc., miscellaneous.

3.3.7. Analysis of Satisfaction with Fieldwork Course Based on Practicum Setting

In the analysis of student satisfaction with the fieldwork course based on practicum setting, no significant difference was found between the attributes such as choice of practicum site, residence during practicum, distance between practicum site and residence (minutes), and previous practicum experience.

Table 8. Satisfaction with fieldwork course based on practicum setting

Item	Category	Satisfaction	
Practicum sites	Public library (large)	3.78 ± 0.610	
	Public library (small)	3.44 ± 0.595	
	University library	3.65 ± 0.535	
	School library	3.98 ± 0.590	
	National library	3.95 ± 0.555	
	Special library	3.51 ± 0.529	
	F	2.826	
	<i>p</i>	0.018	
Choice of practicum sites	Self-selection	3.75 ± 0.540	
	School	3.58 ± 0.643	
	Consultation with school	3.74 ± 0.648	
	Misc.	3.03 ± 0.189	
	F	1.385	
		<i>p</i>	0.250
	Residence during practicum	Living at home	3.71 ± 0.617
Lodging		3.66 ± 0.469	
Living at a relative's house		4.63 ± 0.535	
Misc.		4.06 ± 0.583	
F		0.916	
		<i>p</i>	0.435
Practicum site and residence distance (min)		Less than 30	3.71 ± 0.583
	30-60	3.69 ± 0.557	
	60-89	3.78 ± 0.716	
	More than 90	3.69 ± 0.698	
	F	0.211	
		<i>p</i>	0.889
	Previous practicum experience	Yes	3.60 ± 0.689
No		3.74 ± 0.587	
t		-1.125	
<i>p</i>		0.262	

Values are presented as mean±standard deviation. Misc., miscellaneous.

However, self-selection (3.75) when choosing a practicum site appeared to give more satisfaction than when choosing a site depending on trainees' schools, consultation with school, and so on. With regard to the practicum site, trainees showed higher satisfaction in the order of school library (3.98) and national library (3.95), indicating a significant difference in satisfaction with the fieldwork course based on different practicum sites, at a significance level of 0.05. Table 8 shows further details about respondents' satisfaction with the course based on the practicum setting.

3.3.8. Analysis of Factors Affecting Satisfaction with Fieldwork Course

A multiple regression analysis was performed by setting the mean of satisfaction with the fieldwork course as a dependent variable and using inputs such as each factor of the practicum setting, major-related attributes, and general attributes as independent variables to determine the factors affecting satisfaction with the fieldwork course according to respondents' general attributes, major-related attributes, and different practicum settings, and to reveal the extent to which they influence satisfaction. In the multiple regression model, factors such as satisfaction with major (standardized regression coefficient = 0.340) and interpersonal relationship (standardized regression coefficient = 0.218) were found to influence student satisfaction with the fieldwork course at a significance level of 0.01, along with the religious affiliation factor (standardized regression coefficient = -0.160) at a significance level of 0.05. As a result, the estimated regression coefficient for practicum students' satisfaction with the major is 0.340, and those showing higher satisfaction with their majors were shown to have an increase in satisfaction with the fieldwork course by 0.34 points. Table 9 shows further details about factors affecting student satisfaction with the fieldwork course.

3.3.9. Comparison of Influences on Overall Satisfaction among the Components of Satisfaction with Practicum

A multiple regression analysis was performed by setting overall satisfaction as a dependent variable and using its six components (practicum setting, practicum timeframe and duration, practicum instructor, practicum program, practicum evaluation, and practicum preparation) as independent variables to identify factors affecting the overall satisfaction of respondents and their effects. In the multiple regression model, after excluding the four factors, practicum timeframe and duration, practicum instructor, practicum evaluation, and practicum preparation, the remaining two, practicum setting and practicum program, were adopted as significant variables at a significance level of 0.05, indicating that they have significant effects on overall satisfaction with the fieldwork course. Therefore, among the six components of the overall satisfaction with the fieldwork course, practicum program (standardized coefficient = 0.362) and practicum setting (standardized coefficient = 0.214) were found to have an influence on overall satisfaction. Table 10 shows further details about effects of components of satisfaction with practicum on overall satisfaction.

Table 9. Factors affecting student satisfaction with the fieldwork course

Item	Unstandardized coefficient		Standardized coefficient	t	Significant probability	F value (p)	R squared
	B	Standard error	β				
(Constant)	2.427	0.589		4.118	0.000	3.154 (0.000)	0.246
Gender	-0.066	0.167	-0.031	-0.394	0.694		
Age	-0.095	0.058	-0.140	-1.645	0.102		
Religious affiliation	-0.207	0.098	-0.160	-2.107	0.037*		
Interpersonal relationships	0.249	0.089	0.218	2.797	0.006**		
Motives behind choosing a major	-0.004	0.037	-0.008	-0.106	0.916		
GPA's	-0.085	0.049	-0.143	-1.716	0.088		
Satisfaction with college life	-0.030	0.069	-0.041	-0.433	0.666		
Satisfaction with major	0.237	0.065	0.340	3.619	0.000***		
Career prospects	0.099	0.085	0.101	1.160	0.248		
Practicum sites	0.028	0.030	0.074	0.926	0.356		
Choice of practicum sites	-0.030	0.048	-0.048	-0.629	0.530		
Residence during practicum	0.183	0.118	0.118	1.552	0.123		
Distance between practicum sites and residence	0.023	0.054	0.034	0.425	0.671		
Previous practicum experience	0.159	0.118	0.102	1.347	0.180		

GPA's, grade point averages.
*p<0.05, **p<0.01, ***p<0.001.

Table 10. Effects of components of satisfaction with practicum on overall satisfaction

Model	Unstandardized coefficient		Standardized coefficient	t	Significant probability	F value (p)	R squared
	B	Standard error	β				
(Constant)	0.062	0.345		0.180	0.857	31.859 (0.000)	0.554
Practicum setting	0.255	0.123	0.214	2.071	0.040*		
Practicum timeframe and duration	0.133	0.128	0.091	1.039	0.300		
Practicum instructor	0.055	0.131	0.045	0.419	0.676		
Practicum program	0.424	0.127	0.362	3.343	0.001**		
Practicum evaluation	0.230	0.117	0.169	1.972	0.051		
Practicum preparation	-0.088	0.091	-0.068	-0.975	0.331		

* $p < 0.05$, ** $p < 0.01$.

4. DISCUSSION

4.1. Analysis of Satisfaction with Fieldwork Course

LIS students need the preparation that enables them to apply LIS knowledge to practice through fieldwork courses. Therefore, efforts are needed to support their successful adaptation to field courses as well as to increase their satisfaction. This requires an understanding of factors affecting their satisfaction with the fieldwork course to lay the groundwork and provide baseline data for improving satisfaction.

In the overall satisfaction with the fieldwork course, the factor registering the highest satisfaction with respondents was the practicum instructor (mean 3.92), whereas “satisfaction with practicum preparation” scored relatively low (mean 3.50). This reflects on the relative excellence of library fieldwork instructors. However, this suggests the need for further substantial preparation for fieldwork courses at the school, practicum site, and trainee levels. Yet, no significant difference was observed in the mean of satisfaction among the six components.

4.2. Differences of Satisfaction Based on General Attributes

Despite no significant differences seen in LIS student satisfaction with the fieldwork course by gender and age, respondents with an religious affiliation were found to have higher levels of satisfaction at a statistically significant level, and satisfaction of trainees with strong interpersonal relationships appeared to be the highest, which shows a statistically significant difference in satisfaction with the practicum based on differences in religious affiliation and interpersonal relationships. These may indicate that satisfaction is influenced

by the psychological and interpersonal aspects of the trainees to some extent. Therefore, practicum preparation in the future needs careful consideration of trainees’ psychological and interpersonal aspects to enhance satisfaction with the fieldwork course.

4.3. Differences in Satisfaction Based on Major-related Attributes

There was no significant difference seen in satisfaction with the fieldwork course according to LIS student majors and factors such as the motives for choosing the major and the grade acquired. In contrast, satisfaction with majors among trainees who were very satisfied with their LIS major was found to be high at a statistically significant level, and satisfaction of trainees with positive career prospects was higher than that of those with negative or neutral prospects. This reflected a statistically significant difference in satisfaction with the fieldwork course based on different career prospects and satisfaction with majors. These conclusions suggest that trainees satisfied with the LIS major and with positive career prospects in turn eventually show satisfaction with the fieldwork course. Therefore, more proactive efforts are needed to make trainees feel a sense of pride in their majors so that their overall satisfaction with the practicum improves.

4.4. Differences in Satisfaction Based on Practicum Setting

No significant difference was observed in LIS student satisfaction with the fieldwork course among factors such as choice of practicum site (library), residence during practicum, distance (in minutes) between practicum site and residence, and previous practicum experience, but trainees who chose the national library or school library as their practicum site showed

a relatively higher level of satisfaction. The practicum site selected by a trainee based on familiarity or preference served as a positive factor for enhanced satisfaction. Therefore, vigorous efforts are needed when allocating practicum sites for future fieldwork courses to ensure that trainees can choose a library that is familiar to them or is one they prefer.

4.5. Factors Affecting Satisfaction with Fieldwork Courses

A multiple regression analysis was performed to determine the factors that affect satisfaction with the practicum from among the practicum setting, major-related attributes, and general attributes of LIS students. The results showed that such factors as religious affiliation ($\beta = -0.160$) at a significance level of 0.05, satisfaction with majors ($\beta = 0.340$), and interpersonal relationships ($\beta = 0.218$) at a significance level of 0.01 were found to have a statistically significant influence on satisfaction with fieldwork courses. This suggests that psychological and interpersonal aspects in education and student-oriented education in college majors to develop an interest in LIS majors should be reinforced to increase student satisfaction with the practicum.

4.6. Factors Affecting Overall Satisfaction among Components of Satisfaction with Practicums

Overall satisfaction with the fieldwork course was measured using six variables in this study. Among the six components of the practicum (setting, timeframe and duration, instructor, program, evaluation, and preparation), four variables were excluded and only two variables “practicum program ($\beta = 0.362$)” and “practicum setting ($\beta = 0.214$)” were found to have a statistically significant influence on overall satisfaction. This suggests that a trainee having relatively high satisfaction with the practicum setting tends to have a higher overall satisfaction with the fieldwork course, which is also applicable in the case of high satisfaction with fieldwork course programs because of their systematic and organized manner of functioning, as well as the appropriate design and scope of the course.

The results indicate that strategies for improving satisfaction with practicums in the future include focusing on the “practicum program” and “practicum setting,” to enhance the effectiveness of the LIS fieldwork course by improving satisfaction with the fieldwork course.

5. CONCLUSION

This study was conducted to seek strategies to make fieldwork

courses more effective and satisfactory for students, practicum agencies, and colleges by examining the factors influencing LIS student satisfaction with the fieldwork course and their interrelationships. This will allow for an accurate diagnosis of problems and challenges faced in fieldwork management and methods, and will eventually contribute to planning practicum programs and approaches to improve the effectiveness of the course such that it can provide practical support for future LIS students in building their careers. This paper proposes an improvement strategy in the library fieldwork course for students participating in it through a discussion of the analytical results of factors affecting overall student satisfaction with the practicum. It also analyzed the differences in satisfaction depending on practicum setting, general attributes, and satisfaction with the fieldwork course.

First, from students' overall satisfaction with the fieldwork course, satisfaction with the practicum instructor is highest whereas satisfaction with practicum preparation is relatively low from among the six factors of the practicum affecting satisfaction. Therefore, a more substantial preparation for fieldwork courses should be made at the level of the trainee, practicum site, and college in the future to systemize the pre-training orientation provided for trainees. These improvements can be achieved by developing and reflecting on detailed guidelines of LIS practical training. The following items should be clearly specified: practicum instruction guidelines and tools, relationships with educational institutions, teaching methods for practicums, contents of practicums, practicum processes, training time, the number of trainees, criteria for practicum instructors, criteria for practicum agencies, goals of practicums, and purposes of fieldwork course education.

Second, religious affiliation and interpersonal relationships of the college students are factors affecting student satisfaction with the fieldwork course, indicating that psychological and interpersonal aspects of trainees are also to be considered. Therefore, it is necessary to develop a model for LIS practical training curriculums that includes design of the practicum program, placement of practicum instructors, implementation of pre-training orientation, and student counseling services to foster healthy interpersonal relationships and promote student health and stability by carefully considering the psychological and interpersonal aspects of trainees.

Third, satisfaction with the LIS major among the participants appears to have a positive effect on their satisfaction with the fieldwork course, along with positive career prospects. Therefore, schools need to make a more proactive effort, for example, in introducing a student-oriented intensive education program for majors, dividing students of the same major into small groups

for better communication among them, and implementing extracurricular programs to enhance students' sense of pride in the majors they have chosen.

Fourth, there is a significant difference in satisfaction with the fieldwork course based on the practicum site the student has chosen, suggesting that familiarity with the site or a preference for a particular site is a positive factor for gaining satisfaction. Therefore, schools need to analyze the location and type of libraries during the selection process of practicum sites when liaising with sites for their fieldwork course by choosing the ones that are highly preferred by or are familiar to students.

Fifth, the six components of overall satisfaction with fieldwork courses and their interrelationships were examined, from which only two variables, the "practicum program" and "practicum setting," were found to have a significant influence on overall satisfaction. This indicates that trainees showed higher levels of satisfaction with the practicum setting or with a well-organized fieldwork course offered by a practicum site, resulting in increased overall satisfaction with the LIS fieldwork course. Therefore, it is necessary to have an improvement strategy for satisfaction with practicums by focusing on a strategy that will enhance the "practicum program" and "practicum setting" to improve the educational effects of LIS fieldwork courses. To do this, practical training agencies need to develop a well-structured practical training program, determine its appropriate scope, and implement the fieldwork course in a systematic manner. At the same time, efforts should be made to increase satisfaction with the "practicum setting" by equipping practicum sites with the relevant infrastructure to successfully conduct fieldwork courses and create a cordial and interactive atmosphere in fieldwork education.

REFERENCES

- Aharony, N. (2011). Library and information science students' feedback in an online course. *Journal of Education for Library and Information Science*, 52(4), 305-319.
- Aharony, N., & Raban, D. R. (2008). Economics of information goods: An interdisciplinary subject for Israeli LIS and MBA curricula. *Library & Information Science Research*, 30(2), 102-107.
- Ball, M. A. (2008). Practicums and service learning in LIS education. *Journal of Education for Library and Information Science*, 49(1), 70-82.
- Bean, J. P., & Bradley, R. K. (1986). Untangling the satisfaction-performance relationship for college students. *The Journal of Higher Education*, 57(4), 393-412.
- Bird, N. J., & Crumpton, M. A. (2014). Real learning connections: Questioning the learner in the LIS internship. *Journal of Education for Library and Information Science*, 55(2), 89-99.
- Cherry, J. M., Freund, L., & Duff, W. M. (2013). Students' perceptions of information programs in Canada. *Journal of Education for Library and Information Science*, 54(3), 174-190.
- Coleman, J. G., Jr. (1989). The role of the practicum in library schools. *Journal of Education for Library and Information Science*, 30(1), 19-27.
- Damasco, I. T., & McGurr, M. J. (2008). A survey of cataloger perspectives on practicum experiences. *Cataloging & Classification Quarterly*, 45(4), 43-64.
- Futas, L. (1994). The practicum in collection development: A debate. In P. Johnson, & S. S. Intner (Eds.), *Recruiting, educating and training librarians for collection development* (pp. 145-156). Westport: Greenwood Press.
- Grogan, D. J. (2007). Education for librarianship: Some persistent issues. *Education for Information*, 25(1), 5-26.
- Hearn, J. C. (1985). Determinants of college students' overall evaluations of their academic programs. *Research in Higher Education*, 23(4), 413-437.
- Jung, J. (2011). A case study on the development & application of journal feedback program in pre-librarians' library field experience. *Journal of Korean Library and Information Science Society*, 42(2), 277-297.
- Kim, S. (2011). A study on the development of teacher librarians' teaching practice program in Korea. *Journal of Korean Library and Information Science Society*, 42(2), 193-213.
- Koo, B. (1983). A research on the actual conditions of library practice education in Korea. *Journal of the Korean Society for Library and Information Science*, 10, 3-37.
- Lee, S., & Kim, Y. (2012). A study on library and information school's practical training program: Focused on a case of S public library. *Journal of the Korean Society for Library and Information Science*, 46(4), 147-168.
- Lynch, B. P. (2008). Library education: Its past, its present, its future. *Library Trends*, 56(4), 931-953.
- Marley, J. L. (2007). Gender differences and distance education: Major research findings and implications for LIS education. *Journal of Education for Library and Information Science*, 48(1), 13-20.
- Mediavilla, C. (2006). *Public library internships: Advice from the field*. Lanham: Scarecrow Press.
- Pike, G. R. (1991). The effects of background, coursework, and involvement on students' grades and satisfaction. *Research in Higher Education*, 32(1), 15-30.

- Pryor, B. W. (1999). Three longitudinal impact evaluations of continuing library education: Participant satisfaction, program effects, and future participation. *Journal of Education for Library and Information Science*, 40(1), 10-26.
- Thomas, E. H., & Galambos, N. (2004). What satisfies students? Mining student-opinion data with regression and decision tree analysis. *Research in Higher Education*, 45(3), 251-269.
- Yoo, S. (2012). A case study of job-shadowing internship program for information specialists. *Journal of the Korean BIBLIA Society for library and Information Science*, 23(1), 113-129.
- Yoo, S. (2014). Internee's assessment of library internship program: Capturing of the use of LIS practicum. *Journal of the Korean BIBLIA Society for library and Information Science*, 25(2), 189-205.

Evaluating the Information Access Skills of Students of a College of Education in Ghana

Franklina Adjoa Yebowaah

University Library, University for Development Studies,
Tamale, Ghana
E-mail: fyabowaah@uds.edu.gh

Christopher M. Owusu-Ansah*

University Library, University of Education,
Winneba - Mampong, Ashanti, Ghana
Department of Information Science, University of South Africa,
Pretoria, South Africa
E-mail: cmfum@uew.edu.gh

ABSTRACT

Access to relevant information determines the quality of life of an individual in society and the academe. Academic institutions have established information literacy programmes to enhance students' access to information. With the elevation of colleges of education to tertiary status in Ghana, pre-service teachers are now required to conduct research to improve their reflective thinking in order to be certified. However, the information access skills of students in colleges of education in Ghana are uncertain. This study draws empirical evidence from students in the Nusrat Jahan Ahmadiyya College of Education, Wa, Ghana on their information access skills. Primary data were collected from 303 respondents from the college using a stratified sampling procedure. The data were collected using a questionnaire and analysed using descriptive statistics. The results indicate that information is available to students through different sources such as the Internet, libraries, and contact with friends and authorities. However, students were limited in respect to access to relevant online resources for their academic activities due to inadequate information access skills. Potential solutions are discussed.

Keywords: information access, information literacy, library use, internet access, college of education

Open Access

Received date: January 1, 2019
Accepted date: April 6, 2020

***Corresponding Author:** Christopher M. Owusu-Ansah
Senior Assistant Librarian
University of Education, Winneba
College of Agriculture Education, University Library, P.O. Box 40,
Mampong, Ashanti, Ghana
E-mail: cmfum@uew.edu.gh

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

1. INTRODUCTION

Information seeking has become a significant component of general routine. The increasing capacity of information communication technologies (ICTs) for information generation and use has culminated in the need for the development of factors to determine information quality (Long & Shrikhande, 2007). In the field of academia, access to relevant information determines the quality of one's research output and relevance in the academic environment and society (Okon, Etuk, & Akpan, 2014). Information users are, therefore, concerned with what selection criteria to apply in choosing information sources (Taylor & Dalal, 2017). People, especially college students, obtain information from various sources and for different purposes. Similarly, schools and colleges also devise methods to expose students to different components of information skills (Abdul Karim, Shah, Din, Ahmad, & Khalid, 2013).

Libraries, especially university libraries, contain a diverse repertoire of information sources in the form of print, electronic, and audio-visual resources (Sasikala & Dhanraju, 2011). Today, school libraries enhance the development of creative skills among learners by facilitating access to information. Furthermore, libraries promote individualised access to varied modes of communication capable of supporting the critical thinking and problem-solving needs of a techno-centric generation. Omeluzor, Akibu, Dika, and Ukangwa (2017) indicate that academic libraries enhance the information backbone of universities through the development of an effective and efficient knowledge support system to achieve the lofty goals of the academe. Spenser (2003) also emphasises the teaching role of libraries, stressing the critical role of information literacy in developing students' skills at consuming information for the purpose of addressing learning and research needs.

1.1. Background to the Study

In Ghana, students at different levels have challenges in accessing and evaluating information. For example, Yeboah, Dadzie, and Owusu-Ansah's (2017) research on "information access and evaluation skills" of students of two secondary schools (Opoku Ware Senior High School and Yaa Asantewaa Girls' Senior High School) revealed that students in both schools could not effectively access information for their academic work due to their lack of basic skills in differentiating good information from bad. Besides this, with the elevation of the colleges of education to tertiary status in Ghana, pre-service teachers are now required to conduct research as part of the requirements to be awarded their diploma (Nyarkoh, 2016). This has increased their information access requirements, as

they will need to explore new areas and conduct independent studies. However, the information access skills of college of education students in Ghana have not yet been identified and evaluated. This study draws empirical evidence from students in the Nusrat Jahan Ahmadiyya College of Education (NJACoE), Wa in the Upper West Region of Ghana on their information access and evaluation skills.

1.2. The Nusrat Jahan Ahmadiyya College of Education

The NJACoE started as a girls' academy and later became a mixed school. It was converted into a teacher training college in 1982 and awarded post middle certification. It was upgraded to a post-secondary institution in 1991 and later became a diploma awarding institution in 2004. The school is now a university college. Currently, the school runs three academic programmes: technical/mathematics, general/social science, and science/mathematics. The school has an ICT laboratory with over 100 computers for student use. The student population is 1,264 with 72 staff. Out of the 72 staff, 40 are tutors and 32 are supporting staff.

The school has a library facility with reference materials in fiction, non-fiction, physics, mathematics, chemistry, social studies, English language, biology, ICT, music, technical/vocational, education, general science, and miscellaneous (Nusrat Jahan Ahmadiyya College of Education, 2018). Apart from its print holdings of books, the college, however, cannot boast of standard academic library facilities and services. For instance, the library does not have computers to facilitate access to relevant educational databases. In addition, the library lacks qualified personnel who, among other functions, are responsible for providing training in online searching for academic literature to support teaching and learning. On the basis of this, it may be argued that the NJACoE library is typical of other college of education libraries in Ghana where library services are underdeveloped (Filson & Agyekum, 2014).

1.3. Objectives and Research Questions

The purpose of the study was to evaluate the information access skills of students of the NJACoE, a public teacher tertiary education provider in the Upper East Region of Ghana. The objectives of the study were to:

1. Explore the main sources of scholarly information of the students of the college.
2. Determine how the students accessed scholarly information from the identified sources.

The study was anchored on two important questions:

1. What constitutes the main sources of information for students of NJACoE?
2. To what extent are the information access skills of students of NJACoE adequate to help them conduct independent research?

2. LITERATURE REVIEW

This section explores the literature on the subject in particular respect to the following thematic areas: sources of information; traditional library sources; and Internet sources.

2.1. Sources of Information

In recent times, access to information for academic and other purposes is possible through different sources. In the view of Rieh and Hilligoss (2008), information users consult online search platforms, resort to expert knowledge sources and acquaintances, use libraries, and patronise traditional news sources such as newspapers and television for their information needs. Omeluzor et al. (2017) also suggest that classroom teaching is one of the methods through which students access information. Overall, students possess more knowledge in source identification, needs description, and awareness of potential sources (Abdul Karim et al., 2013). This may be attributed to most universities initiating information literacy programmes for the developing of these skills. In Nigeria, students often access information through the use of information and communication tools such as computers, mobile phones, and Internet (Adetimirin, 2012). A study by Taylor and Dalal (2017) on sex and information literacy identified two sources of information: traditional library sources and Internet sources. However, they indicated that although students used both information sources, they commonly choose the Internet over traditional library sources. The use of Internet sources is determined by relative access, and availability of facilities and services (Almarabeh, Majdalawi, & Mohammad, 2016).

Also, to broaden students' information access as well as to assist students to acquire the skills to differentiate good information from bad, it will be advantageous for institutions to turn their attention to promoting information literacy. This is exactly the position of Yeboah et al. (2017), who suggest the need for integration of information literacy into the curricula of Ghanaian secondary schools, and for librarians to collaborate with relevant stakeholders such as teachers and policy-makers. Similarly, most of the respondents in Shrestha's (2008) study suggest the incorporation of information literacy courses at different stages of higher education.

It has been suggested that the non-integration of information

literacy in academic programmes has contributed to poor information literacy among students (Yeboah et al., 2017). In other words, information literacy retains a central role in the development of critical information skills among students in the current information society (Sasikala & Dhanraju, 2011). Shrestha's (2008) study suggested that information literate people are lifelong learners. According to Taylor and Dalal (2017), information literacy skills involve choosing the right sources of information through the application of relevant quality indicators. A study by Long and Shrikhande (2007) on information literacy and information seeking behaviour among business major students revealed that the set of teaching methods adopted for information literacy delivery culminated in considerable upgrade of information-seeking behaviour among undergraduate students. Additionally, Sasikala and Dhanraju (2011) vouched for the utilitarian role of information literacy as an indispensable skill. Okon et al. (2014) contend further that in the educational sector, this skill is useful for, among other goals, research completion and presentations in the university. Furthermore, students can evaluate information sources by checking for details such as the authority and qualifications of an author of a print book (Yeboah et al., 2017). Julien and Barker (2009) investigated high school students' finding and evaluation information skills. They found that despite unambiguous curricular mandates to develop the information literacy skills of students, their actual skill levels were underdeveloped. Adetimirin (2012) found poor information literacy skills among university students in Nigeria in his investigation of the ICT literacy of university students in Nigeria. To improve these poor outcomes, Yeboah et al. (2017) suggest the establishment of ICT centres in schools to enable students to access quality information to facilitate the development of information literacy (IL) skills among students (Long & Shrikhande, 2007).

2.2. Traditional Library Sources

Libraries have remained an important source of academic information (Yeboah et al., 2017). Academic libraries have the fundamental objective of supporting academic institutions by deploying a highly-effective knowledge-building model to obtain relevant information sources for the educational ideals of universities (Omeluzor et al., 2017). Similarly, school libraries enhance access to a variety of information modes and the concomitant expertise for children to exploit knowledge in all its shapes and forms (Sasikala & Dhanraju, 2011). Hence, Yeboah et al. (2017) are of the view that a school's timetable should have library periods to expose students to the use of library services and information resources to embark on independent learning (Yeboah et al., 2017).

However, the phenomenon of underuse of school libraries has been noted widely in the empirical literature. For instance, Chang et al. (2012) note that secondary school students in Singapore tend not to use the traditional library. Taylor and Dalal (2017) assessed sex and information literacy among college students and note that, overall, students commonly choose other sources of information over traditional library sources. It is argued that library instruction enhances students' ability and confidence to navigate the information search process (Omeluzor et al., 2017). Similarly, Spenser (2003) notes that library instruction helps students to build on their problem-solving skills by connecting their information search skills to their creative potential. Omeluzor et al. (2017) conclude by suggesting that quality instruction is an effective awareness-creation mechanism on the availability of information sources and their potential academic utility.

While students rely heavily on the Internet for information and communication, they also co-use subject experts outside libraries, print sources, and, to some extent, digital library resources (Rieh & Hilligoss, 2008). In this digital age, however, students sparsely depend on library resources, especially scholarly resources, for course-related research; and even fewer, in comparison, use services that require interfacing with human-mediated systems (Head & Eisenberg, 2009). Therefore, libraries are now adapting new approaches and methods in order to enhance information access among students. For instance, Adetimirin (2012) intimated that due to underfunding of university libraries, they were not self-sufficient in the provision of information to their users. Consequently, there was a need to leverage ICTs to complement available library resources to live up to the demands of users. ICTs also enhance the extent to which users can creatively exploit different information (re)sources to achieve standard goals (Omeluzor et al., 2017). It is, therefore, imperative for librarians to adopt viable innovative processes and programmes such as "Ask-a-librarian," blogs, library websites, social media, text messaging, emailing, and radio broadcasting to facilitate information access in libraries. Interestingly, Shrestha's (2008) study of the crucial role of information literacy in Nepal shows that a majority of respondents have used these electronic resources of the library.

2.3. Internet Sources

The technological landscape is evolving rapidly with products, programmes, and applications; these are impacting how individuals gather information, connect with peers, and interact in the classroom (BrckaLorenz, Haeger, Nailos, & Rabourn, 2013). Access to ICT facilities can greatly enhance students' access to information and ultimately result in improved IL

competencies of students (Yeboah et al., 2017). ICTs encompass the Internet, computing, and electronic communications such as radio and television (Almarabeh et al., 2016). A study conducted in Andhra University by Sasikala and Dhanraju (2011) revealed the utilitarian value of ICT skills when using electronic information sources. In respect to the use of the Internet to access information, several studies support Taylor and Dalal's (2017) claim that among the traditional library and the Internet, students' most preferred source of information is the Internet. Adetimirin (2012) pointed out that the use of ICTs provides immense value in the information search process of students. With regard to the use of the Internet, Sasikala and Dhanraju (2011) found that the highest proportion of their respondents were conversant with Internet use. As indicated, almost all students use online search platforms such as Google and Wikipedia for everyday life, especially in their course reading (Head & Eisenberg, 2009). Yeboah et al. (2017) then indicated that the Internet is both an information and communication tool that can be helpful for obtaining academic information and exchanging ideas and information with peers and teachers. Additionally, Metzger, Flanagin, & Zwarun (2003) revealed that college students rely heavily on the Internet for both general and academic information. For instance, a case study in the University of Jordan by Almarabeh et al. (2016) on Internet usage, challenges, and attitudes established that, aside from academic and research studies, communication and entertainment as well as downloading of software were also noted as valuable uses of the Internet.

Among various reasons, Sasikala and Dhanraju (2011) indicated that students are using the Internet for text communication purposes and also to access information on jobs. Students overwhelmingly felt that technology has helped them to understand course materials and to demonstrate understanding and hence they were found to frequently use technology to study on their own (BrckaLorenz et al., 2013). According to Julien and Barker (2009), students prefer the Internet because it is perceived as convenient and familiar. They are also found to be familiar with online search platforms such as Google, email, and Wikipedia (Shrestha, 2008). Finally, in the view of Taylor and Dalal (2017), despite the mixed appreciation of the quality of the Internet as a credible information source for research, its use will continue and will likely increase.

3. METHODOLOGY

This section presents a description of the general procedures adopted in carrying out the study. The main issues covered

include the research design, population and sampling, data collection, and the methods of data analysis.

3.1. Research Design

A cross sectional survey was conducted to gather data for this study. The study applied a descriptive survey method with the aim of providing a good picture of students' access to information.

3.2. Population and Sampling

The population of the study consists of the current student enrolment in the NJACoE. The student population in the 2017/18 academic year is 1,264. This consists of 356 students in Level 100, 453 students in Level 200, and 455 students in Level 300.

The sample size for the study was estimated using a statistical procedure that was proposed by Yamane (1967). The formula is given as:

$$n = \frac{N}{1 + N(e)^2}$$

Where n=sample size; N=sample frame; and e=error or significance level. According to Ahuja (2011), an acceptable error level traditionally is up to ± 0.05 or ± 0.10 (i.e., 5 or 10 percentage point). In this study, N=1264, e=5%. Hence the estimated sample size for the study is

$$n = \frac{1264}{1 + 1264(0.05)^2} = 303$$

The sampling distribution according to stage/levels is shown in Table 1.

From Table 1, a total of 303 respondents were selected, consisting of 85 from Level 100, 109 from Level 200, and 109 from Level 300. A stratified sampling procedure was used in the selection of the respondents. The purpose of the stratification was to give a chance to students from different year groups to be included in the sample.

Table 1. Sampling distribution of respondents

Respondent category	Population	Proportion selected (%)	Sample size
Level 100	356	28.2	85
Level 200	453	35.8	109
Level 300	455	36.0	109
Total	1,264	100	303

3.3. Data Collection and Analysis

Primary data were collected from respondents using a self-administered questionnaire. The questionnaire consisted of three sections: Section one dealt with background information; Section two was on types of information; and Section three concerned sources and access to information. The respondents were identified and copies of the questionnaire were given to them and were retrieved, resulting in a 100% response rate.

The data were cleaned, coded, and entered into the SPSS spreadsheet for further transformation. The analysis was done using descriptive statistics. The frequencies of key variables were generated and their relative percentages estimated. The data were presented in tables to facilitate in discussing the findings.

4. RESULTS AND DISCUSSION

This section presents the results and discussion of the analyses. This includes background information of the respondents, the types of information they access, sources of information, and the relative access to information by the respondents.

4.1. Background Information of Respondents

It was discovered that 75.9% of the respondents were male students while the remaining 24.1% were female students. This means that male students dominate over female students in the college of education. Besides this, the respondents were found to be of a minimum of 17 years and a maximum age of 36 years old. Further statistics in Table 2 shows a mean age of 22.44 (years) with a standard deviation of 2.8. In terms of the programme of study, the results indicate that the students pursue three main programmes. From Table 2, 7.9% of the respondents

Table 2. Background information

Variable	Frequency	Percent			
Sex					
Male	230	75.9			
Female	73	24.1			
Total	303	100			
Programme of study					
Technical and mathematics	24	7.9			
Science and mathematics	82	27.1			
Social studies	197	65.0			
Total	303	100			
Variable	N	Minimum	Maximum	Mean	Standard deviation
Age	303	17	36	22.44	2.814

Table 3. Types and forms of information

Information	Frequency	Percent
Form of academic information		
Online/digital	89	29.4
Printed information	214	70.6
Total	303	100
Form of social information		
Online/digital	214	70.6
Printed information	89	29.4
Total	303	100

pursue technical with mathematics, 27.1% pursue science with mathematics, while the remaining 65.0% study social studies.

4.2. Types of Information

The results indicate that students often use academic and social information. Academic information represents all information needs for academic exercises while social information is information needs for social purposes other than academic. The types of information needs are often accessed through online/digital form or in a printed form. The results in Table 3 revealed that 29.4% of the respondents often use online or digital information for academic purposes. Besides this, 70.6% of the respondents often use printed forms of information for academics. This means that the students prefer (are more conversant with) printed forms of information rather than online/digital forms of information. In terms of information needs for social purposes, the results indicate that 70.6% often use online/digital information while only a small proportion (29.4%) use printed information for their social needs.

The results also imply that when it comes to academic information the students rely greatly on printed forms but use more online information for their social needs. This suggests that the use of online information sources for academic purposes among the students is low. This result is corroborated by previous researchers (Owusu-Ansah, Rodrigues, & Van Der Walt, 2018) who found overwhelming preference for print books among teachers studying through distance learning.

4.2.1. Sources and access to information

The respondents were asked to indicate their specific sources of information, and multiple responses were provided as shown in Table 4. From the findings, the main sources of information include online search platforms (including Web search engines) (49.2%), authorities (such as authors and tutors) (23.4%), friends (45.9%), the college library (35.3%), newspapers (22.1%), and television (54.8%).

Table 4. Sources and access to information

Variable	Frequency	Percent
Sources of information (to students)		
Web search engines	149	49.2
Authorities	71	23.4
Friends	139	45.9
Libraries	107	35.3
Newspapers	67	22.1
Television	166	54.8
Total	303	100
Medium used to access information		
Computers	14	4.6
Mobile phones	235	77.6
Direct contact with source	54	17.8
Total	303	100
Frequency of using online search platforms		
Twice or more in a day	76	25.1
Once in a day	22	7.3
Once in a week	37	12.2
Once in month	130	42.9
Occasional	38	12.5
Total	303	100

Table 4 revealed that several devices and or methods of access were being used by students to access information. The findings show that 4.6% of the students rely on the use of computers to access information, and 77.6% rely on the use of their mobile phones, while 17.8% rely on direct contact with non-digital sources. Some of these non-digital sources include print formats such as books, journals, newspapers; and personal sources.

Results on the frequency of information search using online search platforms were obtained in multiples responses. The results indicate that 25.1% of the respondents use online search platforms to search for information twice or more in a day, 7.3% use them once in a day, 12.2% use them once in a week, 42.9% use them once in a month, and 12.5% use them occasionally. This suggests that information use (with reference to online search platforms) among the students was not regular.

Consistent with past studies such as those by Rieh and Hilligoss (2008) and (as well as) Omeluzor et al. (2017), this study confirmed that students rely on various sources for information such as the use of online search platforms, and consulting authorities together with the use of library resources. The use of different sources of information will widen their information access while exposing them to different perspectives in relation (argument) to a particular subject. This is an improvement in information access over traditional

methods because of the relative convenience that characterizes the use of technology.

It is also important to point out that the participants indicated television as a major source of information, indicated by more than half of the participants (54.8%). The question to ask is whether television as a source of information is used to obtain information related to the students' academic work. As noted by Philo (2014), television remains the main source of information for most people on contemporary events. Despite this, it is interesting that students at a university would consider television as more important a source than libraries, as there remain questions of credibility of information in even the traditional mass media (Philo 2014). Perhaps students in this institution are not satisfied with the information provision in their library, or they lack adequate information search and access skills to take advantage of the enormous volume of information online. Regardless of these postulations, it is necessary that the role of television in information provision to students be explored in the future.

4.3. Use of Traditional Library Sources

Traditional libraries have been sources of relevant academic information to students in different parts of the world. The respondents in this study were, therefore, asked to indicate whether they use the college library. The results indicate that 62.0% had been using the college library to access relevant academic information while 38.0% maintained that they do not use the library resources in the college. When asked to indicate the use of electronic resources in the library, the results revealed

that only 20.1% of the respondents used the electronic resources of the library, a result which is very indicative of low patronage of electronic resources in the library.

Further evidence reported how the students learned to use the library resources. The multiple responses in Table 5 revealed that 25.1% learned to use the library through the assistance of library staff, while 30.7% through their friends who were conversant with the system, and 9.2% often used the library on their own but not without frustrations. However, most (35.0%) maintained that they were conversant with the library and hence were able to use the library without any assistance. This finding is consistent with previous findings of electronic resources usage in Ghana.

The results also imply that the college library has been very useful to students in educating students to access relevant academic information. This is in line with Spenser (2003) who concluded that library instruction enables users to identify and use information to address issues at hand. However, students encounter challenges in their efforts to use library resources. For example, 79.9% of the respondents were not using electronic resources and hence could not be described as educated on how to use electronic resources from the library. On the other hand, a considerable number of users could not access needed information without assistance (55.8%), while 9.2% who reported their ability to use library resources independently complained of frustration. These findings suggest little or no library user education among the students, and as previous research suggests (Adeleke & Emeahara, 2016; Adeniyi Aderibigbe & Adebimpe Ajiboye, 2013), low user education militates against students' ability to use library-based electronic resources.

A question was also posed to students on the use of library catalogues for information access in the college. Library catalogues that are still in use in higher educational institutions in Ghana include bibliographic catalogues, card catalogues, and online public access catalogues. This question was intended to find the perception of students on the usefulness of library catalogues as useful information access tools. From Table 6, multiple responses from the students were that 25.7% consider the bibliographic catalogue as convenient in information access, 16.2% consider the card catalogue an easy way of accessing information, and 33.3% prefer the online catalogue for information access. It was also discovered that 38.6% had no idea regarding the use of library catalogues. With the considerable number of students (38.6%) who were unaware of any type of library catalogue in the library, it could be pointed out that either the library had no library catalogue or that they had but it was not publicised among the students. As noted

Table 5. Use of traditional library resources

Variable	Frequency	Percent
Use of college library		
Yes	188	62.0
No	115	38.0
Total	303	100
Use of library electronic resources		
Yes	61	20.1
No	242	79.9
Total	303	100
Learning to use library resources		
Through the assistance of library staff	76	25.1
Through friends who know how to use the library	93	30.7
I usually find what I want but with frustration	28	9.2
I know how to search for library materials	106	35.0
Total	303	100

by Yang and Hofmann (2010), the library catalogue could be a gateway to all the resources of the library and as such its availability in a library, awareness among students, and readiness to use it can contribute significantly to increasing the use of library resources, including physical and electronic publications owned or accessible through a library.

When asked whether user orientation was essential in information access, especially in library use, 93.7% of the respondents confirmed in the affirmative that they considered user orientation to be essential in their ability to use library resources. Further evidence suggests that 41.3% of the respondents considered user education on library activities to be valuable. Also, 48.2% of the respondents maintained that there should be a guide to the location of relevant academic information in the library; 38.3% indicated that there should be user manuals for the library; and 25.1% indicated that there should be provision of a guide to facilitate the use of the library catalogue. The finding that most students considered user orientation essential to their ability to use library resources was re-echoed in previous studies. Murphy and Black (2013, as cited in Daniel, 2016) indicated in their study that most students found it difficult to access the library guides in their school's learning management system, although among those who found it and used it, they appreciated the fact that the library guide helps them to get started when they were to complete their assignments, as well as enhancing their ability to find important research materials to complete their assignments. There was therefore a call for promotional efforts to create awareness of the potency of library guides in increasing library use.

Table 6. Use of library catalogues to information access

Variable	Frequency	Percent
Bibliographic catalogues	78	25.7
Card catalogues	49	16.2
Online catalogues	101	33.3
Have no idea	117	38.6
User's orientation essential in information access		
Yes	284	93.7
No	19	6.3
Total	303	100
Type of user's orientation preferred		
Briefing in school activities	125	41.3
Guide to the location of resources	146	48.2
Use of manuals	116	38.3
Providing guide to catalogue	76	25.1

The results further suggest that students have admitted their unfamiliarity with the library resources and hence suggest the need for user orientation. This also means that despite their unfamiliarity with the library resources, they understand that it still represents a useful source of academic information. The requirement for college of education students in Ghana to conduct their own research, as indicated by Nyarkoh (2016), had not been met in this study because of students' relative challenge in information access.

4.4. Use of Internet Sources

Students now consider Internet technology a useful mechanism for accessing academic and social information. They achieve this through the use of different online search platforms such as Google, Yahoo, MSN, Google Scholar, Ask, and Bing. The respective frequencies and proportions of students using these online search platforms are shown in Table 7. The findings show that a majority (72.9%) of them use Google to access their needed information. Also, another (widely) fairly used search engine is Yahoo, which is used by 15.8% of the respondents.

The respondents were again asked to rate the effectiveness of the online search platforms used in accessing information. The responses in frequencies and percentages are shown in Table 8. From the table, 43.6% maintain that Google is very effective in information access, and 32.3% indicated that it is effective, while 22.8% said it is somewhat effective. Only 1.3% maintained that Google is not an effective search engine. Beside this, the students also gave a good rating to Yahoo with 21.8%, with 40.3% indicating that it is very effective and effective, respectively. However, the other online search platforms such as MSN and Google Scholar, as shown in Table 8, were not given good ratings by the respondents in terms of their effectiveness in accessing needed information.

The results imply that the numerous several online search platforms available have not all been widely used. The main reason is that their relative effectiveness in searching for needed

Table 7. Use of different online search platforms

Online search platform	Frequency	Percent
Google	221	72.9
Yahoo	48	15.8
MSN	9	3.0
Google Scholar	8	2.6
Ask	16	5.3
Bing	1	0.3
Total	303	100

Table 8. Effectiveness of online search platforms in retrieving information

Search engine	Very effective	Effective	Somewhat effective	Not effective
Google	132 (43.6)	98 (32.3)	69 (22.8)	4 (1.3)
Google Scholar	34 (11.2)	76 (25.1)	163 (53.8)	30 (9.9)
Yahoo	66 (21.8)	122 (40.3)	87 (28.7)	28 (9.2)
MSN	41 (13.5)	43 (14.2)	168 (55.4)	51 (16.8)
Ask	40 (13.2)	64 (21.1)	56 (18.5)	143 (47.2)
Bing	32 (10.6)	67 (22.1)	37 (12.2)	167 (55.1)

Values are presented as number (%).

information varies considerably. Google and Yahoo are more often used because of their relatively higher effectiveness, ease-of-use, and user-friendliness (Kwadzo, 2015). However, it is surprising that Google Scholar was not considered a major important source of scholarly information by the participants. This may be a result of their lack of awareness of the pre-eminence of Google Scholar to search across different platforms to provide openly accessible information, which may be critical to institutions where there is limited or no access to subscription/paywalled models for information access.

5. CONCLUSION AND RECOMMENDATIONS

Access to information is very critical in the academic performance of students. As a result, students often rely on using both purely academic literature as well as social media information to supplement their academic information. The results of this study imply that students use a variety of sources to access information. These sources span traditional (library) and Internet, formal (authoritative), and informal sources of information. However, the uses of these vary in intensity and type of information. The use of traditional library sources is more effective than all the other sources of information, but on the other hand their use is limited because of unfamiliarity with library facilities. Besides this, the findings of this study suggest that the inability of students to access information from various sources means that their information literacy skills are low. The students largely cannot access relevant electronic information thus giving them a limitation on information access. This implies further that they will have challenges in conducting empirical research in line with their curriculum requirements.

On the basis of the results, it is recommended that the information literacy skills of students of the NJACoE in Ghana be improved. This is necessary because they are relatively weak in their ability to access scholarly information. To achieve

this goal, the management of the college should consider implementing and integrating information literacy in their institution's academic curriculum to enhance the information skills of their students to fulfil the empirical research components of pre-service teacher education. Information literacy as a course will afford students of the college to acquire four basic skills: Identify and conceptualise their information needs; locate credible sources of information; improve their search skills; and enhance their evaluation skills. Finally, in order to enhance students' access to quality and credible information resources, the college should seek membership in the Consortium of Academic and Research Libraries in Ghana (CARLIGH) to have access to electronic databases and journals subscribed to by the consortium. This would also expand the institution's access to critical information required for academic research by both students and faculty. Finally, to improve the students' information access skills, adequate provision of information in the traditional format must be made, while enhancing their online searching skills.

REFERENCES

- Abdul Karim, A., Shah, P. M., Din, R., Ahmad, M., & Khalid, F. (2013). Measuring information skills among Malaysian Youth Students: An instrument development. *Asian Social Science*, 9(16), 22-31.
- Adeleke, D. S., & Emeahara, E. N. (2016). Relationship between information literacy and use of electronic information resources by postgraduate students of the University of Ibadan. *Library Philosophy & Practice*, Paper 1381.
- Adeniyi Aderibigbe, N., & Adebimpe Ajiboye, B. (2013). User education program as determinant of electronic information resources usage in Nimbe Adedipe University Library, Nigeria. *The Electronic Library*, 31(2), 244-255.
- Adetimirin, A. E. (2012). ICT literacy among undergraduates in Nigerian universities. *Education and Information Technologies*, 17(4), 381-397.
- Ahuja, R. (2011). *Research methods*. Jaipur: Rawat Publications.
- Almarabeh, T., Majdalawi, Y. Kh., & Mohammad, H. (2016). Internet usage, challenges, and attitudes among university students: Case study of the University of Jordan. *Journal of Software Engineering and Applications*, 9(12), 577-587.
- BrckaLorenz, A., Haeger, H., Nailos, J., & Rabourn, K. (2013, May 18-22). *Student perspectives on the importance and use of technology in learning*. Paper presented at the 2013 Association for Institutional Research Annual Forum in Long Beach, CA.

- Chang, Y., Zhang, X., Mokhtar, I A., Foo, S., Majid, S., Luyt, B., & Theng, Y. (2012). Assessing students' information literacy skills in two secondary schools in Singapore. *Journal of Information Literacy*, 6(2), 19-34.
- Daniel, D. (2016). Embedded library guides in learning management systems help students get started on research assignments. *Evidence Based Library and Information Practice*, 11(1), 76-78.
- Filson, C. K., & Agyekum, B. O. (2014). Colleges of education libraries in Ghana: An evaluative study. *Information and Knowledge Management*, 4(1), 7-19.
- Head, A. J., & Eisenberg, M. B. (2009). Lessons learned: How college students seek information in the digital age. *SSRN Electronic Journal*. doi:10.2139/ssrn.2281478.
- Julien, H., & Barker, S. (2009). How high-school students find and evaluate scientific information: A basis for information literacy skills development. *Library & Information Science Research*, 31(1), 12-17.
- Kwadzo, G. (2015). Awareness and usage of electronic databases by geography and resource development information studies graduate students in the University of Ghana. *Library Philosophy & Practice*, Paper 1210.
- Long, C., & Shrikhande, M. (2007). Information literacy and information seeking behaviour among business majors. *University Library Faculty Publications*, Paper 16.
- Metzger, M. J., Flanagin, A. J., & Zwarun L. (2003). College student web use, perceptions of information credibility, and verification behavior. *Computers & Education*, 41(3), 271-290.
- Nusrat Jahan Ahmadiyya College of Education. (2018). *Basic Statistics*. Wa: NJACoE.
- Nyarkoh, E. (2016). *The degree of autonomy in colleges of education in Ghana. A comparative study of before and after their upgrade to tertiary status* (Unpublished master's thesis). University of Oslo, Oslo, Norway.
- Okon, M. E., Etuk, E. P., & Akpan, U. J. (2014). Information literacy skills and information use by students in two south university libraries in Nigeria. *International Journal of Economics, Commerce and Management*, 2(9), 1-16.
- Omeluzor, S. U., Akibu, A. A., Dika S. I., & Ukangwa, C. C. (2017). Methods, effect and challenges of library instruction in academic libraries. *Library Philosophy and Practice*, Paper 1465.
- Owusu-Ansah, C. M., Rodrigues, A., & Van Der Walt, T. (2018). Factors influencing the use of digital libraries in distance education in Ghana. *Libri*, 68(2), 125-135.
- Philo, G. (2014). *Seeing and believing: The influence of television*. Routledge.
- Rieh, S. Y., & Hilligoss, B. (2008). College students' credibility judgments in the information-seeking process. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital media, youth, and credibility (The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning)* (pp. 49-72). Cambridge, MA: MIT Press.
- Sasikala, C., & Dhanraju V. (2011). Assessment of information literacy skills among science students of Andhra University. *Library Philosophy and Practice*, Paper 626.
- Shrestha, S. (2008). *Study on need and importance of information literacy in Nepal: Special emphasis to students of MLISc, TU, Central Department of Library and Information Science and professionals of Kathmandu Valley* (Doctoral dissertation). Tribhuvan University, Kathmandu, Nepal.
- Spenser, T. (2003). *Information literacy meeting of expert* (Conference report). Retrieved from UNESCO website: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/themes/info_litmeeting_prague_2003.pdf.
- Taylor, A., & Dalal, H. A. (2017). Gender and information literacy: Evaluation of gender differences in a student survey of information sources. *College & Research Libraries*, 78(1), 90-113.
- Yamane, T. (1967). *Elementary sampling theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Yang, S. Q., & Hofmann, M. A. (2010). The next generation library catalog: A comparative study of the OPACs of Koha, Evergreen, and Voyager. *Information Technology and Libraries*, 29(3), 141-150.
- Yeboah, P., Dadzie, P. S., & Owusu-Ansah, C. M. (2017). Information access and evaluation skills of secondary school students in Ghana. *Library Philosophy and Practice*, Paper 1552.

Key Traffic Metrics as a Basis to Measure Library Performance

Olga M. Udartseva*

Scientific and Technological Department, State Public
Scientific-Technological Library of the Siberian Branch of
the Russian Academy of Sciences, Novosibirsk, Russia
E-mail: udartseva@spsl.nsc.ru

ABSTRACT

Webometric research in the Russian library sector is just beginning to gain momentum. This article examines the experience of webometric research in libraries from the perspective of the global practice. In particular, it highlights a number of foreign works, which may have a special practical value for Russian libraries, and emphasizes important webometrics areas for libraries. The purpose of this study is to research the practical application of key performance indicators (KPIs) abroad and conduct a webometric analysis of the websites of some leading Siberian and Far Eastern scientific libraries based on selected KPIs. The study data were collected with SimilarWeb and other analytical tools. The study revealed that key traffic metrics are the basis of webometric research, and identified available promising groundwork for the purpose of their further testing. The shortcomings in the current state of the websites of the Siberian and Far Eastern scientific libraries were noted. Based on the obtained webometric traffic indicators, the ranking of the Siberian and Far Eastern scientific libraries was made.

Keywords: webometry, webometrics research in libraries, web analytics tools, key performance indicators, Siberia and the Far East, traffic sources

Open Access

Received date: March 11, 2018
Accepted date: February 10, 2019

*Corresponding Author: Olga M. Udartseva
Junior Researcher
Scientific and Technological Department, State Public Scientific-Technological
Library of the Siberian Branch of the Russian Academy of Sciences, Voskhod
St., 15, 630102, Novosibirsk, Russia
E-mail: udartseva@spsl.nsc.ru

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

1. INTRODUCTION

Today, we can observe libraries' migration into digital space following potential users. The rapid growth of network information resources around the world, including library resources (websites, social network accounts, blogs, etc.), necessitated their organization and evaluation. On this ground, there arises webometrics and the need for webometric research. To date, webometrics has received several definitions. First, the use of the term "webometrics" in relation to quantitative indicators has been defined—"web statistics." Another webometrics definition refers to it as "a field of research using the methods of bibliometrics of network information resources, structures, and technologies" (Björneborn & Ingwersen, 2001).

Libraries today have to compete with the alternative of instant access to information on the web. That is why libraries are now doing a great job of creating, analysing, and managing web resources (Udartseva, 2018). Libraries are aware of the webometric study potential to provide useful information about user behaviours on their web resources and have several aspects for use: service quality improvement, self-presentation of library products and services in virtual space, and use of empirical data for the implementation of various solutions at the level of web resource management (Lowe, 2012; Paul & Erdelez, 2009). Conducting webometric research, libraries, one way or another, need to tackle a number of pressing issues, the answers to which are a key factor in ensuring library competitiveness in the information services market, being, in particular, the basis for their effective activities in the online space:

- What tools to choose when analysing library web resources?
- What metrics to consider key ones for library web resources?
- What is the trend of web resource traffic?
- Where do users come from (traffic sources)?
- How much time do users spend on the library website on average?
- When do library web resources get the most users?
- How many pages do users view before leaving the website?
- What website pages are the most popular?
- What pages of the library website are considered as target pages?
- What is the target audience of library web resources?
- Who links to library web resources?

Especially today, libraries need theoretical justification of existing issues and their practical resolution. Due to constant transformation of the web space, libraries look for new models and mechanisms of website management. Webometric research will help to better understand the user audience and evaluate library

performance in web space. The world libraries of different levels and types are engaged in the development of the issues raised.

The aim of this research was to:

- study the preferences of scientific libraries in Siberia and the Far East in choosing web-based analytical tools,
- identify key metrics for evaluating library website effectiveness,
- identify the most productive traffic sources for the analysed libraries, and present related findings on audience segmentation depending on the traffic nature, and
- compare the results in order to build a rating of scientific libraries in Siberia and the Far East.

In addition, some metrics correlation was found; this feature was taken into account in this study, which allowed us to give an optimal interpretation of current indicators.

2. LITERATURE REVIEW

2.1. Key Traffic Metrics

Using web analytics as a globally recognized tool, commercial and non-profit organizations seek to improve their current performance indicators (Jacoby & Luqi, 2007; Mihajlova, 2017; Redkina, 2017; Sen, Dacin, & Pattichis, 2006; Srinivasan, Amir, Deshpande, & Zbarsky, 2004). Thus, to be competitive in web space, libraries need not only to find the necessary set of web-based analytical tools, but to set up their own optimal development strategy as well. The strategy is based on key metrics from web analytics tools that allow measurement of the achievement of libraries' strategic goals. For each library, the main question arises: What key metrics will become important in the information market in the near future? This issue concerns large academic libraries (Kann, 2017; Udartseva & Rykhtorova, 2018), regional libraries (Chepushtanova, 2016), university libraries (Chakravarty & Wasan, 2015; Mihajlova, 2017; Tian, 2012), and other libraries that use web-based analytical tools to effectively evaluate activities in web space. In the existing studies, the authors propose different criteria to classify key indicators (Chepushtanova, 2016; Kann, 2017; Udartseva & Rykhtorova, 2018). The key traffic metrics for library web resources, based on the goals and objectives, are considered as follows (Udartseva & Rykhtorova, 2018): visits, visitors, new visitors, traffic sources, average visit time (time on website), visit depth (average indicator), and bounce rate. In many foreign studies related to evaluation of library websites, the emphasis is also placed on these traffic indicators (Fagan, 2014; Jansen, 2009; Vecchione, Brown, Allen, & Baschnagel, 2016).

2.2. Potential of Key Performance Indicators to Measure Library Website Performance

In 2013, the American Library Association's topic "Developing a Web Analytics Strategy for Your Library" was covered by a seminar, which discussed issues of setting goals and choosing suitable key performance indicators (KPIs). Most KPIs in web analytics tools are focused on commercial websites; therefore, libraries need to set separate indicators to assess how users receive information that meets their needs and objectives. This conclusion was made by a group of American scientists of the Boise State University Albertsons Library (USA) (Vecchione et al., 2016). The authors propose to use the bounce rate as a key metric. Web analytics tools typically measure this indicator automatically, but "for libraries, it is a complex metric that needs to be understood in context because library websites also serve as information portals to other websites" (Vecchione et al., 2016, p. 3). The library web team created a control panel in Google Analytics, which included specially selected domains with the changed tracking code. Event tracking is a method of measuring any activity that occurs on a website. The authors describe in detail the technical process of creating a series of labels consistent with each event. Depending on the bounce rate, the web team identified the most popular products and services on the library website. The data obtained have formed the basis for revising the popular content placement on the website main page, which significantly reduced the user path in finding the necessary information, and as a result reduced the bounce rate from 92.8% to 67%. Thus, by promoting popular content and providing users with what they need, the bounce rate can be significantly reduced, and instead the time users spend on the website can be increased (Vecchione et al., 2016).

Association for Research Libraries (ARL) and National Information Standards Organization (NISO) distinguish only one "virtual visit" webometrics (Fagan, 2014). In 2009, Jansen (2009) grouped webometric indicators into four categories, depending on the affiliation to a particular website: commerce, content/media, support/self-service, and lead generation. Fagan (2014) believes that libraries today combine the first three indicator categories, and in detail specifies the value of each indicator to evaluate the library web resources performance. Among the web metrics in the "commerce" category suitable for library web resources, Fagan (2014) highlights the following:

- average order value (comparing the order value for specific subscription resources will allow estimation of its demand),
- average visit cost (total number of visits will allow estimating efficiency/inefficiency of website pages: efficiency/inefficiency of the advertising campaign, the need to redesign pages, and changing of authentication

- mechanisms),
- bounce rate (calculating the percentage of visitors who leave after viewing one page will determine the landing page performance for different traffic sources),
- conversion rate (calculating the percentage of visitors who have completed certain actions will assess the success of the current advertising campaign), and
- loyalty (assessment of overall success, determined by the visit frequency of new and existing users).

In the "content" category for library websites, Fagan (2014) has highlighted the following:

- visit depth (page views/visits will allow evaluation of interaction with the web resource),
- repeat visitors (unique visitors/total visits number will allow assessing the quality of the content of interest to target users),
- new users (new/unique visitors will allow evaluating the website or website pages performance), and
- popular pages (page views/unique visitors show the popularity of the page and may indicate a correspondence between the website goals and visitor interests).

Web metrics for library websites in the "self-service" category include:

- visit depth (low numbers of page views/visits indicate that users quickly find the information they need),
- popular pages (high numbers of page views/unique visitors indicate an area where many people have problems [support issues, not from the website]),
- bounce rate (low percentages of visitors who leave after viewing one page, usually indicating information performance, with some exceptions),
- average time per page (comparison of total time spent on a page/visits with past time periods or similar page types),
- user satisfaction (comparison with previously obtained indicators to check information usability), and
- keywords/phrases for internal search (identifying the most common issues, both with support issues and navigation) (Fagan, 2014).

In the metrics choice for libraries, Fagan proposes to be guided by two questions in relation to each metric: "What can be done to improve them if necessary?" and "Who can take action?" If these questions cannot be answered, the metric usefulness is questionable (Fagan, 2014, p. 31). By choosing indicators that can actually be useful to make a success of a common strategy, libraries can quickly and efficiently manage

their own web resources in the virtual space.

Most of the metrics offered by different library teams and individual scientists as key metrics relate to website traffic analysis in one or another way. This pattern leads to the conclusion that traffic metrics are the basis for library website analysis, so they deserve special attention from the analyst.

2.3. Evaluation of the Library Website Performance Using Tools for Competitive Intelligence

The literature review also showed that competitive intelligence by webometric indicators (tools selection for competitors assessing; collection, processing, and analysis of statistical indicators; comparison/ranking of web resources against each other) remains an insufficiently studied area in the Russian library segment. Most researches aim at studying web resources of libraries where they work (Chepushtanova, 2016; Kann, 2017; Udartseva & Rykhtorova, 2018) due to the closed access to the statistics of other library web resources. However, there is a large number of web-based analytical tools that allow conducting competitive intelligence (SimilarWeb, Alexa, Pr-Cy.ru, etc.).

Their choice depends on the goals and objectives set for the analyst: comprehensive web resource evaluation, optimization assessment, content and popularity evaluation, etc. Authors use different methods and tools in their research (Dastani, Atarodi, & Panahi, 2018; Jowkar & Didegah, 2010; Kann, 2017; Udartseva & Rykhtorova, 2018). Udartseva and Rykhtorova (2018) studied the available analytical functionality and methods of collecting statistics by such tools and identified possible goals of website analysis, noting negative and positive aspects of assessing by such tools. Comparison of SimilarWeb and Alexa tools in the study by Prantl and Prantl (2018) showed that the use of these tools would approximate the significance of a particular website. Thus the study results have shown that the tools of this group have sufficient potential for competitive intelligence (Prantl & Prantl, 2018; Udartseva & Rykhtorova, 2018). Assessment of the Iranian Newspapers websites using Alexa data concluded that the websites require development, and their work is estimated as insufficiently effective in the web environment (Jowkar & Didegah, 2010). Kann (2017) stresses that, using SimilarWeb, it is possible to identify regular attendance of libraries, assessing the entire range of links connecting the library website with its partners and similar pages. It should be noted that for detailed analysis of website use, it is better to apply more complex web-analytical tools with advanced functionality, for example, Yandex.Metrica, Google Analytics, LiveInternet, and AWStats. The technology for collecting data from such tools is slightly different: counters (data about visits is collected using JavaScript code installed on the website pages); log analysers (data on all

server accesses are recorded in log files). However, it will be possible to conduct competitive intelligence with the help of these groups of tools only if the statistics are open for general use, or private access to temporary data is provided.

2.4. Other Library Website Assessment Methods: Webometrics Rating

Dastani et al. (2018) have evaluated the websites of 23 electronic libraries of Iranian medical universities on the basis of evaluation criteria. S is the size (number of pages on the website recorded by Google search engine); R (Rich files) is the value (number of full text files on the website); Sc (Scholar) is academic quality (number of publications and citations recorded by Google Scholar); and V is visibility (number of unique external links to the website recorded by MOZ search engine). The main study conclusion was the proof that the rise in indicators for the selected evaluation criteria increases the website success in the web environment. This method of web resources evaluation is also used to evaluate the websites of libraries and other scientific organizations (Guskov, Kosyakov, & Bykhovtsev, 2016). This analysis method is aimed primarily at quantitative measurement of the website content, and user audience and the website traffic remain virtually unexplored. An alternative approach that complements the existing methodology, in our opinion, is the study of web traffic and attendance resources, so this study will focus on the evaluation of web traffic sources and the analysis of traffic metrics.

3. METHODOLOGY

In this study the author used a quantitative approach. This study was conducted based on a sample of websites of the largest scientific libraries in Siberia and the Far East. The obligatory condition for selection was that the library had the status of a coordinating centre for the libraries of the region, territory, and republic in Siberia and the Far East and that it was working in web space through its official website. Based on these criteria, there were selected 22 library websites of the largest scientific libraries of Siberia and the Far East, which are research institutions in the fields of library science, bibliography and book history, and are organizational-methodical and coordinating centres of libraries of the region, territory, and republic. The relevance of the library website URL address was checked manually during direct visits to the website. At the time of the study, all website URL addresses of the Siberian and Far Eastern scientific libraries, presented in Table 1, were relevant. In order to assess the popularity of web-based analytical tools, which are

preferred by the Siberian and Far Eastern scientific libraries, the Pr-cy.ru analytical tool was used, as well as the analysis of websites by general browsing (de visu) for the purpose of additional evaluation. Pr-cy.ru is a free website analysis service that offers a variety of indicators and is an affordable online tool. In this study, we used it due to availability. When analysing the website, Pr-cy.ru determines the unique identifier of the counter being used by the counter code, which is located on the website pages. Monitoring was conducted in June 2018.

Using web analytics tools makes it possible to analyse the library segment in the web space and identify the main competitors in this area. One of such tools is SimilarWeb, a web service of the Israeli group of developers, which collects, analyses, and provides data on user digital behaviors, allowing creation of a prototype of user behavior and assessment of the audience involvement and loyalty on the website and mobile application. The basis of the web service functioning is the big data processing technologies characteristic of such web analytics tools. SimilarWeb is widely used for competitive intelligence and is especially important in a situation where the web-analytical data describing library websites is not accessible. The accuracy of the data obtained, according to the web service users in Russia, is 90-95% (Mamon, 2017). With the use of SimilarWeb service, data on the attendance of scientific libraries in Siberia and the Far East for a six-month period have been collected. The period (six months) was not chosen by chance. First, SimilarWeb service provides data only for the last six months, and second, this period is optimal, allowing predicting data for the other half of the year, which generally gives a good result. Monitoring was conducted 3-27 July 2018 by key traffic metrics (visits, average visit time, average visit depth, bounce rate, and traffic sources) since they are potentially important for explaining the effective development of libraries on the web. Noting the importance and crucial significance of traffic metrics to assess the efficiency/inefficiency of the library website, we will take them as a basis for our research.

Data from reports of Pr-cy.ru and SimilarWeb services have been entered into an Excel spreadsheet, sorted and constructed to define the regularities of using web analytical tools and current development of scientific libraries of Siberia and the Far East. Webometric indicators without analysis present formal statistics; of particular value is the interpretation of data collected using analytical systems. Interpretation of webometric indicators was based primarily on:

1. definition of terms:

- “visits” means the user interaction with the website for a certain period of time;
- “average visit time” means the user visit duration;

- “average visit depth” means the number of pages viewed during a visit;
 - “bounce rate” means the percentage of visitors who left the website directly from the login page or viewed not more than one website page; and
 - “traffic sources” means other web resources from which the user navigated to the website;
2. quantitative data obtained; and
 3. assessment of the existing experience of Russian and foreign colleagues.

For ease of analysis, the data collected were compiled into separate tables, and the indicators summarized based on the assessment were reformatted into visual elements to make them easier for researchers to understand.

4. RESULTS

Twenty-two websites of the Siberian and Far Eastern scientific libraries have been monitored for the use of web-based analytical tools and their openness. As a rule, statistics from the web-analytical tools of most libraries are closed to “outsiders”; in particular, 15 of the 22 viewed library websites have closed counters. At the time of the study, the following libraries had open access to the statistics of some web analytics tools: N. N. Muravyov-Amursky Amur Regional Scientific Library (LiveInternet), A. S. Pushkin Magadan Regional Universal Scientific Library (LiveInternet), Novosibirsk State Regional Scientific Library (Yandex.Metrica), N. G. Domozhakov National Library (LiveInternet), S. P. Krashennikov Kamchatka Regional Scientific Library (Rating@mail.ru), A. S. Pushkin Zabaikalsky Regional Universal Scientific Library (HotLog), and V. Ya. Shishkov Altay Regional Universal Scientific Library (Rating@mail.ru). This confirms the statement that most libraries close their analytical tool data.

According to Pr-cy.ru, the most commonly used web analytics tools of libraries are Google Analytics, Yandex.Metrica, and LiveInternet. Yandex.Metrica holds the leading position; this analytics system is used by 16 libraries, Google Analytics takes the second place (9 libraries) and, finally, LiveInternet takes the third position; this counter is installed on the websites of five Siberian and Far Eastern scientific libraries (Fig. 1). The data obtained by general browsing on the library website are slightly different, since Pr-cy.ru has not identified such web-analytical tools as Rating@mail.ru, Sputnik, and HotLog. All of these are free Russian-language analytical systems that allow collecting data on website traffic by installing the counter code

on the website pages.

As a result, we can conclude that the Pr-cy.ru marketing tool

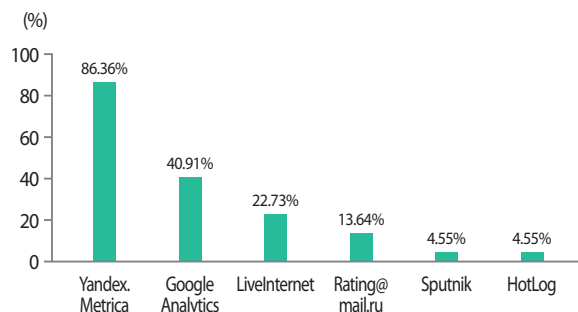


Fig. 1. Popular web analytics tools for scientific libraries in Siberia and the Far East (Russia).

defines only classical acknowledged web analytical systems (Table 1). Another interesting fact is that using general browsing, no information was found on any websites about the collection of statistical data using Google Analytics.

An important indicator of the success and popularity of the library website in virtual space is the increase in its traffic (attendance). To obtain a correct attendance assessment of a library web resource, it is necessary to constantly monitor a number of the selected key traffic metrics. Webometric attendance characteristics of the Siberian and Far Eastern scientific libraries according to SimilarWeb are presented in Table 2, with the exception of data from the website of Sholem Aleichem Birobidzhan Regional Universal Scientific Library (data not available in SimilarWeb), and data from eight libraries

Table 1. Web analytics tools for scientific libraries in Siberia and the Far East (June 2018)

Number	Library name	Website address	Web analytics tools (data from Pr-cy.ru)	Web analytics tools (de visu on the library website)
1	V. Ya. Shishkov Altay Regional Universal Scientific Library	http://akunb.altlib.ru/	Google Analytics LiveInternet Yandex.Metrica	LiveInternet Rating@mail.ru Yandex.Metrica
2	N. N. Muravyov-Amursky Amur Regional Scientific Library	http://www.libamur.ru/	Google Analytics LiveInternet	LiveInternet
3	Sholem-Aleichem Birobidzhan Regional Universal Scientific Library	http://bounb.eao.ru	Yandex.Metrica	-
4	Ugra State Library	http://okriib.ru/	Google Analytics	Rating@mail.ru
5	Krasnoyarsk Region State Universal Scientific Library	http://krslib.ru/	Yandex.Metrica	-
6	Far Eastern State Scientific Library	https://www.fessl.ru/	Yandex.Metrica	Yandex.Metrica
7	A. S. Pushkin Transbaikal Regional Universal Scientific Library	http://www.zabunb.ru/	Yandex.Metrica	Yandex.Metrica HotLog
8	I. I. Molchanov-Sibirsky Irkutsk Regional State Universal Scientific Library	https://www.irklib.ru/	-	Yandex.Metrica
9	S. P. Krasheninnikov Kamchatka Regional Scientific Library	http://www.kamlib.ru/	-	Rating@mail.ru Yandex.Metrica
10	V. D. Fyodorov Kemerovo Regional Scientific Library	http://www.kemrsl.ru	-	-
11	A. S. Pushkin Magadan Regional Universal Scientific Library	http://mounb.ru/	Google Analytics LiveInternet Yandex.Metrica	Yandex.Metrica LiveInternet
12	M. V. Chevalkov Altay Republic National Library	http://www.nbra.ru/	Yandex.Metrica	-
13	Buryat Republic National Library	http://nrb.ru/	Yandex.Metrica	Yandex.Metrica
14	Sakha Republic National Library	https://nlrs.ru/	Google Analytics Yandex.Metrica	-
15	N. G. Domozhakov National Library	http://www.nbdx.ru/	LiveInternet	LiveInternet
16	A. S. Pushkin Tyva Republic National Library	https://tuva-library.ru/	Yandex.Metrica	Sputnik Analytics Yandex.Metrica
17	Novosibirsk State Regional Scientific Library	http://ngonb.ru/	Google Analytics LiveInternet Yandex.Metrica	Yandex.Metrica
18	A. S. Pushkin Omsk State Regional Scientific Library	http://omsklib.ru/	Yandex.Metrica	Yandex.Metrica
19	A. M. Gorky Primorsky Regional Public Library	https://pgpb.ru/	Google Analytics Yandex.Metrica	-
20	Sakhalin Regional Scientific Universal Library	http://www.libsakh.ru/	Google Analytics Yandex.Metrica	-
21	A. S. Pushkin Tomsk Regional Universal Scientific Library	http://www.lib.tomsk.ru/	Google Analytics Yandex.Metrica	Yandex.Metrica
22	D. I. Mendeleev Tyumen Regional Scientific Library	www.tonb.ru/	Yandex.Metrica	-

which are only partially presented.

The high visit rate is shown by the website of the State Universal Scientific Library of the Krasnoyarsk Territory: 55.93 thousand per month. The average time of website visit is 5.17 minutes. It is also the best indicator among the websites of the Siberian and Far Eastern scientific libraries. Only the website of the Novosibirsk State Regional Scientific Library is slightly behind: The average visit time is 5.15 minutes, and the average

visit depth has the highest rate among the libraries analysed at 6.93 pages. The average visit depth for the websites of the Siberian and Far Eastern scientific libraries is 3.58 pages; based on the informational component of library websites, this is a pretty good indicator. The lowest bounce rate (45.78%) is for the website of A. S. Pushkin Omsk State Regional Scientific Library, and the highest (78.57%)—is the website of the Ugra State Library.

Table 2. Attendance of scientific libraries in Siberia and the Far East according to SimilarWeb (1 January-31 June 2018)

Number	Library name	Key traffic metrics							
		Visits (K)	Traffic source				Average visit time (min)	Average visit depth	Bounce rate (%)
			Direct traffic (%)	Referral traffic (%)	Search traffic (%)	Social network traffic (%)			
1	V. Ya. Shishkov Altay Regional Universal Scientific Library	22.5	13.45	11.88	68.03	5.68	1.50	2.75	69.07
2	N. N. Muravyov-Amursky Amur Regional Scientific Library	13.0	28.00	3.00	65.00	1.00	2.49	3.58	51.90
3	Ugra State Library	27.4	18.87	0.94	79.14	1.06	1.40	2.0	78.57
4	Krasnoyarsk Region State Universal Scientific Library	55.93	23.00	1.00	68.00	2.00	5.17	5.39	54.18
5	Far Eastern State Scientific Library	15.3	14.12	2.61	77.40	4.25	1.57	3.31	62.69
6	A. S. Pushkin Transbaikal Regional Universal Scientific Library	-	25.00	2.00	69.00	2.00	-	-	-
7	I. I. Molchanov-Sibirsky Irkutsk Regional State Universal Scientific Library	-	3.60	2.91	54.40	5.67	-	-	-
8	S. P. Krasheninnikov Kamchatka Regional Scientific Library	-	9.00	1.00	87.00	1.00	-	-	-
9	V. D. Fyodorov Kemerovo Regional Scientific Library	24.5	26.00	5.00	65.00	1.00	3.33	3.30	64.58
10	A. S. Pushkin Magadan Regional Universal Scientific Library	-	28.70	0.28	64.08	4.72	-	-	-
11	M. V. Chevalkov Altay Republic National Library	-	28.00	5.00	65.00	0.00	-	-	-
12	Buryat Republic National Library	-	26.09	1.99	68.08	1.92	-	-	-
13	Sakha Republic National Library	-	32.10	12.53	32.79	1.45	-	-	-
14	N. G. Domozhakov National Library	13.3	35.97	4.14	57.27	1.56	2.18	4.12	51.14
15	A. S. Pushkin Tyva Republic National Library	-	15.55	3.33	81.12	0.00	-	-	-
16	Novosibirsk State Regional Scientific Library	51.32	39.00	9.00	46.00	4.00	5.15	6.93	47.49
17	A. S. Pushkin Omsk State Regional Scientific Library	20.9	40.52	7.78	42.70	6.27	3.55	4.04	45.78
18	A. M. Gorky Primorsky Regional Public Library	10.9	15.23	2.25	78.97	2.02	1.58	2.28	68.77
19	Sakhalin Regional Scientific Universal Library	13.0	27.00	3.00	63.00	2.00	3.26	4.38	57.79
20	A. S. Pushkin Tomsk Regional Universal Scientific Library	26.3	24.51	8.12	61.48	5.47	1.42	2.25	73.77
21	D. I. Mendeleev Tyumen Regional Scientific Library	20.9	9.29	6.71	82.66	1.14	1.21	2.21	71.20

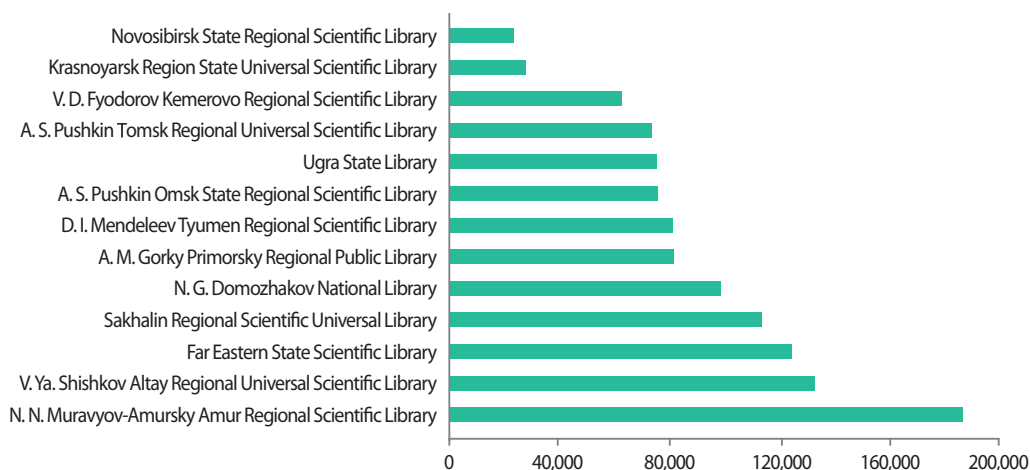


Fig. 2. Ranking of scientific libraries in Siberia and the Far East (Russia) according to SimilarWeb (1 January-31 June 2018).

Traffic sources are the channels through which users come to library web resources. A steady flow of users through these channels is the key to the success of any library web resource. Direct traffic is received by entering the website through a targeted search, typing in the search box the address of a particular website or going from previously saved bookmarks. The websites of A. S. Pushkin Omsk State Regional Scientific Library (40.52%), Novosibirsk State Regional Scientific Library (39.00%), and N. G. Domozhakov National Library (35.97%) have the highest rates of this type of traffic, which is an indicator of the positive image of their permanent audience. Referral traffic (Referrals) is the traffic of visitors who have followed links from third-party portals. The highest rate of this traffic is for the website of the National Library of the Republic of Sakha: 12.53%. The average rate of referral traffic for the Siberian and Far Eastern scientific libraries is 4.5%. High rates of transitions from search engines (Search), so-called organic traffic, are the result of well-conducted promotion of a web resource in search services. The websites of S. P. Krasheninnikov Kamchatka Regional Scientific Library (87.00%), D. I. Mendeleev Tyumen Regional Scientific Library (82.66%), and A. S. Pushkin National Library of the Republic of Tyva (81.12%) stand out for the highest rates of this kind of traffic, which is an indicator of the websites' quality. Traffic from social networks is quite relevant today because users spend a lot of time there, and the task of libraries is to use it correctly. The highest rates of traffic received from social networks are from the websites of A. S. Pushkin Omsk State Regional Scientific Library (6.27%), V. Ya. Shishkov Altay Regional Universal Scientific Library (5.68%), and I. I. Molchanov-Sibirsky Irkutsk Regional State Universal Scientific Library (5.67%). The websites of A. S. Pushkin Omsk

State Regional Scientific Library and V. Ya. Shishkov Altay Regional Universal Scientific Library receive the most traffic from VKontakte social network, along with the website of I. I. Molchanov-Sibirsky Irkutsk Regional State Universal Scientific Library—from Facebook. The average traffic rate from social networks for the Siberia and Far Eastern scientific libraries is only 2.9%, which is the result of the lack of effective library work on social networks. It is obvious that in terms of promoting their own resources and services, some libraries do not work with social networks and do not even have accounts, which affects the overall traffic. Knowledge of traffic sources will allow libraries to conduct marketing activities in virtual space most effectively.

The analysis of traffic metrics allowed ranking the library websites in Siberia and the Far East by attendance, comparing the key attendance indicators and highlighting the most effective websites (Fig. 2).

As on several library websites, some of the traffic metrics were absent in SimilarWeb, and only 13 websites of the Siberian and Far Eastern scientific libraries are presented for ranking, the traffic metrics of which are presented in full. A checklist of libraries is provided in Appendix. Each website selected for the study was carefully evaluated based on the checklist and data table. A high position in the rating of the websites' attendance of the Siberian and Far Eastern libraries is an indicator of competitive advantage among the analysed libraries' websites.

5. DISCUSSION

5.1. Web Analytics Tools

The analysis data showed that among the scientific libraries

of Siberia and the Far East, the most popular web-analytical systems for assessing the performance of their own activities are Yandex.Metrica (86.36%) and Google Analytics (40.91%).

The leading position of Yandex.Metrica analytical system is explained primarily by the fact that the system has a Russian-language interface since its inception. But today, other tools are also being Russified, so in the future, libraries may change their preference and give it to other analytical tools of similar functionality and ease of use. In a pilot study of the websites of scientific research organizations by a group of Russian authors (Guskov, Bykhovtsev, & Kosyakov, 2015), a similar pattern was found in the use of Google Analytics (41%) and Yandex.Metrica (77%) analytical systems. Turner (2010) finds the advantage of these analytics systems in the use of JavaScript code on web pages that allows collecting technical and demographic information. The choice in favour of analytics tools such as Yandex.Metrica and Google Analytics may indicate library interest in assessing the attendance of their websites, given the fact that these tools have great analytical functionality.

5.2. Visits

In our opinion, the first place among the key metrics to evaluate library performance on the web belongs to the visit metric. This metric is interesting to libraries as an indicator that shows the overall usability of the library website by users. As a rule, a high rate of visits has a positive interpretation, indicating the library website popularity (Fagan, 2014). Fagan (2014) clarifies the value of this indicator to assess library web resources performance in the “content” category. However, if this indicator is considered for the purpose of assessing the user “self-service,” then, according to Fagan (2014), it should be concluded that high rates of visits may indicate a poor or insufficiently clear navigation of the web resource (Fagan, 2014). The second version of the indicator interpretation (visits), in our opinion, requires additional qualitative assessment with the help of analytical systems with advanced functionality (visits to individual pages of the website, click map, goals setting, webvisor), which will allow specifying of the data and getting an objective assessment. In addition, the “self-service” evaluation is more focused on the analysis of specific website pages, allowing drawing conclusions about how clear the website is for the users. For the purposes of competitive intelligence, the total number of website visits is estimated. In this case, high values of the webometric indicator—visits—have a positive value, indicating the performance of the specific website in relation to other websites’ libraries. In our study, the website of the State Universal Scientific Library of the Krasnoyarsk Territory demonstrates high attendance, which allows us to talk about its

high popularity among users.

5.3. Average Visit Time

Another important key indicator is the average visit time. We accept that library websites are full of information, and therefore time is required for a thoughtful acquaintance (Rykhotorova & Udartseva, 2018). Based on this, high metric values indicate a general interest of users in the provided library content. The above average data are demonstrated only by three websites of scientific libraries in Siberia and the Far East: Krasnoyarsk Region State Universal Scientific Library, Novosibirsk State Regional Scientific Library, and A. S. Pushkin Omsk State Regional Scientific Library. The average visit time for the websites of scientific libraries in Siberia and the Far East is 2 minutes 51 seconds. If we evaluate this indicator from the point of “self-service” (Fagan, 2014), this is a good result. About three minutes is enough for a modern user not only to quickly solve the task, but also additionally analyse the related information. However, some pages of library websites require deeper studying, so it is important to conduct additional qualitative analysis using the functionality of authoritative analytical systems (Yandex.Metrica, Google Analytics). Thus the optimal session time for overall website assessment is about three minutes, while the optimal time for a particular page of the website will be calculated based on the tasks that the user can solve while staying on this page (number of links, number of words, presentation of visual information, availability of a fill-in form, etc.).

5.4. Average Visit Depth

High indicators of “visit depth” metrics can indicate both poor optimization of the website landing pages and user interest in the proposed content. Such interpretation ambiguity did not allow us to immediately draw the appropriate conclusions, but a deep analysis of the data obtained in comparison with each other revealed some regularity: High indicators of the “average visit time” metric are correlated with high indicators of the “visit depth” metric. Moreover, to obtain an objective assessment of this metric, one more metric should be taken into account: bounce rate. If the library website, along with high rates of visit depth, has a bounce rate that varies in the range of 35-65%, then this indicator can be interpreted as a positive result, indicating the user interest in the content. In turn, if the situation is the opposite, where both the visit depth indicator and the bounce rate are high, the result should be considered negative. It signals existing problems on the website (the website is not optimized for technical devices applied by users, links on the website pages are outdated, incomprehensible design, long page loading,

etc.). Thus the high indicators of the “visit depth” metric of the websites of the Novosibirsk State Regional Scientific Library, State Universal Scientific Library of the Krasnoyarsk Territory, and the Sakhalin Regional Scientific Universal Library in relation to the bounce rates show the audience’s general interest in the content provided.

5.5. Bounce Rate

A complex metric for library websites is the bounce rate (Fagan, 2014; Vecchione et al., 2016). Studies have shown that high bounce rates cannot be interpreted as a positive result (Vecchione et al., 2016). Most scientists are inclined to believe that the lower the bounce rate, the more effective the website is (Fagan, 2014; Guskov et al., 2015; Vecchione et al., 2016). Therefore, libraries should strive to reduce the overall bounce rate on the website. The average bounce rate for the analysed libraries in Siberia and the Far East is 61.3%. Reducing the overall bounce rate for a particular library website can be interpreted as its effective development, improving the current website processes, showing that users do not have problems interacting with the website and that they successfully find the necessary information. Additional monitoring of the bounce rate on website landing pages will help identify current problems (inactive links, non-unique content, slow page loading, inconvenient navigation, pop-ups, etc.), which will increase attendance and reduce the total number of the library website bounces.

5.6. Traffic Source

The study of web traffic of scientific organization websites showed that direct visits to the pages with scientific content are an indicator of the constant audience of the organization website (Guskov et al., 2015). In the course of this study, we noted that the websites of scientific libraries in Siberia and the Far East with high direct traffic rates had higher average visit times than did libraries with lower values. Thus high rates of this traffic can indicate both the web resource visit by the library staff and the well-deserved authority of the library among its regular readers.

Given that new visitors come mainly from search engines (Guskov et al., 2015), a preliminary conclusion can be made that a large audience share for the websites of Siberia and the Far Eastern scientific libraries is new users.

The analysis results show that the most effective traffic sources for libraries are: “direct traffic” (average indicator—23%); and “search traffic” (average indicator—65.53%), which allows making preliminary conclusions about the audience of websites of the Siberian and Far Eastern scientific libraries, distributing it into two segments: regular (23%) and new (65.53%) users.

In the process of research, the key metrics of library website traffic have been identified that will allow evaluation of the library performance on the web, and a possible approach was proposed for interpreting the generalized indicators of key metrics for six months obtained to conduct competitive intelligence using marketing tools.

Attendance estimation by key webometric indicators can be a full-fledged way to assess the performance of library websites, along with the already traditional method of evaluation upon criteria: Size, Rich files, Scholar, and Visibility (Dastani et al., 2018; Guskov et al., 2015). High rates of both methods indicate effective website development in the web environment and raises the ranking position.

6. RECOMMENDATIONS

Library website performance depends largely on the improvement of its webometric indicators, which reproduce the current situation of each website individually. From the discussion above, it is clear that libraries need to review the current strategy of their own websites’ development. For this reason, we provide the following recommendations to help libraries achieve greater efficiency. This study showed that the presented attendance key metrics most fully recreate the picture of the current development of library websites. Based on the data obtained, it is recommended to libraries with low visit rates (A. M. Gorky Primorsky Regional Public Library, N. N. Muravyov-Amursky Amur Regional Scientific Library, Sakhalin Regional Scientific Universal Library, N. G. Domozhakov National Library) to increase website productivity by reconstructing the whole website or its individual pages. In addition, a clear way to attract more users is placing interesting content and links to other popular resources on the pages of the library website, which in turn will contribute to the increase of another no less important indicator—average time spent on the website. Another recommendation is that the content should be up-to-date and constantly updated, which will increase the response rate of users to the published content. While high depth with low bounce rate indicates audience interest in the library website, low depth with high bounce rate indicates problems. As recommendations to libraries whose websites demonstrate such indicators (Ugra State Library, A. S. Pushkin Tomsk Regional Universal Scientific Library, D. I. Mendeleev Tyumen Regional Scientific Library), it is offered to carry out optimization to eliminate the arising problems when users interact with them. Library websites with high bounce rates (Table 2) should pay attention to the website navigation system

to facilitate the search for necessary information.

The increase in quantitative indicators of traffic contributes not only to the growth of the visitors' numbers, but also to the increase of awareness in general, along with conversion of traffic into targeted actions on the website: subscription to a newsletter, seeking advice, ordering literature from electronic catalogues, application for any services, etc. Libraries are recommended to work not only with the most efficient traffic sources ("direct traffic" and "search traffic") to increase their efficiency and promote their own web resources, but also pay attention to the insufficiently productive traffic sources ("social traffic") to conduct a detailed analysis and improve current performance. Based on the study results, we can give several recommendations for further research. In order to study library website convenience, we recommend usability testing. In combination with the results of the work made, it will allow taking into account the website user preferences. Another way is to make a detailed analysis of the website using the advanced functionality of analytical systems (Google Analytics, Yandex.Metrica) that will expand the results of this study regarding the attendance key metrics.

7. CONCLUSION

The experience of Russian libraries in the field of webometrics should be expanded by studying and transforming the existing foreign practice. The literature review showed that the current state of Russian libraries significantly differs from their foreign competitors. Libraries are just beginning to gain experience in this area, are engaged in the selection of web analytical tools, realize the need to use key metrics, and are building the necessary technology to assess the performance of their own web resources, while facing various analytical difficulties.

The study results show that the Siberian and Far Eastern scientific libraries use the analytical systems Yandex.Metrica, Google Analytics, and LiveInternet to collect statistics, which also confirms the hypothesis that scientific organizations are interested in assessing their own resources.

Metrics tracking with web analytics tools is a great way to manage library websites. For successful functioning of a web resource, it is important to conduct a webometric study: to analyse webometric indicators, conduct competitive intelligence, determine one's own target audience, study the factors affecting the implementation of targeted actions on the website, and so on. Literature review (Chepushtanova, 2016; Fagan, 2014; Guskov et al., 2015; Jansen, 2009; Kann, 2017; Udartseva & Rykhtorova, 2018; Vecchione et al., 2016) showed that key traffic

metrics are the foundation for evaluation of library websites' performance. By choosing attendance indicators for monitoring, libraries will be able to usefully conduct webometric studies of their own websites.

Using data from web-based marketing tools to collect and analyse data, such as SimilarWeb, Alexa, Pr-Cy.ru, and others, will allow libraries to provide an analytical overview of competitors in the web environment. Use of SimilarWeb service can be recommended as an alternative source for collecting key traffic metrics: visits, traffic source (direct traffic, referral traffic, search traffic, and social traffic), average visit duration, pages per visit, and bounce rate. Based on the received data, the most effective websites of the Siberian and Far Eastern scientific libraries have been revealed. The important webometric attendance indicators are considered separately, and the interpretation based on the study objectives is given. Analysis of webometric indicators allowed assessing the current state of scientific libraries in Siberia and the Far East. The three leaders are the websites of the Novosibirsk State Regional Scientific Library, State Universal Scientific Library of the Krasnoyarsk Territory, and V. D. Fyodorov Kemerovo Regional Scientific Library. Any library is able to improve its own position by collecting and analysing key indicators of their website traffic.

Analysis has shown that two types of traffic are the most effective for libraries: direct and search. The same conclusion was reached by scientists (Guskov et al., 2015) who studied the websites of universities of the Siberian Branch of the Russian Academy of Sciences (10 websites), which may reflect a certain regularity for the websites of scientific organizations. Further research will confirm or refute the hypotheses previously put forward in this study.

Analysis of traffic metrics showed how effective/inefficient the activity of scientific libraries in Siberia and the Far East is. Based on the traffic indicators, the rating of libraries in Siberia and the Far East was built, thus clearly presenting the current status of website attendance (traffic). This study results indicate the libraries' potential to improve their web resources. The study will give impetus to libraries with insufficient attendance to take the necessary measures to improve current performance and change their own development strategy.

As a result, we note that the generalized statistics of website traffic using SimilarWeb service does not give specifics on the existing problems of the website. Nevertheless, it is interesting, allowing us to assess the current development of the websites of the Siberian and Far Eastern scientific libraries. Projecting the positive experience of a competitive library towards oneself is one of the important reasons to conduct such research.

REFERENCES

- Björneborn, L., & Ingwersen, P. (2001). Perspectives of webometrics. *Scientometrics*, 50, 65-82.
- Chakravarty, R., & Wasan, S. (2015). Webometric analysis of library websites of higher educational institutes (HEIs) of India: A study through google search engine. *DESIDOC Journal of Library & Information Technology*, 35(5), 325-329.
- Chepushtanova, T. A. (2016, April 19-29). Web analyst in assessing the performance of the electronic library (on the example of the electronic library of the Novosibirsk State Regional Scientific Library). In B. A. Deich (Ed.), *Social and cultural practices in modern Russian society: Materials of the scientific forum of teachers, students and graduate students* (pp. 92-95). Novosibirsk State Regional Scientific Library.
- Dastani, M., Atarodi, A., & Panahi, S. (2018). Webometrics ranking of digital libraries of Iranian universities of medical sciences. *International Journal of Knowledge Content Development & Technology*, 8(3), 41-52.
- Fagan, J. C. (2014). The suitability of web analytics key performance indicators in the academic library environment. *The Journal of Academic Librarianship*, 40(1), 25-34.
- Guskov, A. E., Bykhovtsev, E. S., & Kosyakov, D. V. (2015). Alternative webometrics: Study of the traffic of the websites of scientific organizations. *Scientific and Technical Information Processing*, 42(4), 274-289.
- Guskov, A. E., Kosyakov, D. V., & Bykhovtsev, E. S. (2016, June 6-12). On the place of academic libraries in the web-space of science. *The Second World Professional Forum "The Book. Culture. Education. Innovations" ("Crimea-2016")* (pp. 422-429). NPLS&T Russia.
- Jacoby, G. A., & Luqi, V. B. (2007). Intranet model and metrics: Measuring intranet overall value contributions based on a corporation's critical business requirements. *Communications of the ACM*, 50(2), 43-50.
- Jansen, B. J. (2009). Understanding user-web interactions via web analytics. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 1(1), 1-102.
- Jowkar, A., & Didegah, F. (2010). Evaluating Iranian newspapers' web sites using correspondence analysis. *Library Hi Tech*, 28(1), 119-130.
- Kann, S. K. (2017). Evaluation of the library site using web-analytics tools. *Libraries of National Academies of Sciences: Functioning Problems, Development Trends: Scientific-Practical and Theoretical Collection*, 14, 130-139. Retrieved November 15, 2018 from http://nbuv.gov.ua/UJRN/bnan_2017_14_16.
- Lowe, D. (2012). Libraries in the digital age: Embodying the future. *Scientific and Technical Libraries*, 5, 68-80.
- Mamon, A. (2017). *Analyze the quality of the semantics of competitors through SimilarWeb. Collection of rules and practices, methods of collecting semantics*. Retrieved November 15, 2018 from <http://semyadro.pro/sbor/analiziruem-kachestvo-semantiki-konkurentov-cherez-similarweb-4>.
- Mikhailova, A. V. (2017). Culture institutes libraries' web pages analysis as an instrument of information and education resources integration external direction implementation. *Culture and Arts Herald*, 3(51), 49-59.
- Paul, A., & Erdelez, S. (2009). *Web analytics in library practice: Exploration of issues*. Retrieved November 15, 2018 from <https://onlinelibrary.wiley.com/doi/pdf/10.1002/meet.2009.14504603105>.
- Prantl, D., & Prantl, M. (2018). Website traffic measurement and rankings: Competitive intelligence tools examination. *International Journal of Web Information Systems*, 14(4), 423-437.
- Redkina, N. S. (2017). Web-analytics tools for library marketing purposes. *RBA Newsletter*, 80, 164-166.
- Rykhtorova, A. E., & Udartseva, O. M. (2018). Website users' segmentation to promote library resources and services. *Bibliosphere*, 3, 59-67.
- Sen, A., Dacin, P. A., & Pattichis, C. (2006). Current trends in web data analysis. *Communications of the ACM*, 49(11), 85-91.
- Srinivasan, S., Amir, A., Deshpande, P., & Zbarsky, V. (2004). Grammar-based task analysis of web logs. In D. Grossman (Ed.), *Proceedings of the thirteenth ACM international conference on information and knowledge management* (pp. 244-245). ACM Press.
- Tian, R. (2012). Research and evaluation on university library websites of China. *Wuhan Daxue Xuebao (Xinxi Kexue Ban) / Geomatics and Information Science of Wuhan University*, 37(Suppl 2), 151-155.
- Turner, S. J. (2010). Website statistics 2.0: Using google analytics to measure library website effectiveness. *Technical Services Quarterly*, 27(3), 261-278.
- Udartseva, O. M. (2018). An overview of webometrics in libraries: History and modern development tendencies. *Scientific and Technical Information Processing*, 45(3), 174-181.
- Udartseva, O. M., & Rykhtorova, A. E. (2018). Use of web analytics tools in evaluating the performance of ways to promote library resources. *Bibliosphere*, (2), 93-99.
- Vecchione, A., Brown, D., Allen, E., & Baschnagel, A. (2016). Tracking user behavior with google analytics events on an academic library website. *Journal of Web Librarianship*, 10(3), 161-175.

APPENDIX. A checklist of libraries

1. N. N. Muravyov-Amursky Amur Regional Scientific Library
2. Far Eastern State Scientific Library
3. V. D. Fyodorov Kemerovo Regional Scientific Library
4. Krasnoyarsk Region State Universal Scientific Library
5. N. G. Domozhakov National Library
6. Novosibirsk State Regional Scientific Library
7. A. S. Pushkin Omsk State Regional Scientific Library
8. A. M. Gorky Primorsky Regional Public Library
9. Sakhalin Regional Scientific Universal Library
10. V. Ya. Shishkov Altay Regional Universal Scientific Library
11. A. S. Pushkin Tomsk Regional Universal Scientific Library
12. D. I. Mendeleev Tyumen Regional Scientific Library
13. Ugra State Library

Development of Evaluation Perspective and Criteria for the DataON Platform

Suntae Kim*

Department of Library and Information Science,
Jeonbuk National University, Jeonju, Korea
E-mail: kim.suntae@jbnu.ac.kr

ABSTRACT

This study is a preliminary study to develop an evaluation framework necessary for evaluating the DataON platform. The first objective is to examine expert perceptions of the level of DataON platform construction. The second objective is to evaluate the importance, stability, and usability of DataON platform features over OpenAIRE features. The third objective is to derive weights from the evaluation perspective for future DataON platform evaluation. The fourth objective is to examine the preferences of experts in each evaluation perspective and to derive unbiased evaluation criteria. This study used a survey method for potential stakeholders of the DataON platform. The survey included 12 professionals with at least 10 years of experience in the field. The 57 overall functions and services were measured at 3.1 out of 5 for importance. Stability was -0.07 point and usability was measured as -0.05 point. The 42 features and services scored 3.04 points in importance. Stability was -0.58 points and usability was -0.51 points. In particular, the stability and usability scores of the 42 functions and services provided as of 2018 were higher than the total functions were, which is attributed to the stable and user-friendly improvement after development. In terms of the weight of the evaluation point, the collection quality has the highest weight of 27%. Interface usability is then weighted 22%. Subsequently, service quality is weighted 19%, and finally system performance efficiency and user feedback solicitation are equally weighted 16%.

Keywords: evaluation perspective, evaluation criteria, research data platform, DataON, Digital Library

Open Access

Received date: January 19, 2020
Accepted date: April 9, 2020

*Corresponding Author: Suntae Kim
Assistant Professor
Department of Library and Information Science,
Jeonbuk National University, 567, Baekje-daero, Deokjin-gu, Jeonju 54896,
Korea
E-mail: kim.suntae@jbnu.ac.kr

All JISTaP content is Open Access, meaning it is accessible online to everyone, without fee and authors' permission. All JISTaP content is published and distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). Under this license, authors reserve the copyright for their content; however, they permit anyone to unrestrictedly use, distribute, and reproduce the content in any medium as far as the original authors and source are cited. For any reuse, redistribution, or reproduction of a work, users must clarify the license terms under which the work was produced.

1. INTRODUCTION

In the era of the Fourth Industrial Revolution, where data is the center of research, interest in research data is hot. Devices ranging from small sensors to large laboratory instruments produce data called “crude oil” in real time. The era of data-intensive science, where so-called data is the central tool for research, has arrived. As Adams et al. (2017) argues, there is much to prepare for library users in the age of data-intensive science. This is why libraries and information centers should actively participate in the open science movement, which is embodied in open access, open data, and open repositories. In recent years, data is at the heart of the open science movement. In the field of academic research, this data is called research data.

Research data is the data that is collected, observed, or created for purposes of analysis to produce original research results. Research data has significant value in responsible research, which refers to the ability to justify conclusions on the basis of the data acquired and generated through research and that is furnished to other researchers for scrutiny and/or verification (Singh, Monu, & Dhingra, 2018). Research data is an essential part of the scholarly record, and management of research data is increasingly seen as an important role for academic libraries (Tenopir et al., 2017).

The open science movement aims to provide access to and use of publicly funded research data without legal, financial, or technical barriers, leading to various forms of activity in major industrialized countries. Funding agencies, libraries, publishers, and researchers are involved in these activities. As a representative activity, data management plans (hereafter referred to as DMP) have become commonplace, and DMP has begun to be piloted in September 2019 in Korea. The DMP addresses questions about research data types and formats, metadata standards, ethics and legal compliance, data storage and reuse, data management responsibility assignments, and resource requirements (Barsky et al., 2017).

In Korea, the policy direction for research data and DMP has recently been determined. According to the Republic of Korea’s Ministry of Science and ICT (2019), the government’s R&D budget in 2020 was raised to 24.1 trillion won, an increase of 17.3%. In the case of one of the major investment areas of the budget, the data economy revitalization and utilization-based upgrading sector, the big data platform will support the continuous creation of public and industrial innovation services through the supply of high quality data required by the market. In addition, Korea plans to create an ecosystem that creates value at all stages of data accumulation, distribution, and utilization through the linkage between platforms.

As part of this initiative, the Korea Institute of Science and Technology Information (hereinafter referred to as KISTI) is developing the Korea Open Research Data platform (DataON, <https://dataon.kisti.re.kr/>). DataON, which is being developed to systematically collect, manage, preserve, analyze, publish, and service research data produced in Korea, is developed by benchmarking OpenAIRE in Europe. OpenAIRE (<https://www.openaire.eu>) has been building research data platforms for more than 12 years, from 2008 to 2020. The world’s fastest and systematic research data project is under way. OpenAIRE has the largest research (meta) data corpus in the world. It has more than 36 million research and related data sets and provides search service. As of December 2019, it has 31.2 million publications, 1.6 million research data components, 110,000 softwares, and 3.1 million other contents, and as of 2018, it develops and services 42 features. Since 2018, DataON has been developing into KISTI’s own project with the support of the Ministry of Science and ICT. The pilot service was developed from June 2018 and completed pilot operation in June 2019, recently completing the first phase of construction in December 2019. It serves 1.8 million overseas data and 1,000 domestic data sets. KISTI has developed all the features and services of OpenAIRE at a large category level. It also provides a research data analysis environment that is not provided by OpenAIRE.

Such platforms should have a variety of features to enable them to create more value. For example, a developer may want to use the open application programming interface (API) or data provided by the platform. Administrators may want to host data managed by their organization on the platform. In particular, institutions that are small in terms of budget and manpower are likely to make such demands. Researchers may wish to submit research data they produce or collect directly to the platform and have a permanent access to their data. In addition, researchers may wish to use the platform’s data analysis tools, or may wish to use researcher community services. Research funding agencies may use statistical data related to national research data policy implementation and national data governance.

As mentioned above, the DataON platform is related to users of various interest groups. The platform evaluation should be continuously conducted from the platform construction stage to the final operation and function update. Thus, in order to evaluate the DataON platform, evaluation should be conducted using the evaluation viewpoints and evaluation criteria that are considered by various stakeholder groups. According to Saracevic (2004), the evaluation framework should include well-defined evaluation criteria. It should also include valid and reliable measurement methods. Digital library evaluation is characterized by the integration of quantitative and qualitative

evaluation. The research data platform can apply the same data collection method and evaluation method used to evaluate the digital library platform. However, since the core content provided by the DataON platform is research data, there may be differences in the viewpoints and criteria of evaluation even if the existing evaluation framework is used. The purpose of this study is fourfold. The first objective is to examine expert perceptions of the level of DataON platform construction. The second objective is to evaluate the importance, stability, and usability of DataON platform features over OpenAIRE features. The third objective is to derive weights from the evaluation perspective for future DataON platform evaluation. The fourth objective is to examine the preferences of experts in each evaluation perspective and to derive unbiased evaluation criteria. This study uses the evaluation framework of Xie (2006, 2008), which is used to evaluate digital library platforms. It investigates the weighting of the evaluation points presented by Xie to the stakeholder groups of the platform, and presents additional evaluation points and evaluation criteria.

2. PREVIOUS RESEARCH

The research data platform aims to provide researchers with research data services. The research data platform has the key functions of collecting, storing, managing, preserving, and publishing data. Thus, it is very similar to the digital library platform. However, only content and users are specialized in the field of research.

Unlike traditional library evaluation, digital library platform evaluation uses interface design, system performance, sustainability, user effect, and user participation (Xie & Matusiak, 2016). Digital library platforms need to adapt to the rapidly changing user needs and digital environment. Through this, the functions and services should be included in the digital library platform. For this reason an only way to evaluate a digital library platform cannot exist, and it must be evaluated from various angles. The same applies to the research data platform.

Platform evaluation involves detailed evaluation of the value and meaning of the platform. The evaluation of the platform is a systematic evaluation of how much the platform is achieving its intended purpose. It should be assessed in terms of how much the stakeholder group's needs related to the platform are being accommodated. Evaluation should also weigh its direction. There should be assessment of whether it is being developed and operated to meet platform objectives. Meanwhile, guidelines are needed for systematic evaluation. For example, guidelines for evaluation objectives, timing of evaluation, evaluation

targets, evaluation methods, and how the evaluation results are reflected in the platform development process are needed (Xie & Matusiak, 2016). Among them, the purpose of evaluation is very important because it is the basis of platform evaluation and the foundation of evaluation. Platform building includes planning phases, prototype development, platform building, platform testing, platform service launch and operation, and platform upgrade processes. At each stage, the platform can be evaluated. Thus, each stage may have a different objective. In addition, evaluation criteria and measurement methods may be different depending on the evaluation purpose (Buttenfield, 1999). In addition, evaluation criteria and methods are different depending on what is evaluated.

On the other hand, research on the evaluation framework for evaluating digital libraries has been progressing since the advent of digital libraries, as follows. According to Xie and Matusiak (2016), Tsakonas, Kapidakis, and Papatheodorou (2004) suggested usability, usefulness, and system performance as evaluation criteria. Saracevic (2004) presented content, technology, interfaces, processes/services, users, and contexts from an evaluation perspective. Fuhr et al. (2007) developed the Digital Library Evaluation Framework by incorporating Saracevic's four perspectives (construct, context, criteria, and methodology) and key questions related to why, what, and how. Candela et al. (2007) presented content, functionality, quality, policy, users, and architecture from an evaluation perspective. Xie (2006, 2008) presented collection quality, system performance efficiency, interface convenience, service quality, and user feedback from an evaluation perspective. Zhang (2010) presented content, description, interface, service, user, and context as an evaluation perspective. Zhang also conducted further research and validation of Saracevic's evaluation framework from the perspective of digital library interest groups (administrators, developers, librarians, researchers, and user groups). Tsakonas and Papatheodorou (2011) presented effectiveness, performance measurement, service quality, outcome assessment, and technical excellence as an assessment perspective. Lagzian, Abrizah, and Wee (2013) presented six evaluation points consisting of resource, motivation, location, process, people, and time and 36 evaluation criteria as evaluation points.

The core entities of the DataON platform are systems, collections, services, interfaces, and users. DataON platform evaluation needs to evaluate the system itself that composes the platform, and also needs to evaluate the data collection built into the system. Logical collection composition is provided by various types of services, and services can be accessed through various interfaces. The user using the interface may be a researcher or a system. Considering the core entities of the above

platform, it was determined that the weight of each viewpoint can be investigated by using the evaluation viewpoint presented by Xie among various evaluation frameworks presented in the previous research.

3. RESEARCH METHOD

This study is a preliminary study to develop an evaluation framework necessary for evaluating the DataON platform. This study used a survey method for potential stakeholders of the DataON platform. The survey included 12 professionals with at least 10 years of experience in the field. The survey was composed of three software developers, three data managers, three researchers, and three staff from the funding organization. In order to understand the DataON platform, KOR platform function evaluation was also performed against OpenAIRE function.

To examine expert perceptions of the level of the DataON platform building phase, the platform building phase (platform planning, platform prototyping, platform building, platform testing, platform service launch and operation, and platform upgrade) proposed by Buittenfield (1999) was used. For the review of DataON platform function compared to OpenAIRE function, the importance, stability, and usability of each function were evaluated on a 5-point scale.

In order to derive weights from the evaluation point of view and to further propose evaluation criteria, a user-driven evaluation model developed by Xie (2006, 2008) was used. Analytic Hierarchy Process (hereinafter referred to as AHP) method, known as a famous multi-criteria decision-making method to gain the weight of evaluation perspectives, was used also.

4. RESEARCH DATA PLATFORM CAKE AND STAKEHOLDERS

Xie (2006, 2008) proposed a user-driven evaluation model. The evaluation criteria were presented from five perspectives: interface convenience, collection quality, service quality, system performance efficiency, and user feedback. In this study, the hierarchical structure of the platform corresponding to Xie's point of view is constructed as shown in Fig. 1.

The bottom of Fig. 1 shows the hierarchy of the platform. In this paper, this hierarchy is called platform cake. The platform cake has a system layer at the bottom and a collection layer above it. A collection is a logical organization of resources served by the platform. These collections are provided as services that

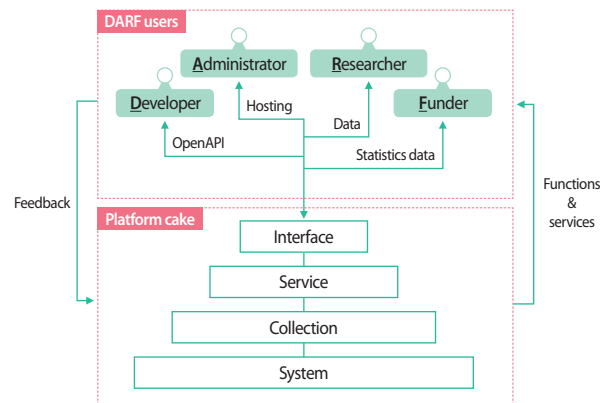


Fig. 1. Research data platform and its stakeholders. Developers may want to use the open application programming interface (API), data, etc. Administrators may want to use the platform's resource for their data to be hosted. A researcher may want to reuse the research data, analytic tools, community services, and so on. A funder may want to use the statistics data, etc. which can give the insights related to national research data policy building. Platform cake is a hierarchical representation of Xie's evaluation perspective.

work in a variety of ways. Thus, the service is located above the collection hierarchy. The service is provided as an interface of the platform, such as a search service performed by a user or an open API that is used systematically.

The top of Fig. 1 represents the stakeholder group surrounding the DataON platform. In this paper, the developer group, manager group, researcher group, and research fund support group are set as the core interest groups of the DataON platform.

Developer groups are very interested in APIs that can leverage research data or metadata managed by the platform. Types of developers can range from developers developing systems from other institutions that can work with the platform, to individual developers developing mashup services. A group of managers can deposit large amounts of research data from their institutions on the DataON platform. Administrators can also hope that platform data is retrieved from their institution's search service. A group of researchers can deposit their research data into the DataON platform. Research funding groups may be interested in the management and utilization of research data generated from research grants.

5. NATIONAL RESEARCH DATA PLATFORM: DATA ON EVALUATION PERSPECTIVES AND CRITERIA DEVELOPMENT

The survey was conducted on the same experts who participated in the evaluation of OpenAIRE and DataON

Table 1. Recognition status of each stakeholder group about the platform construction stage

Platform build phase	Developer			Administrator			Researcher			Funder		
	D1	D2	D3	A1	A2	A3	R1	R2	R3	F1	F2	F3
Platform planning												
Platform prototype development								0		0		
Platform building	0								0		0	
Platform testing				0	0	0	0					0
Platform service launch and operation		0	0									
Platform upgrade												

functions and services. We investigated the current DataON platform construction stage. The survey results are shown in the table below. Table 1 shows the recognition status of each stakeholder group about the platform construction stage. Two experts judged the level of DataON at the prototype development stage, and three experts considered the platform construction stage. Five experts judged it to be a test stage and two experts judged it to be a service start-up and operation stage. The data manager group determined that all DataON levels were in the testing phase. Taken together, DataON is currently at the platform building and testing stage.

Experts provided evidence for evaluating the level of the platform. The following is a summary of the developer group's comments on the platform. First, the implementation of basic functions is considered to be completed, and it is necessary to collect the base data for DataON operation. Second, it is concisely built on essential functions as a national research data platform and seems to focus on stable operation. Third, DataON needs to develop additional functions of OpenAIRE.

The following is an aggregate of comments from the management group on the platform. First, it seems that the platform has been established and the service has been started, but if it is diagnosed by referring to the contents and user status currently registered, it is determined that a certain group of users is testing it. Second, there is not much data construction, so it seems a bit too much to service in general. It would be better if an additional test is performed for each user, and more data needs to be built. Third, not all services are currently working, and there are errors and possibilities for improvements in the interface.

The following is a summary of researchers' opinions on the platform. First, more features need to be added for platform services, and errors need to be corrected. Second, although it shows the function implemented, it is necessary to reflect the opinion of many demands for service. Third, the frame of the platform is established, but functional improvement is in progress.

The following is an aggregate of opinions from a group of funding agencies on the platform. First, it is judged as a function definition and implementation stage. Second, after the relevant data DB is built, it is not verified, so it is judged as the platform construction stage. Third, it is estimated that there will be many trial and errors and inquiries when the researcher requests data registration and actually conducts the management operation in the research institute.

5.1. DataON Analysis

In order to analyze the functions and services of the DataON platform, we used importance, stability, and usability as the criteria for evaluation. Significance refers to the major extent of the function or service given the purpose of the research data platform. Stability refers to the degree to which a function or service operates reliably as intended. Usability refers to the degree to which features or services are easy to use. In the evaluation method, the importance, stability, and usability of each function and service were measured on a 5-point scale. Fig. 2 shows the result of comparing the importance, stability, and usability of the functions and services of OpenAIRE and DataON. The functions and services provided by OpenAIRE and DataON are composed of a total of 57 by allowing duplicates. As of 2018, OpenAIRE is providing 42 functions, and 21 of them (50%) are also known to be provided by DataON. Meanwhile, DataON provides 15 functions that are not provided by OpenAIRE. Likert scale 5-point scale was used to evaluate the importance, stability, and usability of DataON platform function compared to OpenAIRE function. The importance was measured from 1 to 5 points and the stability and usability from -2 to 2 points. The closer the importance is to 5 points, the higher the importance. If the stability and usability are -2, OpenAIRE is very good. In addition, it is designed to mean that DataON is very excellent in the case of 2 points.

The importance, stability and usability of all 57 functions were measured. In addition, we measured the importance, stability,

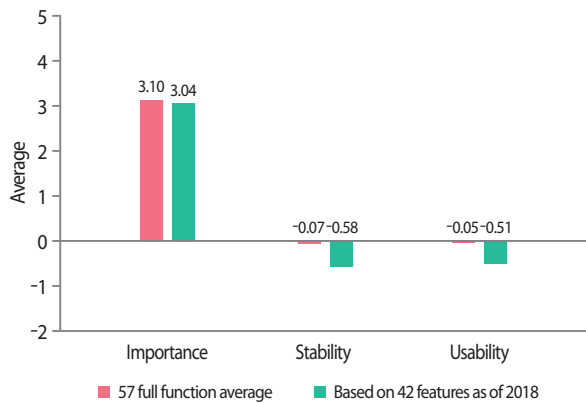


Fig. 2. Comparison of the importance, stability, and usability of functions and services of OpenAIRE and DataON.

and usability of 21 features implemented in DataON among 42 features of OpenAIRE, which were developed in 2018. All 57 functions were rated 3.06 out of 5 in importance. Stability was measured at 0.11 points and usability at 0.05 points. The 21 functions scored 3.12 in importance. Stability was 0.25 point and usability was measured to be 0.27 point. The importance is more than three points, and the functions being developed in OpenAIRE and DataON can be judged to be developed and operating well, reflecting the requirements of stakeholder groups. Both stability and usability are close to zero. Thus, there is no big difference between the functional stability and ease of use of OpenAIRE and DataON. However, when looking at the scores of the 21 features and services in DataON, which are minor differences, the DataON scores were slightly higher. This is because DataON is developing benchmarking of OpenAIRE function and applying the latest web technology and new requirements.

5.2. Analytic Hierarchy Process Design and Execution

As the first step in AHP, the hierarchical structure with a goal at the top level, the attribute/criteria at the second level and alternatives at the third were developed. Fig. 3 shows the hierarchical structure model for AHP analysis. Level 1 proposed the selection of a research data platform. As a component of Level 2, the evaluation viewpoint developed by Xie (2006, 2008) was used without modification. In addition, Level 3 uses various research data platforms that can replace DataON, including DataON.

Twelve experts in the DataON platform’s stakeholder group responded to the questions of the relative importance of perspectives for evaluating the platform, assuming platform selection. A normalized pair-wise comparison matrix was

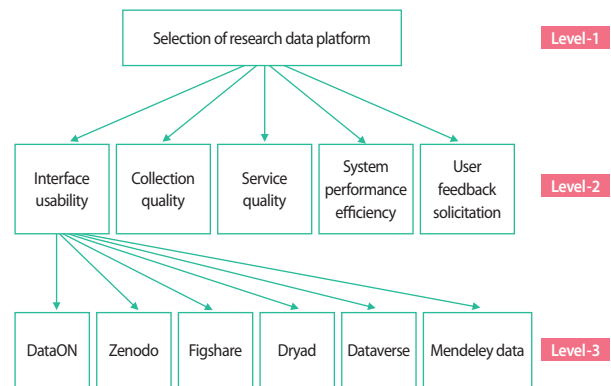


Fig. 3. A hierarchical structure for the Analytic Hierarchy Process.

created with the help of scale of relative importance. The scale is like as below.

- 3 - Very important
- 2 - Important
- 1 - Equal important
- 1/2 - Not important
- 0 - Not important at all

Table 2 shows a normalized pair-wise comparison and consistency index matrix prepared by normalizing the questionnaire data. Of the 12 respondents, 11 responded consistently with a consistency index of 0.1 or less. Only one researcher had a slightly higher consistency index of 0.127. In this study, the data of one researcher with a relatively high consistency index was considered as meaningful data and included in the weighted average calculation by evaluation criteria.

Fig. 4 and Table 3 below show the results of surveying the weight of the platform evaluation perspective for the stakeholder groups related to the DataON platform. In Fig. 4 and Table 3, ‘D’ means developer, ‘R’ means researcher, ‘A’ means administrator, and ‘F’ means funder.

Fig. 4 shows the weights of each evaluation perspective used to evaluate the platform. It is shown by respondents in the stakeholder group related to DataON platform. The consistency index that shows whether respondents consistently responded to the weighting survey by evaluation point of view is as follows. The administrator group responded most consistently with a consistency index average of 0.04. The developer and funder groups responded consistently, following the administrator group, with a consistency index average of 0.05. On the other hand, the average of the consistency index of the researcher

Table 2. Normalized pair-wise comparison and consistency index matrix related to evaluation perspectives

Interviewee	Interface usability	Collection quality	Service quality	System performance efficiency	User feedback solicitation	Consistency index
D1	0.284	0.162	0.161	0.172	0.221	0.077
D2	0.171	0.325	0.197	0.145	0.162	0.040
D3	0.140	0.305	0.214	0.097	0.244	0.023
A1	0.302	0.255	0.227	0.135	0.080	0.017
A2	0.262	0.249	0.191	0.167	0.131	0.060
A3	0.236	0.356	0.191	0.124	0.093	0.040
R1	0.241	0.352	0.170	0.145	0.092	0.060
R2	0.160	0.129	0.151	0.253	0.307	0.065
R3	0.226	0.263	0.257	0.145	0.110	0.127
F1	0.226	0.292	0.168	0.146	0.168	0.019
F2	0.151	0.276	0.213	0.263	0.098	0.036
F3	0.275	0.319	0.121	0.143	0.143	0.095

D, developer; A, administrator; R, researcher; F, funder.

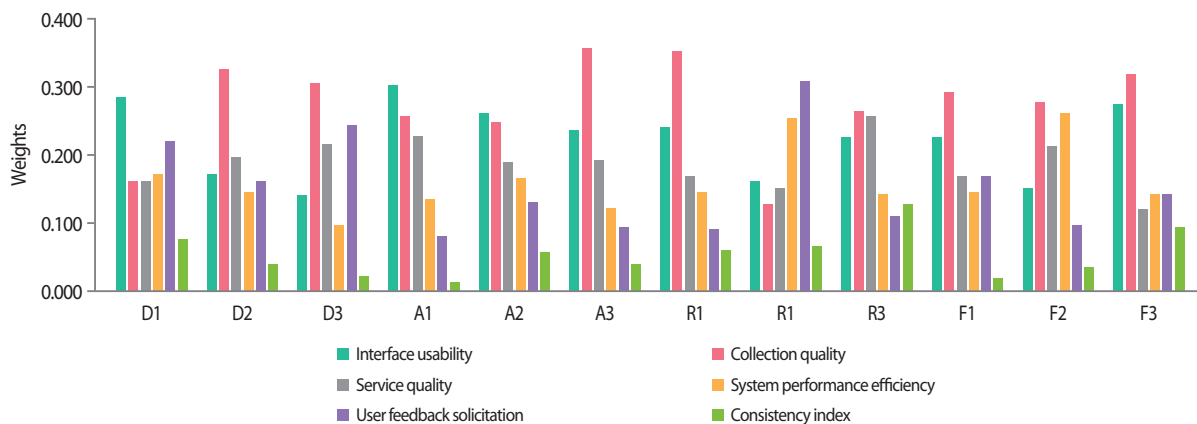


Fig. 4. Weights for the evaluation perspectives for evaluating the platform by stakeholder groups related to the DataON platform. D, developer; A, administrator; R, researcher; F, funder.

group was 0.08, which is a big difference from other groups. This is due to the slightly higher 0.127 consistency index that a researcher responded to.

Table 3 shows the weights of the evaluation points for evaluating the platform by the group of questionnaire respondents in the stakeholder group related to the DataON platform, in tones of color. The darker the purple, the higher the weight. The closer to red, the lower the weight. White is close to the middle of the weight. The developer group assigned the highest weight of 26.4% to the collection quality point of view, followed by the next highest weight of 20.9% to the user feedback solicitation point of view. The lowest weight gave 13.8% to the system performance efficiency perspective. The administrator

group gave 28.7% of the top weight to the collection quality perspective, followed by 26.7% of the next highest weight to the interface usability perspective. As the lowest weight, 10.2% was given to the viewpoint of user feedback solicitation, which was given the second highest weight in the developer group.

All DARF groups gave the highest weights in terms of collection quality. The administrator, researcher, and funder groups, except the developer group, gave the next highest weight to the interface usability perspective. Only the developer group gave the next highest weight to the user feedback solicitation perspective. In contrast, the administrator, researcher, and funder groups gave the lowest weight to the user feedback solicitation perspective.

Table 3. Evaluation point weight of each group, expressed in tones

Group class	Interface usability	Collection quality	Service quality	System performance efficiency	User feedback solicitation
D Avg.	19.8	26.4	19.1	13.8	20.9
A Avg.	26.7	28.7	20.3	14.2	10.2
R Avg.	20.9	24.8	19.3	18.1	17.0
F Avg.	21.8	29.6	16.7	18.4	13.6

D, developer; A, administrator; R, researcher; F, funder; Avg., average.

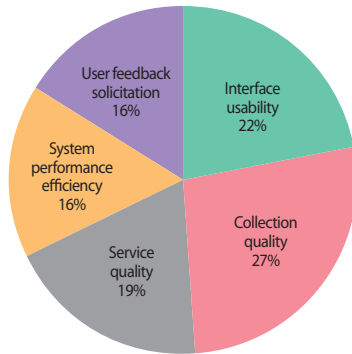
**Fig. 5.** Weights for platform assessment perspectives by interest group.

Fig. 5 shows the weighted data of each stakeholder group by evaluation perspective. Collection quality has the highest weight with 27%. Interface usability is then weighted 22%. Subsequently, service quality is weighted 19%, and finally system performance efficiency and user feedback solicitation are equally weighted 16%.

5.3. Investigating Preferences and Evaluation Criteria by Expert Group on Evaluation Perspectives

Table 4 shows preferred evaluation viewpoints by stakeholder groups of the DataON platform. In terms of interface usability, the DARF experts preferred it. On the other hand, only five of 12 experts preferred user feedback solicitation. In terms of collection quality, both administrators and funders were preferred, and developers and researchers were preferred by two and one, respectively. In terms of service quality, the entire funder was preferred. In terms of system performance efficiency, the entire group of administrators and researchers were preferred.

Table 5 shows the evaluation perspectives further suggested by the experts. The persistence perspective was proposed by eleven former experts, and the platform policy perspective was proposed by six experts. In addition, the effect on the user view, diverse collection and service, coherence with national policy direction, and hardware build support power were proposed.

Table 4. Preferred evaluation points by interest group of DataON platform

Evaluation perspective	Developer			Administrator			Researcher			Funder		
	D1	D2	D3	A1	A2	A3	R1	R2	R3	F1	F2	F3
Interface usability	0	0	0	0	0	0	0	0	0	0	0	0
Collection quality	0	0		0	0	0	0			0	0	0
Service quality		0	0	0	0		0		0	0	0	0
System performance efficiency	0	0		0	0	0	0	0	0	0	0	
User feedback solicitation			0			0	0	0	0	0		

Table 5. Suggested additional evaluation perspectives for each stakeholder group of DataON

Evaluation perspective	Developer			Administrator			Researcher			Funder		
	D1	D2	D3	A1	A2	A3	R1	R2	R3	F1	F2	F3
Persistence	0	0	0	0	0	0	0	0	0	0		0
Effect on the user							0			0		
Platform policy			0		0	0	0	0				0
Diverse collection and service	0											
Coherence with National Policy Direction			0									
Hardware build support power			0									

Table 6. Suggesting evaluation criteria corresponding to Xie's evaluation viewpoints for each stakeholder group of DataON platform

Evaluation perspective	Additional evaluation criteria
Interface usability	(D) mouse click count, ease of use, lesser click count, function closeness (A) facet, responsive web function (R) extensibility, visibility, quantitative display, and notification of progress (F) simplicity, user interface compatibility, waiting time, association term suggestion, intuitive manual
Collection quality	(D) ease to combine, description, limitation to use (A) openness, retention period, purpose (R) accuracy, verification methods, and quality assurance (F) security
Service quality	(D) policy, legal issues to consider (A) N/A (R) SNS connectivity, demand reflection, data quality control (documents and references) (F) N/A
System performance efficiency	(D) stability, response time (A) search speed, server reliability, response speed (R) compatibility, stability in repetition, stability in simultaneous use, environment restoration before and after search, network stability diagnosis or display (F) response time, stability, accuracy time, search speed
User feedback solicitation	(D) speed, help desk (A) real time consultation (R) preference, when to reflect (F) feedback time

D, developer; A, administrator; R, researcher; F, funder; N/A, not applicable; SNS, social networking service.

Table 6 shows the evaluation criteria proposed by experts in response to Xie's evaluation viewpoint for each stakeholder group of the DataON platform. In terms of interface usability, the simplicity of the interface and the response speed of web functions are proposed. In terms of collection quality, openness and usage restrictions, verification, and security related to the use of content have been proposed. In terms of service quality, social networking service (SNS) linkages and policies have been proposed. In terms of system performance efficiency, stability, response time, and search speed are proposed.

In this study, the evaluation criteria were examined without providing Xie's evaluation criteria to experts. Thus, there may be criteria that overlap with Xie's evaluation criteria. The results of this study may be used in further derivation and weighting studies of evaluation criteria, which will be conducted in subsequent studies.

6. DISCUSSION

According to Mayernik (2015), institutional support for data and metadata management in a single organization or discipline is uneven. Thus, for organizations or academic fields that require national support in terms of size or budget, a research data platform such as DataON is essential. Various platforms

such as DataON are being created by the efforts of countries and nonprofit organizations to support open science. For example, in the area of biodiversity research, the GFBio project addresses the challenges of data management caused by the size and heterogeneity of data. The GFBio project is collaborating with museums, archives, biodiversity researchers, and computer scientists to align with the data life cycle (Diepenbroek et al., 2014). The multi-disciplinary research data platform is the Open Science Framework (OSF). According to Foster and Deardorff (2017), the OSF is a tool that promotes open, centralized workflows by enabling capture of different aspects and products of the research lifecycle, including developing a research idea, designing a study, storing and analyzing collected data, and writing and publishing reports or papers. DataON is Korea's national research data platform. Thus, it is necessary to continuously benchmark the functions presented by OSF. Compared with the function of OSF, DataON provides research data analysis function which is not provided by OpenAIRE. However, it does not provide a function to support writing a report or a thesis. In particular, the community functions of DataON need to support the research data lifecycle. In other words, if the researcher's idea development and research design are supported from the beginning, it will be able to secure many users and activate the platform.

On the other hand, according to Tenopir et al. (2017), many

European libraries offer or plan to provide consulting rather than technical or practical research data service at the time. This may be considered to mean the expansion of a platform that can directly consult researchers beyond the platform's ability to provide an environment for storing and reusing research data due to the maturity and generalization of science and technology. Thus, it is necessary to further develop platform functions related to this, and further research on evaluation viewpoint and evaluation criteria is needed.

7. CONCLUSION

In this study, we conducted a preliminary study to develop an evaluation framework for evaluating the DataON platform under development in Korea. The survey was conducted with 12 experts from the stakeholder group surrounding the DataON platform. DataON is currently at the platform building and testing stage. As of 2018, OpenAIRE is providing 42 functions, and 17 of them (40.5%) are also known to be provided by DataON. As a result, the 57 overall functions and services were measured at 3.1 out of 5 for importance. Stability was -0.07 point and usability was measured as -0.05 point. The 42 features and services scored 3.04 points in importance. Stability was -0.58 points and usability was -0.51 points. Both criteria were measured with negative scores in stability and usability, indicating that the OpenAIRE function is more stable and convenient to use. In particular, the stability and usability scores of the 42 functions and services provided as of 2018 were higher than the total functions, which is attributed to the stable and user-friendly improvement after development.

As the first step in AHP, the hierarchical structure with a goal at the top level, the attribute/criteria at the second level and alternatives at the third were developed. Twelve experts in the DataON platform's stakeholder group responded to the questions of the relative importance of perspectives for evaluating the platform, assuming platform selection. All DARG groups gave the highest weights in terms of collection quality. The administrator, researcher, and funder groups, except the developer group, gave the next highest weight to the interface usability perspective. Only the developer group gave the next highest weight to the user feedback solicitation perspective. In contrast, the administrator, researcher, and funder groups gave the lowest weight to the user feedback solicitation perspective. When the weighted values of each stakeholder group's evaluation points are combined, collection quality has the highest weight with 27%. Interface usability is then weighted 22%. Subsequently, service quality is weighted 19%, and finally

system performance efficiency and user feedback solicitation are equally weighted 16%. On the other hand, in the research on additional evaluation points, the persistence, platform policy, effect on the user view, diverse collection and service, coherence with national policy direction, and h/w build support power were proposed. The simplicity of the interface and the response speed of web functions, openness and usage restrictions, verification, and security related to the use of the content, SNS linkages and policies, stability, response time, and search speed are proposed as an evaluation criterion.

This study proposed weights for evaluation points for evaluating DataON from experts in the interest group surrounding the DataON platform, and suggested additional evaluation points and examined evaluation criteria. Thus, the results of this study can be used as basic data in developing an evaluation framework for the DataON platform.

REFERENCES

- Adams Becker, S., Cummins, M., Davis, A., Freeman, A., Giesinger Hall, C., Ananthanarayanan, V., ... Wolfson, N. (2017). *NMC Horizon report: 2017 library edition*. Retrieved November 3, 2019 from <https://www.issueelab.org/resources/27498/27498.pdf>.
- Buttenfield, B. (1999). Usability evaluation of digital libraries. *Science & Technology Libraries*, 17(3-4), 39-59.
- Barsky, E., Brown, H., Ellis, U., Ishida, M., Janke, R., Menzies, E., ... Vis-Dunbar, M. (2017). *UBC Research data management survey: Health sciences: Report*. Retrieved from University of British Columbia website: <https://open.library.ubc.ca/cIRcle/collections/ubclibraryandarchives/494/items/1.0348070>.
- Candela, L., Castelli, D., Pagano, P., Thanos, C., Ioannidis, Y., Koutrika, G., ... Schuldt, H. (2007). Setting the foundations of digital libraries: The DELOS manifesto. *D-Lib Mag*, 13(3/4).
- Diepenbroek, M., Glöckner, F. O., Grobe, P., Güntsch, A., Huber, R., König-Ries, B., ... Triebel, D. (2014, September 22-26). Towards an integrated biodiversity and ecological research data management and archiving platform: The German federation for the curation of biological data (GFBio). In E. Plödereder, L. Grunske, E. Schneider, & D. Ull (Eds.), *Informatik 2014* (pp. 1711-1721). Gesellschaft für Informatik e.V.
- Foster, E. D., & Deardorff, A. (2017). Open Science Framework (OSF). *Journal of the Medical Library Association*, 105(2), 203-206.
- Fuhr, N., Tsakonias, G., Aalberg, T., Agosti, M., Hansen, P.,

- Kapidakis, S., ... Sølberg, I. (2007). Evaluation of digital libraries. *International Journal on Digital Libraries*, 8(1), 21-38.
- Lagzian, F., Abrizah, A., & Wee, M. C. (2013). An identification of a model for digital library critical success factors. *The Electronic Library*, 31(1), 5-23.
- Mayernik, M. S. (2015). Research data and metadata curation as institutional issues. *Journal of the Association for Information Science and Technology*, 67(4), 973-993.
- Ministry of Science and ICT. (2019). *Announcement of 2020 ministry of transitional and monetary affairs budget and government R&D budget*. Retrieved November 9, 2019 from <http://www.korea.kr/common/download.do?fileId=187894400&tblKey=GMN>.
- Saracevic, T. (2004, October 4-5). Evaluation of digital libraries: An overview. In M. Agosti, & N. Fuhr (Eds.), *DELOS WP7 Workshop on the Evaluation of Digital Libraries* (pp. 13-30). Springer.
- Singh, N. K., Monu, H., & Dhingra, N. (2018, February 21-23). Research data management policy and institutional framework. *2018 5th International Symposium on Emerging Trends and Technologies in Libraries and Information Services (ETTLIS)* (pp. 111-115). IEEE.
- Tenopir, C., Talja, S., Horstmann, W., Late, E., Hughes, D., Pollock, D., ... Allard, S. (2017). Research data services in European academic research libraries. *LIBER Quarterly*, 27(1), 23-44.
- Tsakonas, G., Kapidakis, S., & Papatheodorou, C. (2004, October 4-5). Evaluation of user interaction in digital libraries. *DELOS WP7 Workshop on the Evaluation of Digital Libraries* (pp. 45-60). Elsevier.
- Tsakonas, G., & Papatheodorou, C. (2011). An ontological representation of the digital library evaluation domain. *Journal of the American Society for Information Science and Technology*, 62(8), 1577-1593.
- Xie, H. I. (2006). Evaluation of digital libraries: Criteria and problems from users' perspectives. *Library & Information Science Research*, 28(3), 433-452.
- Xie, H. I. (2008). Users' evaluation of digital libraries (DLs): Their uses, their criteria, and their assessment. *Information Processing & Management*, 44(3), 1346-1373.
- Xie, I., & Matusiak, K. K. (2016). Evaluation of digital libraries. In I. Xie, & K. K. Matusiak (Eds.), *Discover digital libraries* (pp. 281-318). Amsterdam: Elsevier.
- Zhang, Y. (2010). Developing a holistic model for digital library evaluation. *Journal of the American Society for Information Science and Technology*, 61(1), 88-110.

Call for Paper

Journal of Information Science Theory and Practice (JISTaP)

We would like to invite you to submit or recommend papers to **Journal of Information Science Theory and Practice** (JISTaP, eISSN: 2287-4577, pISSN: 2287-9099), a fast track peer-reviewed and no-fee open access academic journal published by Korea Institute of Science and Technology Information (KISTI), which is a government-funded research institute providing STI services to support high-tech R&D for researchers in Korea. JISTaP marks a transition from Journal of Information Management to an English-language international journal in the area of library and information science.

JISTaP aims at publishing original studies, review papers and brief communications on information science theory and practice. The journal provides an international forum for practical as well as theoretical research in the interdisciplinary areas of information science, such as information processing and management, knowledge organization, scholarly communication and bibliometrics.

We welcome materials that reflect a wide range of perspectives and approaches on diverse areas of information science theory, application and practice. Topics covered by the journal include: information processing and management; information policy; library management; knowledge organization; metadata and classification; information seeking; information retrieval; information systems; scientific and technical information service; human-computer interaction; social media design; analytics; scholarly communication and bibliometrics. Above all, we encourage submissions of catalytic nature that explore the question of how theory can be applied to solve real world problems in the broad discipline of information science.

Co-Editors in Chief: Gary Marchionini & Dong-Geun Oh

Please click the "Online Submission" link in the JISTaP website (<http://www.jistap.org>), which will take you to a login/ account creation page. Please consult the "Author's Guide" page to prepare your manuscript according to the JISTaP manuscript guidelines.

Any question? Ji-Young Kim (managing editor) : jistap@kisti.re.kr

Information for Authors

The Journal of Information Science Theory and Practice (JISTaP), which is published quarterly by the Korea Institute of Science and Technology (KISTI), welcomes materials that reflect a wide range of perspectives and approaches on diverse areas of information science theory, application and practice. JISTaP is an open access journal run under the Open Access Policy. See the section on Open Access for detailed information on the Open Access Policy.

A. Originality and Copyright

All submissions must be original, unpublished, and not under consideration for publication elsewhere. Once an article is accepted for publication, all papers are accessible to all users at no cost. If used for other researches, its source should be indicated in an appropriate manner and the content can only be used for uncommercial purpose under Creative Commons license.

B. Peer Review

All submitted manuscripts undergo a single-blind peer review process in which the identities of the reviewers are withheld from the authors.

C. Manuscript Submission

Authors should submit their manuscripts online via Article Contribution Management System (ACOMS). Online submission facilitates processing and reviewing of submitted articles, thereby substantially shortening the paper lifecycle from submission to publication. After checking the manuscript's compliance to the Manuscript Guidelines, please follow the "Online Submission" hyperlink in the top navigation menu to begin the online manuscript submission process.

D. Open Access

With the KISTI's Open Access Policy, authors can choose open access and retain their copyright or opt for the normal publication process with a copyright transfer. If authors choose open access, their manuscripts become freely available to public under Creative Commons license. Open access articles are automatically archived in the KISTI's open access repository (KPubS, www.kpubs.org). If authors do not choose open access, access to their articles will be restricted to journal users.

E. Manuscript Guidelines

Manuscripts that do not adhere to the guidelines outlined below will be returned for correction. Please read the guidelines carefully and make sure the manuscript follows the guidelines as specified. We strongly recommend that authors download and use the manuscript template in preparing their submissions.

Manuscript Guidelines

1. Page Layout :

All articles should be submitted in single column text on standard Letter Size paper (21.59 × 27.94 cm) with normal margins.

2. Length :

Manuscripts should normally be between 4,500 and 9,000 words (10 to 20 pages).

3. File Type :

Articles should be submitted in Microsoft Word format. To facilitate the manuscript preparation process and speed up the publication process, please use the manuscript template.

4. Text Style :

- Use a standard font (e.g., Times New Roman) no smaller than size 10.
- Use single line spacing for paragraphs.
- Use footnotes to provide additional information peripheral to the text. Footnotes to tables should be marked by superscript lowercase letters or asterisks.

5. Title Page :

The title page should start with a concise but descriptive title and the full names of authors along with their affiliations and contact information (i.e., postal and email addresses). An abstract of 150 to 250 words should appear below the title and authors, followed by keywords (4 to 6).

Author1

Affiliation, Postal Address. E-mail

Author2

Affiliation, Postal Address. E-mail

ABSTRACT

A brief summary (150-250 words) of the paper goes here.

Keywords : 4 to 6 Keywords, separated by commas.

6. Numbered Type :

1. INTRODUCTION

All articles should be submitted in single column text on standard letter size paper (21.59 × 27.94 cm) with normal margins[1 . Text should be in 11-point standard font (e.g., Times New Roman) with single line spacing.

[1 Normal margin dimensions are 3 cm from the top and 2.54 cm from the bottom and sides.

2. SECTIONS

The top-level section heading should be in 14-point bold all uppercase letters.

2.1. Subsection Heading 1

The first-level subsection heading should be in 12-point bold with the first letter of each word capitalized.

2.1.1. Subsection Heading 2

The second-level subsection heading should be in 11-point italic with the first letter of each word capitalized.

7. Figures and Tables :

All figures and tables should be placed at the end of the manuscript after the reference list. To note the placement of figures and tables in text, “Insert Table (or Figure) # here” should be inserted in appropriate places. Please use high resolution graphics whenever possible and make sure figures and tables can be easily resized and moved.

Figure

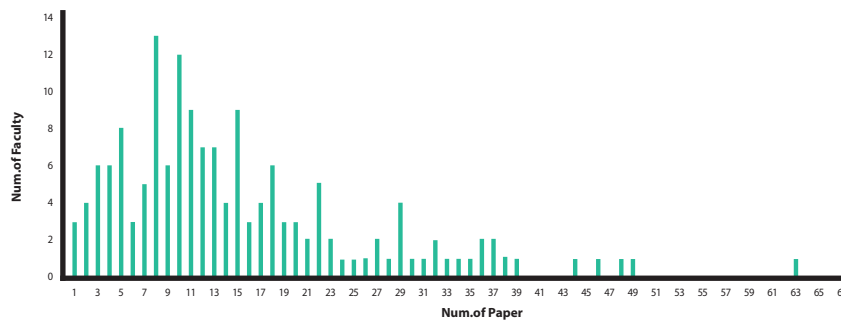


Fig. 1. Distribution of authors over publication count.

Table

Table 1. The title of table goes here

Study	Time period study	Data
Smith Wesson (1996)	1970 - 1995	684 papers in 4 SSCI journals
Reeves [a (2002)	1997 - 2001	597 papers in 3 SSCI journals
Jones Wilson [b (2011)	2000 - 2009	2,166 papers in 4 SSCI journals

[a Table footnote a goes here

[b Table footnote b goes here

8. Acknowledgements :

Acknowledgements should appear in a separate section before the reference list.

9. Citations :

Citations in text should follow the author-date method (authors' surname followed by publication year).

- Several studies found... (Barakat et al., 1995; Garfield, 1955; Meho & Yang, 2007).
- In a recent study (Smith & Jones, 2011)...
- Smith and Jones (2011) investigated...

10. Reference List :

Reference list, formatted in accordance with the American Psychological Association (APA) style, should be alphabetized by the first authors last name.

Journal article

- Author, A., Author, B. & Author, C. (Year). Article title. *Journal Title*, volume(issue), start page-end page.
- Smith, K., Jones, L. J., & Brown, M. (2012). Effect of Asian citation databases on the impact factor. *Journal of Information Science Practice and Theory*, 1(2), 21-34.

Book

- Author, A., & Author, B. (Year). *Book title*. Publisher Location: Publisher Name.
- Smith, K., Jones, L. J., & Brown, M. (2012). *Citation patterns of Asian scholars*. London: Sage.

Book chapter

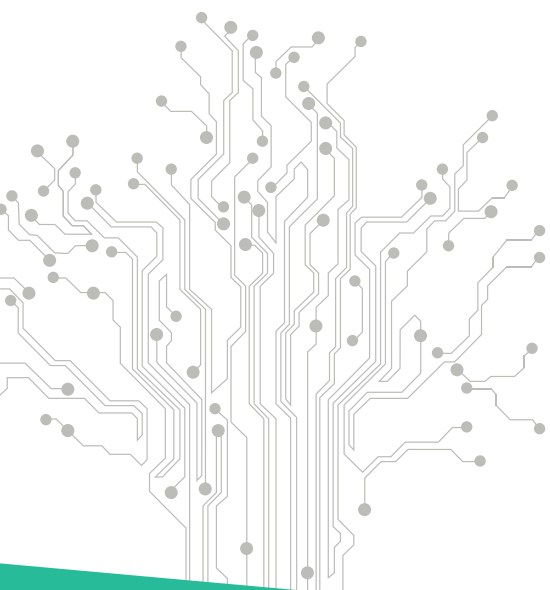
- Author, A., & Author, B. (Year). Chapter title. In A. Editor, B. Editor, & C. Editor (Eds.), *Book title* (pp. xx-xx). Publisher Location: Publisher Name.
- Smith, K. & Brown, M. (2012). Author impact factor by weighted citation counts. In G. Martin (Ed.), *Bibliometric approach to quality assessment* (pp. 101-121). New York: Springer.

Conference paper

- Author, A., & Author, B. (Year, Month Day). Article title. In A. Editor & B. Editor (Eds.), *Conference title* (pp. xx-xx). Publisher Name.
- Lin, H., Sun, B., Wu, J. & Xiong, H. (2016, June 24-26). Topic detection from short text: A term-based consensus clustering method. In B. Yang (Ed.), *2016 13th International Conference on Service Systems and Service Management (ICSSSM 2016)* (pp. 1-6). IEEE.

Online document

- Author, A., & Author, B. (Year). Article title. Retrieved *month day, year* from URL.
- Smith, K. & Brown, M. (2010). The future of digital library in Asia. *Digital Libraries*, 7,111-119. Retrieved *May 5, 2010*, from <http://www.diglib.org/publist.htm>.



JISTaP

Journal of Information Science
Theory and Practice

<http://www.jistap.org>



66, Hoegi-ro, Dongdaemun-gu, Seoul, Republic of Korea (ZIP code: 02456)
Tel. +82-2-3299-6102 Fax. +82-2-3299-6067 <http://www.jistap.org>