International Journal of Information Science and Management
Vol. 20, No. 1, 2022, 91-118
https://dorl.net/dor/20.1001.1.20088302.2022.20.1.6.9

Original Research

A Framework for Assessing the Quality of Wikipedia Articles: A Meta-synthesis of the Literature

Fatemeh Fahimnia
Associate Professor, Department of Information and Knowledge Science, faculty of Management, University of Tehran. Tehran. Iran
fahimnia@ut.ac.ir ORCID iD: https://orcid.org/0000-0002-8991-5391

Mansoureh Damerchiloo
Ph.D. Student, Department of Information and Knowledge Science, faculty of Management, University of Tehran. Tehran, Iran
Corresponding Author:
mansoureh.damirchi@gmail.com ORCID iD: https://orcid.org/0000-0002-8991-53910000-0002-9565-8491

Mohammad Khandan
Assistant Professor, Department of Information and Knowledge Science, faculty of Management, University of Tehran. Tehran, Iran
khandan@ut.ac.ir ORCID iD: https://orcid.org/0000-0003-3138-2470

Mahshid Eltemasi
Assistant Professor, Department of Information and Knowledge Science, faculty of Management, University of Tehran. Tehran, Iran
eltemasi@ut.ac.ir ORCID iD: https://orcid.org/0000-0002-7675-2626

Received: 20 January 2021
Accepted: 17 March 2021

Abstract

This study aimed to design and validate an information quality assessment framework based on the systematic review of the literature. A meta-synthesis method was applied to identify features and dimensions for designing the framework for assessing the quality of Wikipedia. Following this, the validity of the obtained framework was evaluated by the Kappa Test of Agreement. The statistical population consists of all scientific documents related to the quality of information and a sample of 39 documents selected based on CASP. MAXQDA 11 was used to analyze data. Nine dimensions were identified and classified across 3 levels. The features and components necessary to evaluate each dimension were identified and explained. The strong level approved the framework of expert agreement (0.654). The proposed framework can be used to assess the Wikipedia quality independently of specialists, improve the quality of articles through identified effective features, and ultimately, build tools and practical guidelines to assess the quality of information.

Keywords: Information quality, assessment, Wikipedia, Meta-synthesis.

Introduction

Today, Web 2.0 and social media have become an integral part of people's life. The success and popularity of Web 2.0 lie in its user-generated content (UGC) (Dalip, Gonçalves, Cristo & Calado, 2017). Web 2.0 allows users to interact and collaborate in a social media environment as content creators (Weitzel, de Oliveira & Quaresma, 2014) to generate and share their content (Kim, Gudewicz, Dighe & Gilbertson, 2010).

The increasing interest of users in UGC has led to uncontrollable growth in content

production by users, together with an increase in content-sharing sites (Kim et al., 2010). In the past, there were legal and approved mechanisms to check and determine the quality of information such as expert review processes or data cleaning and text editing, but, in Web 2.0, these methods serve little if any purpose, as UGC-based applications make it possible to produce and publish information by anonymous users quickly and with minimal monitoring. There is no specific standard for evaluating the UGC; but misinformation will lead to wrong decisions, costly for individuals and society. As such, it is essential to consider the quality of the user-created content.

One of the largest and most popular user-generated content on the Web is Wikipedia, the online and free multilingual encyclopedia. Although there are some criticisms on the open nature of its editing process that inherently poses risks of inaccuracies, a study showed that Wikipedia articles are ranked in the top 10 results of most search engines (Laurent & Vickers, 2009). According to Internet traffic information from Alexa, Inc[1], Wikipedia ranks as one of the most visited websites.

However, as with all UGC applications, determining the quality of Wikipedia articles is not an easy task for human users. The high number of Wikipedia articles has made it impossible for editors to evaluate the quality of all Wikipedia articles. Another problem is that Wikipedia includes a wide range of topics, and judging their quality requires experts in various subject areas that are not usually available. On the other hand, Wikipedia is a dynamic resource, and any editing done on an article may affect its quality. In other words, the evolving nature of Wikipedia results in increased complexity in assessing the quality of articles. Therefore, the existence of a framework for the automatic evaluation of Wikipedia is of considerable importance.

Contribution of this study

This study contributes to the general knowledge in the field of information quality assessment. The proposed framework in the present study is a comprehensive conceptual framework that can be applied to the assessment of Wikipedia article quality, can help end-users of Wikipedia determine the quality of information in the specific subject area, and help researchers better comprehend the assessment of information quality. In addition, it can help editors of Wikipedia improve the quality of content by recognizing the factors affecting the quality of information, as the effective features have been identified, and editors can improve the quality of articles by paying attention to them. Also, the framework can provide the basic knowledge for software programmers to design and build a suitable tool for automatically evaluating and measuring the quality of Wikipedia articles.

Objectives of study

The primary purpose of the current study is to determine the main features and dimensions of Wikipedia information quality to develop a conceptual framework. In this regard, the present study pursues the following objectives:

- ✓ Determining the main features and dimensions of Wikipedia quality
- ✓ Proposing a framework to assess the quality of Wikipedia articles based on their components and features
- ✓ Validating the proposed framework

Literature Review

Due to the importance of "information quality", several studies have been conducted in this field. These can be divided into two groups. Firstly, user-based studies that identify the aspects of information quality from the perspective of the end-user (Han, 2018; Orehovacki, 2010; Shahbazi, Farajpahlou, Osareh, & Rahimi, 2019; Tao, LeRouge, Smith & De Leo, 2017); secondly, product-based studies that seek to identify the aspects of information quality according to the characteristics of the information source (Blumenstock, 2008a, 2008b; Włodzimierz Lewoniewski, Khairova, Węcel, Stratiienko & Abramowicz, 2017; Włodzimierz Lewoniewski & Węcel, 2017; Włodzimierz Lewoniewski, Węcel, & Abramowicz, 2017; Stvilia, Gasser, Twidale & Smith, 2007; Stvilia, Twidale, Smith & Gasser, 2005; Warncke-Wang, Cosley & Riedl, 2013). In most of these studies (user-based or product-based), one or more aspects of information quality are identified, and only a few studies provide a framework or categorize the identified aspects. Wang and Strong's research (1996) is one of the first of such studies. According to their proposed framework, the aspects of information quality can be classified into four categories: 1) Intrinsic IQ - which includes aspects such as accuracy, objectivity, believability, and reputation, all of which embody internal characteristics of information; 2) Contextual IQ - which is related to the quality requirements of the information that must be considered in the context of information products such as relevancy, timeliness, completeness, and value-adding; 3) Representational IQ - which emphasizes how to display information and includes interpretability, ease of understanding, concision, and consistency of representation; and 4) Accessibility IQ - which addresses aspects such as accessibility and security of access.

In order to create a comprehensive and practical method for evaluating information quality, Lee, Strong, Kahn & Wang (2002) proposed the AIM quality evaluation method, which consists of three components. Its primary component is a 2x2 model that includes product quality, service quality, formal specifications, and customer expectations. This means that depending on whether the information is considered a product or a service, its appearance may be evaluated or consumer expectations met. The second component is a questionnaire to measure the quality of information and includes all aspects that are important for information consumers and managers. Many of these are components of the aspects in the 2x2 model. The third component includes two analytical techniques to interpret the questionnaire results and ultimately apply the gathered data to improve the quality of information.

Another framework proposed for information quality is the InfoQual framework. It is based on semiotic theory[2] and describes form-, meaning-, and use-related aspects of information (Price & Shanks, 2004).

Naumann and Rolker (2005) proposed a new, assessment-oriented way to classify information quality criteria by reviewing relevant literature. The results showed that the three main sources of assessing the quality of information are: user perception; the source of information, and the process of accessing information; and based on this, the criteria for evaluating information, which was divided into the three categories of subjective criteria, objective criteria, and process criteria respectively. In the Subjective method, the scores are determined only by the users and based on their perspective, experience, and background, while in the Objective method, the quality of information is determined by careful analysis of the information source. In the Process method, scores of process criteria are determined by querying, such as response time.

Yaari, Baruchson-Arbib & Bar-Ilan (2011) divided the criteria of information quality assessment in Wikipedia articles into two measurable and non-measurable groups. Each group classified the identified aspects into two categories: based on the article page and revision history.

Leite, Gonçalves, Teixeira & Rocha (2016) developed a model for assessing the data quality of health websites. At first, they reviewed the literature to identify the common aspects and characteristics in assessing the quality of the information in health websites; then, using the Delphi method, the respective weights of each aspect were determined. Twenty-three aspects were identified and organized in three categories: intrinsic (8 aspects), contextual (10 aspects), and representational (5 aspects).

Given that quality is a multidimensional concept, Dalip et al. (2017) proposed a general, multi-view framework to assess the quality of user-generated content on Web 2.0. In that study, quality is regarded as a combination of independent assessments, and each assessment is based on several indicators or criteria. The proposed framework, called MVIEW, consists of three steps: a) Collecting and determining a set of information quality aspects that are classified in several categories, including article length, readability, relevance, writing style, structure, editing history, and user; b) The evaluation of each of the mentioned categories in the first step by experts and specialists; and c) The final phase, in which the opinions of experts are combined using weighting methods.

In reviewing information quality frameworks and models, it was observed that, in most cases, the quality of information had been considered in general, whereas UGC-based applications have different features, and each of these applications has its characteristics. Due to the importance and popularity of Wikipedia, several studies have been conducted on the quality of its articles, but in most of them, only one or more aspects of information quality have been examined (Betancourt, Segnini, Trabuco, Rezgui & Jullien, 2016; Blumenstock, 2008a, 2008b; Lu, Zhang, & Li, 2013; Schmidt & Zangerle, 2019; P. Wang & Li, 2020). On the other hand, as far as we know, most studies have applied survey methods (questionnaires), interviews, or observations, and none of them has considered the meta-synthesis method to identify the components of information quality. However, in the era of information explosion, meta-synthesis is one of the most effective research methods because it allows the researcher to apply findings of related qualitative studies with a structural way in which to integrate the results and to create a comprehensive interpretation of them that will provide a systematic approach and enable the exploration of new and fundamental topics and concepts.

Therefore, this study is intended to develop a framework for assessing the quality of Wikipedia by identifying the factors affecting the quality of information on Wikipedia, using meta-synthesis and careful review of the literature in the field of information quality. Given that information quality assessment is a two-way process involving the information source itself on the one hand and the end-user on the other. Since the present study seeks to provide a framework for automatically assessing the information quality, and independent of the human factor, it has only considered the components and features of the information source itself.

Methodology

The present study is a qualitative meta-synthesis study. Meta-synthesis is an appropriate method for investigating the previous studies to explore the relationship between variables and discover the factors affecting the subject under study. In other words, this method is used to

combine previous studies for integrating the qualitative findings of primary research (Sandelowski & Barroso, 2006). Meta-synthesis can discover powerful interpretations and clarify a more remarkable ability to generalize and improve levels of abstraction, all of which help researchers revise or reject existing theories (Branger, O'Connell & Peacock, 2018; Sherwood, 1999). It can be referred to as the "research of research" (Hansen, Draborg & Kristensen, 2011) because through this method, we can systematically summarize, synthesize and compile primary findings studies on a specific topic (Finfgeld, 2003; Sandelowski & Barroso, 2006; Walsh & Downe, 2005). By using meta-synthesis, researchers can present in-depth interpretations of phenomena, propose new concepts, and enhance the process of theory building (Finfgeld, 2003; Sandelowski & Barroso, 2006).

This paper uses meta-synthesis as a research methodology because its specifications are appropriate to the purposes of this paper. Using this method, we can summarize, synthesize and integrate primary findings of previous studies and form general knowledge of the quality assessment. By using meta-synthesis and the seven steps of Sandelowski and Barroso (Sandelowski & Barroso, 2006), components and aspects were identified to design the framework for assessing the quality of Wikipedia.

In the present study, IQ dimension was defined as any feature of the IQ concept. For measuring a specific dimension of an information entity, the dimension needs to have been constructed as a measurable feature of the entity (Stvilia et al., 2007). The features of Wikipedia articles include text features, review features, and network features. Text Features such as the length, number of sections, or readability of a Wikipedia article are extracted from the text of the article. Review features are extracted from the history of each article, e.g., the number of edits or editors, the last or first edit date, and the citation or page rank are considered network features (Bassani & Viviani, 2019b). Finally, according to Stvilia (2006), we define the framework as a multidimensional structure consisting of general concepts, relationships, classifications, and methods that can be used as a resource and guide for developing information quality measurement models in a specific field.

Meta-synthesis implementation steps include: setting up research questions, searching for related literature, selecting documents, extracting information, analyzing and synthesizing the findings, quality control, and providing findings.

For setting up research questions, fundamental questions were raised, and the document domain was determined through these questions. (Table 1)

Table 1

Questions for the first step of meta-synthesis

What	Identifying the components, features and dimensions of information quality in Wikipedia.
Who	All scientific documents, including scientific papers, theses and book chapters that addressed the components, aspects, features and criteria of information quality in Wikipedia.
When	Wikipedia was launched in 2001, so, all existing documents published from 2001 to 2019 (the date of the research) were considered
How	Thematic review of documents, coding of the text, and categorization of concepts.

Then a systematic literature search for relevant studies was conducted using the databases: Web of Science, Scopus, Emerald, and Proquest; because of their relative comprehensiveness

in scientific indexing resources. The search query was determined by reviewing related research and identifying repetitive keywords and their synonyms. In fact, a variety of related keywords were used to increase the comprehensiveness of the results, which can be summarized as follows:

"Information quality" AND (Wikipedia OR "Web 2.0" OR "user-generated content") AND (evaluation OR evaluating OR assessment OR assessing OR measurement OR measuring OR aspects OR criteria OR factors)

Searching was performed on 15 January 2020, and the results were entered into a scientific reference manager (Endnote) for the automatic removal of duplicates. In the end, 406 documents were considered as the population.

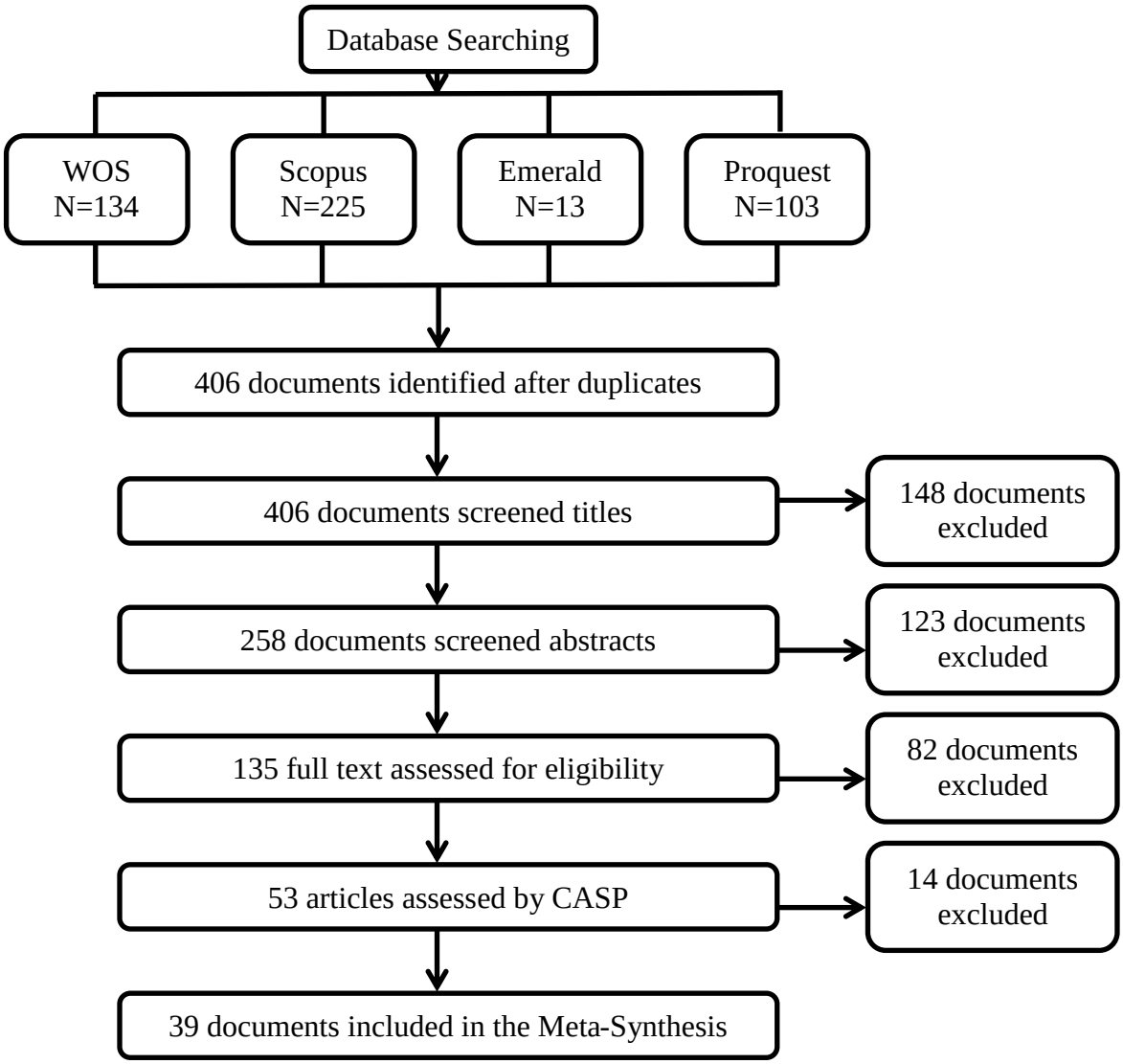

Figure 1: Flowchart for selecting studies

For selecting documents, the relationship between the retrieved documents and the research objectives was reviewed by screening the title, abstract, and content to determine whether the manuscript pertains to dimensions, criteria, and assessment procedures of information quality. At each stage, the documents that were not relevant were removed from the research process.

Then, to evaluate the methodological quality of the documents, we chose the Critical Appraisal Skills Program (CASP)[3], the most frequently used instrument (Fadahunsi et al., 2019; Keshavarz, Esmaeili Givi & Norouzi, 2020; Ludvigsen, Hall, Meyer, Fegran, Aagaard & Uhrenfeldt, 2016). This program is an indicator that helps the researcher determine the accuracy, validity, and importance of qualitative studies. This tool presents 10 questions that deal with some of the principles or assumptions that characterize qualitative research, including the aims of the research, methodology, research design, sampling, data collection, reflexivity (research partnership elations), ethical issues, data analysis, and the findings and value of the research. Based on the 50-point Critical Appraisal Skills Program, scores are considered for documents in this way: excellent article: 41-50, very good article: 31-40, good article: 21-30, average article: 11-20, and poor article: 0-10. In order to use the highest quality documents, the documents with fewer than 30 points were excluded. Finally, 39 documents were selected, and their bibliographic information can be seen in appendix A. The highest and lowest scores of documents are 50 and 32, respectively. For the sake of clarity, the selection process was also presented in a flowchart. (Fig. 1).

To extract information: At first, each document was read carefully, and then a coding process was performed line-by-line by the open coding method. The coding process was done with MAXQDA 11 (See appendix B and C for more details). In total, 48 features were identified (appendix D)

For analyzing and synthesizing findings, each feature is categorized, taking into account the concept of each one. Since the findings of this step are the basis for designing the framework, they must be categorized with special care and attention. Grouping findings according to their topics will help us see how the topical groups are related to each other, while merging the findings allows us to differentiate the prevailing relationships between findings.

Results

The present study used an open coding method to analyze and synthesize the findings.. The identified features in nine dimensions include believability, popularity, credibility, informativeness, timeliness, completeness, structure, consistency, and understandability. The dimensions were then classified into three categories based on Wang and Strong's model (1996). Because it is one of the complete models of quality dimension classification that some studies have used (Katerattanakul & Siau, 1999; Klein, 2002; Leite et al., 2016; Stvilia et al., 2007). Although it was initially developed in traditional information systems, it has also been successfully applied to information on the World Wide Web (Klein, 2002). On the other hand, since the dimensions of this model have been identified from the point of view of consumers, it allows information systems managers to better meet and respond to users' information quality needs (Katerattanakul & Siau, 1999). All stages, steps, and dimensions, together with their code, are shown in Table 2.

Table 2
Categorization of identified features

Categories	dimensions	Codes	References
Intrinsic IQ	Believability	Number of images, Type of editors and Number of references	(Bassani & Viviani, 2019b; Friberg & Reinhardt, 2009; W. Lewoniewski, 2019; Włodzimierz Lewoniewski, Węcel & Abramowicz, 2019; Liu & Ram, 2018; Yang & Chao, 2018).
	Popularity	Page rank, Citations, Page view, Number of Translated version, Number of links, Type of editors	(Bassani & Viviani, 2019a, 2019b; Dalip, André Gonçalves, Cristo & Calado, 2009; Dang & Ignat, 2016; de La Robertie, Pitarch & Teste, 2017; Hammwöhner, 2007; Włodzimierz Lewoniewski et al., 2019; Lu et al., 2013; Nakatani, Jatowt, Ohshima, & Tanaka, 2009; Saengthongpattana, Supnithi, & Soonthornphisaj, 2018; Stvilia, 2006; Stvilia et al., 2005; P. Wang & Li, 2020)
	Credibility	Number of Citations, Number of links, Type of editors and Number of references	(Chevalier, Huot, & Fekete, 2010; Dalip et al., 2017; Dang & Ignat, 2016; de La Robertie, Pitarch & Teste, 2015; de La Robertie et al., 2017; Dondio & Barrett, 2007; G. Haralabopoulos, I. Anagnostopoulos & S. Zeadally, 2016; Giannis Haralabopoulos, Ioannis Anagnostopoulos, & Sherali Zeadally, 2016; Lederman, Fan, Smith & Chang, 2014; W. Lewoniewski, 2019; Włodzimierz Lewoniewski et al., 2019; McGuinness, Zeng, Da Silva, Ding, Narayanan & Bhaowal, 2006; Stvilia, 2006; Warncke-Wang et al., 2013; Yaari et al., 2011; Zhang, Sun, & Xie, 2015)
Contextual IQ	Informativeness	Information noise, infobox, Number of categories, and Number of images	(Cozza, Petrocchi & Spognardi, 2016; Fidler & Lavbič, 2017; Hammwöhner, 2007; Lex et al., 2012; Stvilia, 2006; Stvilia et al., 2005; Warncke-Wang et al., 2013)
	Timeliness	Date of last modification, Age of article, and Number of broken links	(Dalip et al., 2017; Friberg & Reinhardt, 2009; Schmidt & Zangerle, 2019; Stvilia, 2006; Stvilia et al., 2005; Zhang et al., 2015)
	Completeness	Number of references, Number of edits, Number of links, Number of images, Length of the article, and Number of broken	(Betancourt et al., 2016; Chevalier et al., 2010; Cozza et al., 2016; Dalip et al., 2009; Dalip et al., 2017; Friberg & Reinhardt, 2009; W. Lewoniewski, 2019; Włodzimierz Lewoniewski et al., 2019; Saengthongpattana et al., 2018; Stvilia, 2006; Stvilia et al., 2005;

Categories	dimensions	Codes	References
		links	Warncke-Wang et al., 2013)
Presentational IQ	Structured	Number of sections, Number of categories, and Infobox	(Bassani & Viviani, 2019a; Cozza et al., 2016; Dalip et al., 2017; Friberg & Reinhardt, 2009; Halfaker, Kittur, Kraut, & Riedl, 2009; Kane & Ransbotham, 2012; Lukyanenko, Parsons & Wiersma, 2014; Schmidt & Zangerle, 2019)
	Consistency	Age of article, Number of reverts, Number of edits and Number of discussions in last three mouths	(Arazy, Nov, Patterson & Yeo, 2011; Bassani & Viviani, 2019a, 2019b; Chevalier et al., 2010; Dalip et al., 2011; Dondio & Barrett, 2007; Stvilia, 2006; Stvilia et al., 2005; Stvilia, Twidale, Smith, & Gasser, 2008; Viégas, Wattenberg, & Dave, 2004; P. Wang, Li, & Wu, 2021; Warncke-Wang et al., 2013)
	Understandability	Readability, Number of images, and Number of edits	(Bassani & Viviani, 2019a, 2019b; Dalip et al., 2009; Dalip et al., 2017; Friberg & Reinhardt, 2009; G. Haralabopoulos et al., 2016; Kane & Ransbotham, 2012; W. Lewoniewski, 2019; Stvilia, 2006; Stvilia et al., 2005; P. Wang & Li, 2020; P. Wang et al., 2019; Zhang et al., 2015)

Validation and credibility of the framework were investigated. The obtained framework has a relative validity because it is based on previous studies whose contents have been carefully examined. However, to determine the extent of the agreement on the components of the obtained framework, ensure their accuracy and bring the framework closer to the scientific and fundamental perspective, we decided to validate it separately. In the present study, a survey of five experts in LIS education in Iran was conducted using the Kappa method to validate the proposed framework.

The results of calculating the Kappa coefficient related to this process using SPSS software are presented in Table 3. The measurement of the achieved agreement regarding the components of the research model was calculated to be over 0.65 with the standard error of 0.1 and the Sig of 000. Thus, the validity of the proposed framework was confirmed with a good measurement of the significance agreement (higher than 0.6). The highest and lowest agreement belongs to understandability (0.74) and popularity (0.52), respectively.

Table 3
Results of the kappa probabilistic test

	Items	The calculated alignment					
	Valid	Missing	Measurement of agreement	Standard error	Sig		
	Count	Percent	Count	Percent			
Total	54	100%	0	0	0.654	0.1	000
Believability	4	100%	0	0	0.68	0.1	000
Popularity	12	100%	0	0	0.52	0.1	000
Credibility	6	100%	0	0	0.59	0.1	000
Informativeness	5	100%	0	0	0.58	0.1	00
Timeliness	5	100%	0	0	0.74	0.1	00
Completeness	7	100%	0	0	0.73	0.1	00
Structured	4	100%	0	0	0.68	0.1	00
Consistency	8	100%	0	0	0.62	0.1	00
Understandability	3	100%	0	0	0.74	0.1	00

Based on the findings, the conceptual framework for assessment of the information quality in Wikipedia was developed, as shown in Figure 2.

According to the proposed framework, information quality can be automatically evaluated according to three aspects - intrinsic, contextual, and representational - as follows:

1. Intrinsic IQ: In this category, the quality of information can be assessed in terms of credibility, believability, and popularity.
2. Contextual IQ: In this category, the quality of information can be evaluated in terms of completeness, timeliness, and informativeness (value-added).
3. Presentational IQ: In this category, the ability to understand, structure, and consistency of information can be evaluated.

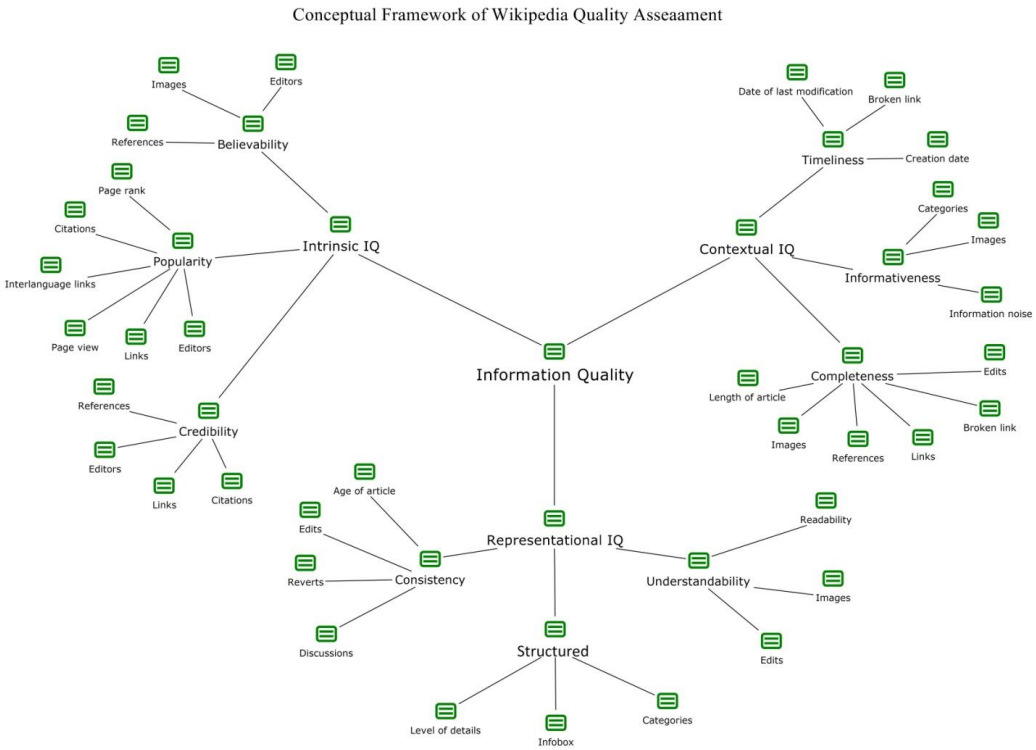

Figure 2: Framework of the information quality of Wikipedia

Discussion

The primary purpose of the current study was to identify the basic features and dimensions of Wikipedia articles to develop a conceptual framework for automatically assessing information quality. Nine dimensions of information quality relating to features of Wikipedia articles were identified. These were classified into three categories: Intrinsic, Contextual, and Presentational, based on the proposed framework of Wang and Strong (1996). However, they classified the dimensions of information quality into four categories, but as Agarwal and Yiliyasi (2010) demonstrated, the fourth category, accessibility, is not considered in social networks.

- The intrinsic IQ reflects the inherent nature of the data. It means the data itself know data quality. Believability or objectivity, popularity or authority, and credibility or reliability were classified in this category.

Believability: Experts believe there is no significant difference between objectivity and believability (Yang & Chao, 2018). So, we merged these two concepts, and the new definition of believability includes objectivity. Objectivity may be defined as the extent to which information is fair and impartial in a manner (Lewoniewski, 2019; Yang & Chao, 2018) and whether a clear and realistic description of the facts, without deforming their form by personal feelings or other effects, is presented (Fidler & Lavbič, 2017). Believability can be expected from an article that experienced authors/editors have contributed to, because an editor who contributes to different articles is likely to have experience in presenting different points of view and can therefore contribute to the article without bringing bias to it. Thus, it can be concluded that the type of editors involved in an article have an impact on its believability (Bassani & Viviani, 2019b; Friberg & Reinhardt, 2009; Lewoniewski, 2019; Włodzimierz Lewoniewski et al., 2019; Liu & Ram, 2018). Studies have also shown that the number of

resources (Friberg & Reinhardt, 2009; Lewoniewski, 2019) and the number of images (Lewoniewski, 2019; Włodzimierz Lewoniewski et al., 2019) in an article can help assess its objectivity.

Popularity: Studies have shown that the popularity of an article is influenced by the type of editors who have been involved with it (Dang & Ignat, 2016; de La Robertie et al., 2017; Włodzimierz Lewoniewski et al., 2019; Saengthongpattana et al., 2018; Stvilia, 2006; Stvilia et al., 2005). Inter-language links[4] of an article are also a good indicator of the reputation and authority of the article because popular articles have significantly more inter-language links than other articles. In other words, high-quality articles are more likely to be translated because of their popularity (Dalip et al., 2009; Hammwöhner, 2007; Włodzimierz Lewoniewski et al., 2019; Lu et al., 2013). Also, if other web pages link a page, it can mean that the page is popular. So, page rank can be used to determine the popularity of a page (Dalip et al., 2009; Stvilia, 2006; P. Wang & Li, 2020). In addition, the number of page visits (Włodzimierz Lewoniewski et al., 2019; Stvilia, 2006; Stvilia et al., 2005), links (Nakatani et al., 2009; Stvilia et al., 2005), and citations (Bassani & Viviani, 2019a, 2019b; Dalip et al., 2009; Stvilia, 2006; Stvilia et al., 2005) affect the popularity of the Wikipedia articles.

Credibility: The references allow the verification of the content. References help the readers of a text to check where the information comes from and whether the appropriate number of references supports the content. Thus, it is one of the most critical factors in assessing the reliability and credibility of an article (Dang & Ignat, 2016; de La Robertie et al., 2015; Giannis Haralabopoulos et al., 2016; Lederman et al., 2014; Lewoniewski, 2019; Włodzimierz Lewoniewski et al., 2019; McGuinness et al., 2006; Warncke-Wang et al., 2013; Yaari et al., 2011; Zhang et al., 2015). It also seems that an article containing internal and external links is more credible because the links invite readers to check the accuracy of the content to make users feel more secure (Chevalier et al., 2010; Dang & Ignat, 2016; Yaari et al., 2011). The number of citations can also be a sign of the reliability and credibility of an information source, because by citing an article, the author is indicating trust in the article (Dalip et al., 2017; G. Haralabopoulos et al., 2016; McGuinness et al., 2006). And the last feature that affects the reliability of Wikipedia is the type of authors who have contributed to the article (Dalip et al., 2017; de La Robertie et al., 2015, 2017; Dondio & Barrett, 2007; G. Haralabopoulos et al., 2016; Lederman et al., 2014; McGuinness et al., 2006; Stvilia, 2006; Zhang et al., 2015).

- The contextual IQ emphasizes the requirement that information quality must be considered within the context. Informativeness or value-added, timeliness, and completeness or comprehensiveness fall into this category.

Informativeness: Informativeness may be defined as the amount of information in an information object and is calculated as a ratio of the size of the informative content (after tokenizing and removing markup labels and stop words) to the overall size of an information object (Lex et al., 2012; Stvilia, 2006; Stvilia et al., 2005; Warncke-Wang et al., 2013). Two other features that can be value-added for an article are the number of pictures (Fidler & Lavbič, 2017; Stvilia et al., 2005; Warncke-Wang et al., 2013) and categories (Cozza et al., 2016; Hammwöhner, 2007). Categories refer to the tagging of Wikipedia articles by the Media Wiki categories and show the category to which a page belongs (Cozza et al., 2016), improving the quality of its retrieval (Hammwöhner, 2007).

T**imeliness:** The date of the last modification is a potential indicator for this aspect. We can assume the content is relatively up-to-date if the last change happened only a short time ago

(Dalip et al., 2017; Friberg & Reinhardt, 2009; Schmidt & Zangerle, 2019; Stvilia, 2006; Stvilia et al., 2005; Zhang et al., 2015). Timeliness is also affected by the creation date of the article (Friberg & Reinhardt, 2009; Stvilia, 2006; Stvilia et al., 2005; Zhang et al., 2015), together with outdated or unreachable links (Friberg & Reinhardt, 2009).

Completeness: The word count of an article is a straightforward metric that gives a good insight into the amount of information contained. Thus, the length of the article can be considered for assessing the completeness (Betancourt et al., 2016; Chevalier et al., 2010; Dalip et al., 2009; Dalip et al., 2017; Friberg & Reinhardt, 2009; Lewoniewski, 2019; Stvilia, 2006; Stvilia et al., 2005). The number of images can also be used to evaluate the comprehensiveness, as it seems that a complete article on a specific topic should include some images (Friberg & Reinhardt, 2009; Włodzimierz Lewoniewski et al., 2019; Saengthongpattana et al., 2018). Another feature is that links in the article makes further readings possible and increase the comprehensiveness of an article (Betancourt et al., 2016; Cozza et al., 2016; Friberg & Reinhardt, 2009; Saengthongpattana et al., 2018; Stvilia et al., 2005; Warncke-Wang et al., 2013). Conversely, broken links in an article will have a negative effect on this item (Betancourt et al., 2016; Cozza et al., 2016; Stvilia et al., 2005; Warncke-Wang et al., 2013).

- The presentational IQ highlights the importance of the presentation and the use of information. It includes structured or structure, consistency or stability, and understandability or comprehensibility.

Structured: According to Wikipedia standards, good articles should be organized and detailed adequately. In other words, in a structured article, the content should be organized into sections that can logically be included in an article (Bassani & Viviani, 2019a; Dalip et al., 2017; Friberg & Reinhardt, 2009; Halfaker et al., 2009; Kane & Ransbotham, 2012; Schmidt & Zangerle, 2019). The other structural elements of a page are the number of categories (Cozza et al., 2016; Lukyanenko et al., 2014) and info-boxes in an article (Schmidt & Zangerle, 2019). An Infobox is a table usually located in the top right-hand corner of articles and presents a summary of the aspects that the article contains.

Consistency: It is difficult to measure the stability or consistency of Wikipedia articles. First of all, it is very time-consuming to examine the difference between various versions of an article. On the other hand, it is not possible to accurately distinguish between changing an entire sentence and adding a single comma (Viégas et al., 2004). It seems that the number of edits is a good indicator of the stability of an article. Supposedly, mature articles are less edited and do not change widely, and there is not much discussion about them. Therefore, the number of edits and discussions in previous months can be considered to determine the consistency of an article (Bassani & Viviani, 2019a, 2019b; Chevalier et al., 2010; Dalip et al., 2011; Dondio & Barrett, 2007; Viégas et al., 2004; P. Wang et al., 2019; Warncke-Wang et al., 2013). The ratio between the age of an article and the number of edits can also be used to verify the average length of time for which an article is stable (Arazy et al., 2011; Dalip et al., 2011; Dondio & Barrett, 2007; Stvilia et al., 2008). Another feature is the number of reverts in an article. Reverting is where the last edit is reversed, and/or the effects of one or more edits are undone. It is a term that is generally reversed for vandalism or other disruptive edits. The number of reverts indicates the volatility of an article, and, consequently, a low number of reverts can be seen to demonstrate consistency (Stvilia, 2006; Stvilia et al., 2005).

Understandability: The minimum scholastic level for understanding the content of a given article is determined by measuring the readability. In other words, good articles are well written,

understandable, and free of unnecessary complexity (Bassani & Viviani, 2019a, 2019b; Dalip et al., 2017; Friberg & Reinhardt, 2009; G. Haralabopoulos et al., 2016; Kane & Ransbotham, 2012; Lewoniewski, 2019; Stvilia, 2006; Stvilia et al., 2005; P. Wang & Li, 2020; Zhang et al., 2015). In addition, further editing will make the text more comprehensible and legible (Dalip et al., 2017; P. Wang et al., 2019), while images will help to make the content clearer and more understandable (Dalip et al., 2009; Friberg & Reinhardt, 2009).

Overall, the framework of information quality assessment resulting from the meta-synthesis has two characteristics as follows:

First, the framework is based on the source's perspective. The proposed framework is a product based on the characteristics of the information source. It includes data quality dimensions that allow the end-user to evaluate the quality of information entity based on its characteristics. Considering that user-based assessment requires an expert in the subject area and a significant amount of time to determine the quality of information, a product-based assessment could solve this problem. In other words, in the age of information explosion, the existence of a framework that could assist in assessing quality independently of the specialists has a significant time-saving impact on the non-specialist user when identifying the quality of information.

Second, the proposed framework contains all of the quality dimensions and features identified in the 39 previous studies that have been reviewed in the present study, which to the best of our knowledge, covers all the quality dimensions contained in previous frameworks.

Conclusion

In this paper we applied a meta-synthesis method to propose a framework for IQ assessment. The framework is based on previous research studies and thus features powerful interpretation and clarifies greater generalizability.

The proposed framework contributes to general knowledge in information quality assessment and has advantages for Wikipedia consumers and contributors.

This framework is proposed based on objective criteria assessment; the end-user can evaluate Wikipedia articles in any subject without consulting a specialist or other human agent.

The end-user can evaluate Wikipedia articles at three levels and from nine aspects, using the framework to identify high-quality information. The use of high-quality information leads to more accurate decisions which ultimately benefit individuals and society. Furthermore, this framework identifies features that increase or decrease the quality of Wikipedia article information, and editors and contributors of Wikipedia can improve the quality of its articles by considering these features. According to studies completed to date, Wikipedia is one of the most widely used information sources-improving its information quality is therefore of considerable importance.

In addition, the developed conceptual framework based on previous studies can be used as a basis for future research, enabling the building of tools and the creation of practical guidelines for automatically assessing the quality of information. In other words, the most important achievement of this research lies in listing the components that can be used to evaluate different aspects of Wikipedia information quality. This research can also be considered fundamental knowledge for preparing appropriate tools for evaluating and measuring the quality of the information in other studies.

We are planning to test the developed framework empirically on a representative sample of

Wikipedia content in future work. It is hoped that in this way, more research results can be publicized.

Endnotes

1. https://www.alexa.com/topsites
2. The philosophical theory of signs
2. A systematic review and meta-ethnography identify how effective, cost-effective, accessible and acceptable self-management support is.
2. This feature indicates the number of versions of an article in other languages.

References

Agarwal, N. & Yiliyasi, Y. (December 2010). Information quality challenges in social media. Paper presented at *the International Conference on Information Quality*. *(ICIQ)*, Retrieved from http://mitiq.mit.edu/ICIQ/Documents/IQ%20Conference%202010/Papers/3A1_IQChallengesInSocialMedia.pdf

Arazy, O., Nov, O., Patterson, R. & Yeo, L. (2011). Information quality in Wikipedia: The effects of group composition and task conflict. *Journal of management information systems*, 27(4), 71-98. https://doi.org/10.2753/MIS0742-1222270403

Bassani, E. & Viviani, M. (April 2019a). Automatically assessing the quality of Wikipedia contents. Paper presented at *the Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*. Retrived from https://dl.acm.org/doi/pdf/10.1145/3297280.3297357

Bassani, E., & Viviani, M. (September 2019b). Quality of Wikipedia articles: Analyzing features and building a ground truth for supervised classification. Paper presented at *the Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*. Vienna, Austria, Retrieved from https://www.scitepress.org/Link.aspx?doi=10.5220/0008149303380346

Betancourt, G. G., Segnini, A., Trabuco, C., Rezgui, A. & Jullien, N. (Agust 2016). Mining team characteristics to predict Wikipedia article quality. Paper presented at *the OpenSym 2016: 12th International Symposium on Open Collaboration*. Berlin, Germany. pp.1-9. ffhal-01354368f. Retrieved from https://hal.archives-ouvertes.fr/hal-01354368/file/mining-team-characteristics%20vf.pdf

Blumenstock, J. E. (2008a). Automatically assessing the quality of Wikipedia articles. *UC Berkeley: School of Information*. Retrieved from https://escholarship.org/uc/item/18s3z11b

Blumenstock, J. E. (April 2008b). Size matters: word count as a measure of quality on wikipedia. Paper presented at *the Proceedings of the 17th international conference on World Wide Web.*, Beijing, China, Retrieved from http://www.jblumenstock.com/files/papers/jblumenstock_www08.pdf

Branger, C., O'Connell, M. E. & Peacock, S. (2018). Protocol for a meta-integration: investigating positive aspects of caregiving in dementia. *BMJ open*, 8(7), e021215. http://dx.doi.org/10.1136/bmjopen-2017-021215

Chevalier, F., Huot, S. & Fekete, J. D. (2010). WikipediaViz: Conveying article quality for casual Wikipedia readers. In *2010 IEEE Pacific Visualization Symposium (PacificVis)*, (PP.49-56). https://doi.org/10.1145/2957792.2971802

Cozza, V., Petrocchi, M. & Spognardi, A. (2016). A matter of words: NLP for quality

evaluation of Wikipedia medical articles. In: Bozzon A., Cudre-Maroux P., Pautasso C. (eds), *Web Engineering. ICWE 2016. Lecture Notes in Computer Science*, vol 9671. Springer, Cham, pp.448-456. https://doi.org/10.1007/978-3-319-38791-8_31

Dalip, D. H., Gonçalves, M. A., Cristo, M. & Calado, P. (June 2009). Automatic quality assessment of content created collaboratively by web communities: A case study of wikipedia. Paper presented at *the Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, Austin TX, USA. retrieved from https://dl.acm.org/doi/pdf/10.1145/1555400.1555449

Dalip, D. H., Gonçalves, M. A., Cristo, M. & Calado, P. (2017). A general multiview framework for assessing the quality of collaboratively created content on web 2.0. *Journal of the Association for Information Science and Technology*, 68(2), 286-308. https://doi.org/10.1002/asi.23650

Dalip, D. H., Santos, R. L., Oliveira, D. R., Amaral, V. F., Gonçalves, M. A., Prates, R. O., . . . De Almeida, J. M. (June 2011). GreenWiki: A tool to support users' assessment of the quality of Wikipedia articles. Paper presented at *the Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, Ottawa, Ontario, Canada. Retrieved from https://dl.acm.org/doi/pdf/10.1145/1998076.1998190

Dang, Q. V. & Ignat, C. L. (November 2016). Measuring quality of collaboratively edited documents: The case of Wikipedia. Paper presented at *the Proceedings - 2016 IEEE 2nd International Conference on Collaboration and Internet Computing, IEEE CIC*, Pittsburgh, PA, USA. Retrived from https://www.archives-ouvertes.fr/hal-01388614/document

de La Robertie, B., Pitarch, Y. & Teste, O. (2015). Measuring article quality in wikipedia using the collaboration network. Paper presented at *the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM).* (PP. 464-471).

de La Robertie, B., Pitarch, Y. & Teste, O. (2017). Structure-Based Features for Predicting the Quality of Articles in Wikipedia.In: Kawash J., Agarwal N., Özyer T. (eds), 115-140, *Prediction and Inference from Social Networks and Social Media. Lecture Notes in Social Networks*. Springer, Cham. https://doi.org/10.1007/978-3-319-51049-1_6

Dondio, P. & Barrett, S. (2007). Computational trust in web content quality: A comparative evalutation on the Wikipedia project. *Informatica (Ljubljana)*, 31(2), 151-160. https://doi.org/10.21427/D70W52

Fadahunsi, K. P., Akinlua, J. T., O'Connor, S., Wark, P. A., Gallagher, J., Carroll, C., . . . O'Donoghue, J. (2019). Protocol for a systematic review and qualitative synthesis of information quality frameworks in eHealth. *BMJ open*, 9(3), e024722. http://dx.doi.org/10.1136/bmjopen-2018-024722

Fidler, M. & Lavbič, D. (2017). Improving information quality of Wikipedia articles with cooperative principle. *Online information review*, 41(6), 797-811. DOI: 10.1108/OIR-01-2016-0003

Finfgeld, D. L. (2003). Metasynthesis: The state of the art-so far. *Qualitative Health Research*, 13(7), 893-904. https://doi.org/10.1177/1049732303253462

Friberg, T. & Reinhardt, W. (2009). An empirical study on criteria for assessing information quality in corporate wikis. In: Bowen, P., Elmagarmid,A.K., Österele, H., Sattler, K. (eds), *Proceedings of the 2009 International Conference on Information Quality, ICIQ 2009*,

Hasso Plattner Institute, University of Potsdam, Germany, pp.217-230. Retrieved from http://mitiq.mit.edu/ICIQ/Documents/IQ%20Conference%202009/Papers/7-A.pdf

Halfaker, A., Kittur, A., Kraut, R. & Riedl, J. (October 2009). A jury of your peers: Quality, experience and ownership in Wikipedia. Paper presented at *the Proceedings of the 5th International Symposium on Wikis and Open Collaboration, WiKiSym*. Association for Computing Machinery, New Yourk, United State. Retrieved from https://dl.acm.org/doi/pdf/10.1145/1641309.1641332 (accessed 21 Aug 2019).

Hammwöhner, R. (2007). Interlingual aspects of wikipedia's quality. In: Robbert, M.A., O'Hare, R., Markus, L., Klein, B.D. (eds), *Proceedings of the 12th International Conference on Information Quality*, MIT, Cambridge, USA, (PP. 477-488). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=AE80FE32D4DFCDB5ABED10902B69CCD3?doi=10.1.1.93.1943&rep=rep1&type=pdf

Han, W.-M. (2018). Evaluating perceived and estimated data quality for Web 2.0 applications: a gap analysis. *Software Quality Journal*, 26(2), 367-383. https://doi.org/10.1007/s11219-017-9365-7

Hansen, H. P., Draborg, E. & Kristensen, F. B. (2011). Exploring qualitative research synthesis. *The Patient: Patient-Centered Outcomes Research*, 4(3), 143-152. https://doi.org/10.2165/11539880-000000000-00000

Haralabopoulos, G., Anagnostopoulos, I. & Zeadally, S. (2016). The challenge of improving credibility of user-generated content in online social networks. *Journal of Data and Information Quality (JDIQ)*, 7(3), 13, 1-4. https://doi.org/10.1145/2899003

Kane, G., & Ransbotham, S. (December 2012). Codification and collaboration: Information quality in social media. Paper presented at *the Proceedings of 33th International Conference on Information Systems (ICIS)*, Orlando, Florida, USA. Retrieved from https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1012&context=icis2012

Katerattanakul, P. & Siau, K. (December 1999). Measuring information quality of web sites: Development of an instrument. *ICIS '99: Proceedings of the 20th international conference on Information Systems*, (PP. 279-285). Association for Information Systems, Atlanta, GA, United State.

Keshavarz, H., Esmaeili Givi, M. & Norouzi, Y. (2020). Credibility evaluation of scientific information on websites: Designing and evaluating an exploratory model. *Journal of Librarianship and Information Science*, 52(4), 1086-1101. https://doi.org/10.1177/0961000620903103

Kim, J. Y., Gudewicz, T. M., Dighe, A. S. & Gilbertson, J. R. (2010). The pathology informatics curriculum wiki: Harnessing the power of user-generated content. *Journal of Pathology Informatics*, 13(1). https://doi.org/10.4103/2153-3539.65428

Klein, B. (2002). When do users detect information quality problems on the World Wide Web? *AMCIS 2002 Proceedings*, 152. Retrieved from https://aisel.aisnet.org/amcis2002/152

Laurent, M. R., & Vickers, T. J. (2009). Seeking Health Information Online: Does Wikipedia Matter? *Journal of the American Medical Informatics Association, 16*(4), 471-479. https://doi.org/10.1197/jamia.M3059

Lederman, R., Fan, H., Smith, S. & Chang, S. (2014). Who can you trust? Credibility assessment in online health forums. *Health Policy and Technology, 3*(1), 13-25. https://doi.org/10.1016/j.hlpt.2013.11.003

Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: a methodology for information quality assessment. *Information & Management,* 40(2), 133-146. https://doi.org/10.1016/S0378-7206(02)00043-5

Leite, P., Gonçalves, J., Teixeira, P. & Rocha, Á. (2016). A model for the evaluation of data quality in health unit websites. *Health Informatics Journal*, 22(3), 479-495. https://doi.org/.1177/1460458214567003

Lewoniewski, W. (2019). Measures for quality assessment of articles and infoboxes in multilingual wikipedia.In Abramowicz W., Paschke A. (eds) *Business Information Systems Workshops. BIS 2018. Lecture Notes in Business Information Processing*, vol 339. Springer, Cham, pp.619-633. https://doi.org/10.1007/978-3-030-04849-5_53

Lewoniewski, W., Khairova, N., Węcel, K., Stratiienko, N. & Abramowicz, W. (2017). Using morphological and semantic features for the quality assessment of Russian Wikipedia. In R. Damaševičius & V. Mikašytė (Eds.), *Information and Software Technologies. ICIST. Communications in Computer and Information Science,* vol. 1756, Springer, Cham. pp.550-560. https://doi.org/10.1007/978-3-319-67642-5_46

Lewoniewski, W., & Węcel, K. (2017). Relative quality assessment of Wikipedia articles in different languages using synthetic measure. In W. Abramowicz (Ed.), *Business Information Systems Workshops. BIS. Lecture Notes in Business Information Processing* , vol. 303, Springer, Cham. pp.282-292. https://doi.org/10.1007/978-3-319-69023-0_24

Lewoniewski, W., Węcel, K. & Abramowicz, W. (2017). Relative Quality and Popularity Evaluation of Multilingual Wikipedia Articles. *Informatics,* 4(4), 43. https://doi/org/10.3390/informatics4040043

Lewoniewski, W., Węcel, K. & Abramowicz, W. (2019). Multilingual Ranking of Wikipedia Articles with Quality and Popularity Assessment in Different Topics. *Computers*, 8(3), 60. https://doi.org/10.3390/computers8030060

Lex, E., Voelske, M., Errecalde, M., Ferretti, E., Cagnina, L., Horn, C., . . . Granitzer, M. (October 2012). Measuring the quality of web content using factual information. Paper presented at *the Proceedings of the 2nd joint WICOW/AIRWeb workshop on web quality*, Madrid, Spain, retrieved from https://dl.acm.org/doi/pdf/10.1145/2184305.2184308

Liu, J. & Ram, S. (2018). Using big data and network analysis to understand Wikipedia article quality. *Data and Knowledge Engineering,* 115, 80-93. https://doi.org/10.1016/j.datak.2018.02.004

Lu, Y., Zhang, L. & Li, J. (2013) Evaluating article quality and editor reputation in Wikipedia. In: Qi G., Tang J., Du J., Pan J.Z., Yu Y. (eds), *Linked Data and Knowledge Graph. CSWS 2013. Communications in Computer and Information Science*, vol 406. Springer, Berlin, Heidelberg, pp.215-227. https://doi.org/10.1007/978-3-642-54025-7_19

Ludvigsen, M. S., Hall, E. O., Meyer, G., Fegran, L., Aagaard, H. & Uhrenfeldt, L. (2016). Using Sandelowski and Barroso's meta-synthesis method in advancing qualitative evidence. *Qualitative Health Research,* 26(3), 320-329. https://doi.org/10.1177/1049732315576493

Lukyanenko, R., Parsons, J. & Wiersma, Y. F. (2014). The IQ of the crowd: Understanding and improving information quality in structured user-generated content. *Information Systems Research,* 25(4), 669-689. https://doi.org/10.1287/isre.2014.0537

McGuinness, D. L., Zeng, H., Da Silva, P. P., Ding, L., Narayanan, D. & Bhaowal, M. (May 2006). Investigations into trust for collaborative information repositories: A wikipedia case study. Paper presented at *the Proceedings of the Workshop on Models of Trust for the Web*, Edinburgh, Scotland, UK, Retrieved from https://ebiquity.umbc.edu/_file_directory_/papers/274.pdf

Nakatani, M., Jatowt, A., Ohshima, H. & Tanaka, K. (2009). Quality evaluation of search results by typicality and speciality of terms extracted from wikipedia.In Zhou X., Yokota H., Deng K., Liu Q. (eds.), *Database Systems for Advanced Applications. DASFAA 2009. Lecture Notes in Computer Science*, vol 5463. Springer, Berlin, Heidelberg, pp.570-584. https://doi.org/10.1007/978-3-642-00887-0_50

Naumann, F. & Rolker, C. (2005). *Assessment methods for information quality criteria*. Berlin: Mathematisch-Naturwissenschaftliche Fakultät II, Institut für Informatik.

Orehovacki, T. (2010). Proposal for a set of quality attributes relevant for Web 2.0 application success. In V. Luzar-Stiffler, I. Jarec & Z. Bekic (eds.), *Proceedings of the ITI 2010, 32nd International Conference on Information Technology Interfaces* , University Computing Centre, Zagreb, pp.319-326.

Price, R. & Shanks, G. (2005). A semiotic information quality framework : development and comparative analysis. *Journal of Information Technology,* 20(2), 88-102. https://doi.org/10.1057/palgrave.jit.2000038

Saengthongpattana, K., Supnithi, T. & Soonthornphisaj, N. (November 2018). Quality Classification of ASEAN Wikipedia Articles using Statistical Features. Paper presented at *the 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing, iSAI-NLP*, Pattaya, Thailand. Retrieved from https://saki.siit.tu.ac.th/isai-nlp2018/uploads_final/51__0ce6368ae89c4b419ff2d92827ef2d42/FinalVersion-1.pdf

Sandelowski, M. & Barroso, J. (2006). *Handbook for synthesizing qualitative research*. springer publishing company.

Schmidt, M. & Zangerle, E. (2019). Article quality classification on Wikipedia: Introducing document embeddings and content features. Paper presented at *the Proceedings of the 15th International Symposium on Open Collaboration, OpenSym*. Skövde, Sweden, retrieved from https://b2n.ir/p55859

Shahbazi, M., Farajpahlou, A., Osareh, F. & Rahimi, A. (2019). Development of a scale for data quality assessment in automated library systems. *Library & Information Science Research,* 41(1), 78-84. https://doi.org/10.1016/j.lisr.2019.02.005

Sherwood, G. (1999). Meta-synthesis: merging qualitative studies to develop nursing knowledge. *International Journal of Human Caring,* 3(1), 37-42. https://doi.org/DOI:10.20467/1091-5710.3.1.37

Stvilia, B. (2006). *Measuring information quality*. Ph. D. Thesis, University of Illinois at Urbana-Champaign, Ann Arbor. Retrived from http://hdl.handle.net/2142/81539

Stvilia, B., Gasser, L., Twidale, M. B. & Smith, L. C. (2007). A framework for information quality assessment. *Journal of the American Society for Information Science and technology,* 58(12), 1720-1733. https://doi.org/10.1002/asi.20652

Stvilia, B., Twidale, M. B., Smith, L. C. & Gasser, L. (November 2005). Assessing Information Quality of a Community-Based Encyclopedia. Paper presented at *the Proceedings of 10th International Conference on Information Quality, ICIQ*, Cambridge, MA, United States, retrieved from https://b2n.ir/z83958

Stvilia, B., Twidale, M. B., Smith, L. C. & Gasser, L. (2008). Information quality work organization in Wikipedia. *Journal of the American Society for Information Science and technology*, 59(6), 983-1001. http://dx.doi.org/10.1002/asi.20813

Tao, D., LeRouge, C., Smith, K. J. & De Leo, G. (2017). Defining information quality into health websites: A conceptual framework of health website information quality for educated young adults. *JMIR human factors*, 4(4), p.e25. https://doi.org/10.2196/humanfactors.6455

Viégas, F. B., Wattenberg, M. & Dave, K. (April 2004). Studying cooperation and conflict between authors with history flow visualizations. Paper presented at *the Proceedings of the SIGCHI conference on Human factors in computing systems*. Vienna, Austria. retrieved from https://alumni.media.mit.edu/~fviegas/papers/history_flow.pdf

Walsh, D. & Downe, S. (2005). Meta-synthesis method for qualitative research: A literature review. *Journal of Advanced Nursing*, 50(2), 204-211. https://doi.org/10.1111/j.1365-2648.2005.03380.x

Wang, P. & Li, X. (2020). Assessing the quality of information on wikipedia: A deep-learning approach. *Journal of the Association for Information Science and Technology*, 71(1), 16-28. https://doi.org/10.1002/asi.24210

Wang, P., Li, X. & Wu, R. (2021). A deep learning-based quality assessment model of collaboratively edited documents: A case study of Wikipedia. *Journal of information science*. 47(2), 176-191. https://doi.org/10.1177/0165551519877646

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of management information systems*, 12(4), 5-33. https://doi.org/10.1080/07421222.1996.11518099

Warncke-Wang, M., Cosley, D. & Riedl, J. (Aguat 2013). Tell me more: An actionable quality model for Wikipedia. Paper presented at *the Proceedings of the 9th International Symposium on Open Collaboration*. Hong Kong, China. rerieved from http://files.grouplens.org/papers/wikisym2013_warnckewang-cosley-riedl.pdf

Weitzel, L., de Oliveira, J. P. M. & Quaresma, P. (2014). Measuring the reputation in user-generated-content systems based on health information. *Procedia Computer Science*, 29, 364-378. https://doi.org/10.1016/j.procs.2014.05.033

Yaari, E., Baruchson-Arbib, S. & Bar-Ilan, J. (2011). Information quality assessment of community generated content: A user study of Wikipedia. *Journal of information science*, 37(5), 487-498. https://doi.org/10.1177/0165551511416065

Yang, H.-L. & Chao, A. F. (2018). Sentiment annotations for reviews: An information quality perspective. *Online information review*, 24(5), 579-594. https://doi.org/10.1108/OIR-04-2017-0114

Zhang, Y., Sun, Y. & Xie, B. (2015). Quality of health information for consumers on the web: A systematic review of indicators, criteria, tools, and evaluation results. *Journal of the Association for Information Science and Technology*, 66(10), 2071-2084. https://doi.org/10.1002/asi.23311

Appendix A
Bibliographic information of 39 documents were selected

1. Arazy, O., Nov, O., Patterson, R., & Yeo, L. (2011). Information quality in Wikipedia: The effects of group composition and task conflict. *Journal of management information systems, 27*(4), 71-98.
2. Bassani, E., & Viviani, M. (2019). *Automatically assessing the quality of Wikipedia contents.* Paper presented at the Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing.
3. Bassani, E., & Viviani, M. (2019). *Quality of Wikipedia articles: Analyzing features and building a ground truth for supervised classification.* Paper presented at the IC3K 2019 - Proceedings of the 11th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management.
4. Betancourt, G. G., Segnini, A., Trabuco, C., Rezgui, A., & Jullien, N. (2016). *Mining team characteristics to predict Wikipedia article quality.* Paper presented at the Proceedings of the 12th International Symposium on Open Collaboration, OpenSym 2016.
5. Chevalier, F., Huot, S., & Fekete, J. D. (2010). *WikipediaViz: Conveying article quality for casual wikipedia readers.* Paper presented at the IEEE Pacific Visualization Symposium 2010, PacificVis 2010 - Proceedings.
6. Cozza, V., Petrocchi, M., & Spognardi, A. (2016). *A matter of words: NLP for quality evaluation of Wikipedia medical articles.* Paper presented at the International Conference on Web Engineering.
7. Dalip, D. H., André Gonçalves, M., Cristo, M., & Calado, P. (2009). *Automatic quality assessment of content created collaboratively by web communities: a case study of wikipedia.* Paper presented at the Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries.
8. Dalip, D. H., Gonçalves, M. A., Cristo, M., & Calado, P. (2017). A general multiview framework for assessing the quality of collaboratively created content on web 2.0. *Journal of the Association for Information Science and Technology, 68*(2), 286-308.
9. Dalip, D. H., Santos, R. L., Oliveira, D. R., Amaral, V. F., Gonçalves, M. A., Prates, R. O., . . . De Almeida, J. M. (2011). *GreenWiki: A tool to support users' assessment of the quality of Wikipedia articles.* Paper presented at the Proceedings of the ACM/IEEE Joint Conference on Digital Libraries.
10. Dang, Q. V., & Ignat, C. L. (2016). *Measuring quality of collaboratively edited documents: The case of Wikipedia.* Paper presented at the Proceedings - 2016 IEEE 2nd International Conference on Collaboration and Internet Computing, IEEE CIC 2016.
11. de La Robertie, B., Pitarch, Y., & Teste, O. (2015). *Measuring article quality in wikipedia using the collaboration network.* Paper presented at the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM).
12. de La Robertie, B., Pitarch, Y., & Teste, O. (2017). Structure-Based Features for Predicting the Quality of Articles in Wikipedia *Prediction and Inference from Social Networks and Social Media* (pp. 115-140): Springer.
13. Dondio, P., & Barrett, S. (2007). Computational trust in web content quality: A comparative evalutation on the Wikipedia project. *Informatica (Ljubljana), 31*(2), 151-160.
14. Fidler, M., & Lavbič, D. (2017). Improving information quality of Wikipedia articles

with cooperative principle. *Online information review, 41*(6), 797-811.

15. Friberg, T., & Reinhardt, W. (2009). *An empirical study on criteria for assessing information quality in corporate wikis.* Paper presented at the Proceedings of the 2009 International Conference on Information Quality, ICIQ 2009.

16. Halfaker, A., Kittur, A., Kraut, R., & Riedl, J. (2009). *A jury of your peers: Quality, experience and ownership in Wikipedia.* Paper presented at the Proceedings of the 5th International Symposium on Wikis and Open Collaboration, WiKiSym 2009.

17. Hammwöhner, R. (2007). Interlingual aspects of wikipedia's quality.

18. Haralabopoulos, G., Anagnostopoulos, I., & Zeadally, S. (2016). The challenge of improving credibility of user-generated content in online social networks. *Journal of Data and Information Quality, 7*(3). doi: 10.1145/2899003

19. Kane, G., & Ransbotham, S. (2012, 16-19 Dec). *Codification and collaboration: Information quality in social media.* Paper presented at the the Proceedings of 33th International Conference on Information Systems (ICIS), Orlando, Florida, USA.

20. Lewoniewski, W. (2019) Measures for quality assessment of articles and infoboxes in multilingual wikipedia. *Vol. 339. Lecture Notes in Business Information Processing* (pp. 619-633).

21. Lewoniewski, W., Węcel, K., & Abramowicz, W. (2019). Multilingual ranking of wikipedia articles with quality and popularity assessment in different topics. *Computers, 8*(3). doi: 10.3390/computers8030060

22. Lex, E., Voelske, M., Errecalde, M., Ferretti, E., Cagnina, L., Horn, C., . . . Granitzer, M. (2012). Measuring the quality of web content using factual information. 7. doi: 10.1145/2184305.2184308

23. Lih, A. (2004). Wikipedia as participatory journalism: Reliable sources? metrics for evaluating collaborative media as a news resource. *Nature, 3*(1), 1-31.

24. Liu, J., & Ram, S. (2018). Using big data and network analysis to understand Wikipedia article quality. *Data and Knowledge Engineering, 115*, 80-93. doi: 10.1016/j.datak.2018.02.004

25. Lu, Y., Zhang, L., & Li, J. (2013) Evaluating article quality and editor reputation in Wikipedia. *Vol. 406. Communications in Computer and Information Science* (pp. 215-227).

26. Lukyanenko, R., Parsons, J., & Wiersma, Y. F. (2014). The IQ of the crowd: Understanding and improving information quality in structured user-generated content. *Information Systems Research, 25*(4), 669-689.

27. McGuinness, D. L., Zeng, H., Da Silva, P. P., Ding, L., Narayanan, D., & Bhaowal, M. (2006). *Investigations into trust for collaborative information repositories: A wikipedia case study.* Paper presented at the Proceedings of the Workshop on Models of Trust for the Web.

28. Nakatani, M., Jatowt, A., Ohshima, H., & Tanaka, K. (2009). *Quality evaluation of search results by typicality and speciality of terms extracted from wikipedia.* Paper presented at the International Conference on Database Systems for Advanced Applications.

29. Saengthongpattana, K., Supnithi, T., & Soonthornphisaj, N. (2018). *Quality Classification of ASEAN Wikipedia Articles using Statistical Features.* Paper presented at the 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing, iSAI-NLP 2018 - Proceedings.

30. Schmidt, M., & Zangerle, E. (2019). *Article quality classification on Wikipedia: Introducing document embeddings and content features.* Paper presented at the Proceedings of the 15th International Symposium on Open Collaboration, OpenSym 2019.

31. Stvilia, B. (2006). *Measuring information quality*. (3223727 Ph.D.), University of Illinois at Urbana-Champaign, Ann Arbor. Retrieved from https://search.proquest.com/docview/305328745?accountid=42543 ProQuest Dissertations & Theses Global database.

32. Stvilia, B., Twidale, M. B., Smith, L. C., & Gasser, L. (2005). *Assessing Information Quality of a Community-Based Encyclopedia*. Paper presented at the Proceedings of the International Conference on Information Quality. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.78.6243&rep=rep1&type=pdf

33. Stvilia, B., Twidale, M. B., Smith, L. C., & Gasser, L. (2008). Information quality work organization in Wikipedia. *Journal of the American Society for Information Science and technology, 59*(6), 983-1001.

34. Viégas, F. B., Wattenberg, M., & Dave, K. (2004). *Studying cooperation and conflict between authors with history flow visualizations*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems. https://alumni.media.mit.edu/~fviegas/papers/history_flow.pdf

35. Wang, P., & Li, X. (2020). Assessing the quality of information on wikipedia: A deep-learning approach. *Journal of the Association for Information Science and Technology, 71*(1), 16-28. doi: 10.1002/asi.24210

36. Wang, P., Li, X., & Wu, R. (2021). A deep learning-based quality assessment model of collaboratively edited documents: A case study of Wikipedia. *Journal of information science*. doi: 10.1177/0165551519877646

37. Warncke-Wang, M., Cosley, D., & Riedl, J. (2013). *Tell me more: an actionable quality model for Wikipedia*. Paper presented at the Proceedings of the 9th International Symposium on Open Collaboration. http://files.grouplens.org/papers/wikisym2013_warnckewang-cosley-riedl.pdf

38. Yaari, E., Baruchson-Arbib, S., & Bar-Ilan, J. (2011). Information quality assessment of community generated content: A user study of Wikipedia. *Journal of information science, 37*(5), 487-498. doi: 10.1177/0165551511416065

39. Zhang, Y., Sun, Y., & Xie, B. (2015). Quality of health information for consumers on the web: A systematic review of indicators, criteria, tools, and evaluation results. *Journal of the Association for Information Science and Technology, 66*(10), 2071-2084. doi: 10.1002/asi.23311

Appendix B

A sample of basic information and ratings of articles based on the CASP

Number	Doc. Type	Resource	Extracted information	Research method	Score
1	research article	Wang & Li,) (2020	1. Popularity ✓ Page rank of an article, which measures the popularity of the article 2. understandable ✓ The readability scores in Table 4 represent the grade level or education level that readers must attain to understand the text 3. Article age, citation count, citation count per text length, discussion count, image number, number of characters, number of external links, links per text length, mean time between two reviews, review count, ratio between the number of modified lines and the total number of lines and translation count affect the quality of the article .	Quantitative research	43
2	Conference paper	Bassani &) Viviani, (2019a	1. Stability and Maturity ✓ how many times and in which way the article has been modified. They can measure the degree of maturity and stability of an article, since no extensive corrections could indicate good-quality articles having reached amaturity level, while a lack of stability could indicate different kinds of controversies 2. Popularity ✓ citations can provide evidences of the popularity of the articles. In addition, a high-quality article is expected to be used as a reference point for articles dealing with interconnected topics. 3. According to theWikipedia quality standards, 4 a good article must be reasonably clear, organized adequately 4. Understandable ✓ Readability Features, indicating the degree of readability of the articles, i.e., the minimum scholastic level that is necessary to understand their contents. The intuition behind	Qualitative research	40

Number	Doc. Type	Resource	Extracted information	Research method	Score
			these features is that "good articles should be well written, understandable, and free of unnecessary complexity 5. It is supposed that a good-quality text in a mature stage is reasonably neither too short (incomplete topic coverage), nor excessively long (verbose content) 6. point to appropriate references and links		
3	Conference paper	Bassani &) Viviani, (2019b	1. Accuracy ✓ Text Features provide the best level of Accuracy 2. Objectivity ✓ the total number of unique users that have contributed to the article. More contributors an article has, more objective its content is supposed to be 3. stability and maturity ✓ how many times and in which way the article has been modified. They can measure the degree of maturity and stability of an article, ✓ since no extensive corrections could indicate good-quality articles having reached a maturity level, while a lack of stability could indicate different kinds of controversies (e.g., with respect to neutrality, correctness, etc.). ✓ Last three-months review count*: the number of reviews made in the last three months. This feature could indicate that the content of an article is controversial, the article is about evolving events or it is in the beginning of its editorial process; (28) Last three-months review rate (Dondio et al., 2006): the percentage of reviews made in the last three months; 4. Popularity ✓ Network features are extracted from the articles graph, which is built by considering citations among	Qualitative research	42

Number	Doc. Type	Resource	Extracted information	Research method	Score
			articles. These citations can provide evidences of the popularity of the articles. 5. Understandable ✓ Readability Features. They are numerical indicators of the US grade level, i.e., the comprehension level that a reader must possess to understand what is debated in a text. They were first used, to tackle the considered problem		
4	research article	Lewoniewski, Węcel, & Abramowicz, (2019	1. Objectivity ✓ Objectivity can expected from article, which was jointly created by a large number of different users. So, the most popular measure is the number of unique authors ✓ larger number of authors can positively influence the objectivity of the article ✓ Objectivity: to what extent the content of the article meets the criterion of a neutral point of view, does it contain pictures and other multimedia materials related to this article 2. Completeness ✓ One of the most popular measure for this dimension is content volume measured by articles. length Length can be measured in different ways: bytes, characters, words and others. ✓ we can also measure completeness (because articles on a specific topic should contain images) 3. credibility ✓ Therefore, one of the most commonly used measure related to credibility is number of the references in Wikipedia articles ✓ Credibility: whether the information provided can be checked with reliable sources 4. Popularity 5. understandable ✓ Readability: to what extent the text is understandable and free from	Quantitative research	44

Number	Doc. Type	Resource	Extracted information	Research method	Score
			unnecessary complexity 6. Timeliness: ✓ to what extent the article describes the current state of a certain reality (degree to which information is up-to-date). ✓ Measures of this quality dimension can be related to currency and volatility of the information		

Appendix C
A sample of coding of extracted features

code	Resource	frequency
Credibility	(B. De La Robertie, Pitarch, & Teste, 2015; Lewoniewski, 2019; McGuinness et al., 2006; Besiki Stvilia, 2006; Warncke-Wang, Cosley, & Riedl, 2013; Zhang, Sun, & Xie, 2015) (Dondio & Barrett, 2007)	7
N. Of editors	(Bassani & Viviani, 2019b; Chevalier, Huot, & Fekete, 2010; B. De La Robertie et al., 2015; Hammwöhner, 2007; Lewoniewski, 2019; Lewoniewski et al., 2019; Lih, 2004; Liu & Ram, 2018; Saengthongpattana, Supnithi, & Soonthornphisaj, 2018; Wang & Li, 2020; Wang, Li, & Wu, 2021; Yaari, Baruchson-Arbib, & Bar-Ilan, 2011) (Kane & Ransbotham, 2012) (Arazy, Nov, Patterson, & Yeo, 2011)	14
unique editors	(Daniel H Dalip, Gonçalves, Cristo, & Calado, 2017; B. De La Robertie et al., 2015; Baptiste de La Robertie, Pitarch, & Teste, 2017; Friberg & Reinhardt, 2009; Lewoniewski et al., 2019; Liu & Ram, 2018; McGuinness et al., 2006; Besiki Stvilia, 2006; B. Stvilia, Twidale, Smith, & Gasser, 2005; Wang et al., 2019; Yaari et al., 2011) (Haralabopoulos, Anagnostopoulos, & Zeadally, 2016) (Lederman, Fan, Smith, & Chang, 2014) (Besiki Stvilia, Twidale, Smith, & Gasser, 2008) (Dondio & Barrett, 2007)	15
registered editors	(Daniel H Dalip et al., 2017; Besiki Stvilia, 2006; B. Stvilia et al., 2005; Wang & Li, 2020; Wang et al., 2019; Yaari et al., 2011)	6
Consistency)this code includes: consistency, stability and maturity(	(Bassani & Viviani, 2019a, 2019b; Chevalier et al., 2010; D. H. Dalip, Gonçalves, Cristo, & Calado, 2009; Hammwöhner, 2007; Besiki Stvilia, 2006; B. Stvilia et al., 2005) (D. H. Dalip et al., 2011; Dondio & Barrett, 2007) (Besiki Stvilia et al., 2008) (Viégas, Wattenberg, & Dave, 2004) (Kane & Ransbotham, 2012) (Arazy et al., 2011)	13
N. Of categories	(Hammwöhner, 2007; Schmidt & Zangerle, 2019; Warncke-Wang et al., 2013) (Cozza, Petrocchi, & Spognardi, 2016) (Lukyanenko, Parsons, Wiersma, & Maddah, 2019)	5

length of the article	(Bassani & Viviani, 2019a; Betancourt, Segnini, Trabuco, Rezgui, & Jullien, 2016; Chevalier et al., 2010; D. H. Dalip et al., 2009; Daniel H Dalip et al., 2017; Friberg & Reinhardt, 2009; Hammwöhner, 2007; Lewoniewski, 2019; Lu, Zhang, & Li, 2013; Saengthongpattana et al., 2018; Schmidt & Zangerle, 2019; Besiki Stvilia, 2006; B. Stvilia et al., 2005; Wang & Li, 2020; Wang et al., 2019; Yaari et al., 2011) (Kane & Ransbotham, 2012)	17

Appendix D
48 featured were identified by coding process

1. Credibility
2. Number of editors
3. Unique editors
4. Number of unique editors to edits
5. Registered editors
6. Anonymous editors
7. Consistency
8. Number of categories
9. Length of article
10. Number of links
11. External links
12. Internal links
13. Broken links
14. Edits
15. Edits to age
16. Translation count
17. Popularity
18. Authority
19. Citations
20. Structured
21. Sections
22. Complexity
23. Clarity
24. Understandability
25. Readability
26. References
27. Believability
28. Objectivity
29. Number of reverts
30. Median Revert Time
31. Complete
32. Reliability
33. Validity
34. Up to date
35. Timeliness
36. Edits in a period
37. Creation date
38. Last modified date
39. Page rank
40. Image
41. Discussion
42. Page view
43. Unbiased
44. Infobox
45. Information noise
46. Informativeness
47. volatility
48. age