# ANALYSING INFORMATION QUALITY OF WIKIPEDIA ARTICLES

W.Chinthani Sugandhika Sirisoma

208108X

Degree of Master of Science

Department of Information Technology

University of Moratuwa
Sri Lanka

May 2022

# ANALYSING INFORMATION QUALITY OF WIKIPEDIA ARTICLES

W.Chinthani Sugandhika Sirisoma

208106X

Dissertation submitted in partial fulfilment of the requirements for the degree Master of Science in Information Technology

Department of Information Technology

University of Moratuwa
Sri Lanka

May 2022

## DECLARATION OF THE CANDIDATE AND SUPERVISORS

I declare that this is my own work, and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: UOM Verified Signature Date: 30/05/2022

The above candidate has carried out research for the Masters Dissertation under my supervision.

Name of the principal supervisor: Dr. (Mrs.) Supunmali Ahangama

Signatu UOM Verified Signature Date: 02/06/2022

Name of the co-supervisor Dr. Sapumal Ahangama

UOM Verified Signature

Signature Date: 07.06.2022

## DEDICATION

I dedicate this dissertation to my parents and my loving husband for always being there with me, cheering me up and standing by my side through all the good times and bad.

## ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to all those who helped me to make this research successful with all their guidance and advice. Firstly, I greatly thank my Principal Supervisor, Dr. (Mrs.) Supunmali Ahangama and Co-supervisor Dr. Sapumal Ahangama for their enormous guidance, advice and encouragement gave for me to conduct this research. Next, I would like to give my heartfelt gratitude to my panel of reviewers Dr. Amal Shehan Perera and Dr. C. R. J. Amalraj for their valuable review comments, suggestions and advice during the progress reviews. I would also thank University of Moratuwa for offering me the grant SRC/LT/2021/24 for carrying out my research successfully. Further, I would like to thank the three participants who participated in conducting the inter-rater reliability test. Lastly, I would like to thank all the academic and non-academic staff of the Department of Information Technology who helped me in diverse ways to make this research a success and a reality.

# ABSTRACT

User Generated Content (UGC) is growing in significance for information sharing along with the introduction of Web 2.0. Being one of the largest UGC databases in the world, Wikipedia also stands as the largest community-based collaborative encyclopedia ever created. However, Wikipedia's open-source and collaborative structure presents a serious information quality (IQ) concern. Malicious users take advantage of Wikipedia's popularity on the World Wide Web (WWW) when conducting malicious activities such as link spamming. Wikipedia is therefore often discouraged for use in academic-related activities and research. However. there are some high-quality articles that are both rich in information and quality. Statistical models and machine learning algorithms have been used in existing methods for determining Wikipedia's IQ. However, the outcomes of these models are not satisfactory. Therefore, in this study a novel theoretical model for evaluating IQ is presented, based on Google's E-A-T framework. The model comprises three IQ constructs Expertise, Authority and Trustworthiness. A collection of IQ dimensions that affect the aforementioned three IQ constructs as well as 45 IQ attributes to assess the IQ dimensions were identified and presented based on empirical findings and study results. A Selenium 3.14 web automation script was used to automatically and inexpensively extract the IQ attributes from Wikipedia articles' content and metadata statistics. The data study employed a sample of 2000 articles from six WikiProjects, including 1000 Featured Articles (FA) and 1000 non-FA articles. The suggested model's classification and clustering accuracies were compared to those of three previously published models. The proposed model was compared with three previously published models in terms of classification and clustering accuracy. It received classification and clustering accuracies of 95% and 93% respectively, which is a drastic improvement over the existing models. Furthermore, an average inter-rater agreement of 84% was observed. Accordingly, this comprehensive experiment fairly validates the effectiveness of the suggested model. This study contributes to the related knowledge area by introducing a novel framework to assess Wikipedia articles' IQ.

# TABLE OF CONTENTS

## LIST OF FIGURES

# LIST OF TABLES

## LIST OF ABBREVIATIONS

| Abbreviation | Description |
|---|---|
| IQ | Information Quality |
| UGC | User Generated Content |
| WWW | World Wide Web |
| FA | Featured Article |
| SERP | Search Engine Result Pages |
| SEO | Search Engine Optimization |
| ORES | Online Objective Revision Evaluation Service |
| VIF | Variance Inflation Factor |
| EFA | Exploratory Factor Analysis |
| SVM | Support Vector Machine |
| KNN | K-Nearest Neighbour |

# CHAPTER 1 INTRODUCTION

User Generated Content (UGC) is becoming increasingly crucial for knowledge sharing as Web 2.0 persists. Wikipedia is the world's first and largest community-based collaborative encyclopedias, as well as one of the world's largest UGC databases, which is created and maintained by volunteers all over the world. Due to its open-source and collaborative nature, Wikipedia has a serious problem with Information Quality (IQ). This section provides an overview of the study. Firstly, the chapter explains about the background of the study and the problem. Secondly, the formulated research questions are discussed. Afterwards, the study's contribution, implications and limitations are discussed.

## 1.1. Background of the study

Wikipedia is the world's largest and freely available online encyclopedia to date. The term 'Wikipedia' is formed from the two words 'wiki' and 'encyclopedia'. Wiki has the meaning of *quick*. Wikipedia contained 6,348,346 articles as of October 2021. It is a multilingual encyclopedia and supports 325 languages [1]. Wikipedia was first launched on January 15 2001, by Jimmy Wales and Larry Sanger. Currently, Wikipedia is hosted by the Wikimedia Foundation which is an American non-profit organization. Wikipedia is known as a 'collaborative encyclopedia' because its content is collaboratively created by numerous volunteers. This community of volunteers is known as Wikipedians/Wikipedia community and currently, there are 100,422,346 registered members in it. According to Statista, Wikipedia is the second most visited website in the world with 2.5 billion unique monthly visitors while trailing behind only Google [2].

It is a known fact that the traditional encyclopedia is composed of the relevant domain experts and thus, the IQ of the presented content is of a high standard. However, due to the collaborative feature and open-source model of Wikipedia, it often suffers from many IQ issues. For an instance, vandalism is a serious issue seen on Wikipedia. Wikipedia has defined vandalism as any modification made to the content deliberately to compromise its integrity. A famous example of Wikipedia's vandalism is Seigenthaler biography incident [3]. Some anonymous users had inserted a piece of information in Seigenthaler's biography saying that Seigenthaler is suspected of John

F. Kennedy’s assassination. Seigenthaler is a political figure in the USA as well as the founding editorial director of USA Today. This vandal edit has remained on Wikipedia for four months. Following this incident, he has criticized Wikipedia as “a flawed and irresponsible research tool”. Donald Trump, Hillary Clinton, Bill Clinton and Paul Davis Ryan are some other politicians whose biographies have been vandalized on Wikipedia. Since Wikipedia content is referred to by many other websites' web-based systems this wrong information is penetrated through other sources as well. For example, Wikipedia links appear on top of the SERP for many queries on Google. Furthermore, link spamming is another type of attack. Deleting the existing content, adding humour and obscenities within the content, and changing the formatting are some other types of vandal activities. Due to the popularity of Wikipedia in WWW, attackers place inappropriate and proportional external links within Wikipedia’s content. Thus, link spamming would bring large traffic to the landing sites resulting in monetary benefits for the spammers. Further, broken or unavailable destination links and non-copyrighted images and graphs are some other information quality issues. Therefore, the quality of information presented on Wikipedia is always questionable.

Due to this overall IQ issue in Wikipedia, academic professionals do not encourage their students to refer to Wikipedia for academic activities and research purposes [4], [5]. Certain higher educational institutes and universities have policies against referencing Wikipedia [6]. However, it should be worth mentioning that there are high-quality articles as well on Wikipedia. However, these quality articles are also missed out because of the existing overall IQ issues. Wikipedia has its own quality grading system for assessing the IQ of articles. Based on community voting, the articles are categorized into seven classes from best to worse. They are, Featured Articles (FA), A-Class, Good Articles (GA), B-Class, C-Class, Start class, and Stub class [7]. Among these classes, FA articles are considered to be the best quality articles which have undergone an extensive review process. Currently, there are 6,059 FA articles and 35 971 GA articles. However, these quality articles are also ignored purposefully as a result of the overall IQ problem in Wikipedia.

This study focuses on the IQ assessment of the English language Wikipedia. The English Wikipedia currently has 6,360,413 articles. It has 42,063,344 registered users and 1,087 admin users. The average number of new articles per day is 588. However, the community cannot alone monitor thousands of edits made every day. It requires a lot of time and effort. Therefore, it is required to have an efficient IQ assessment process to support the Wikipedia community.

## 1.2. Problem identification

Since around 2010, researchers have tried assessing the IQ of Wikipedia by introducing various theoretical frameworks and statistical models. They have focused on different IQ criteria such as informativeness, the reputation of authors and editors, and word count. Further, some studies have adopted various Machine Learning-based approaches [8]–[11]. However, these approaches have not received significant classification performance. Some of the approaches are not automated and require substantial time and human effort. Furthermore, prior studies have rarely focused on how certain factors like language, the authority of the content, and verifiability affect the IQ of Wikipedia [12].

## 1.3. Research question and hypotheses

Therefore, this study attempts to introduce a novel theoretical framework based on Google's E-A-T quality framework to better classify Wikipedia articles by quality.

RQ1 - What are the Information Quality determinants of Wikipedia?

- RQ1-a – What is the relationship between the Expertise of a Wikipedia article and its Information Quality?
- RQ1-b – What is the relationship between the Authority of a Wikipedia article and its Information Quality?
- RQ1-c – What is the relationship between the Trustworthiness of a Wikipedia article and its Information Quality?

RQ2 – How the Google's E-A-T model can be used to better classify Wikipedia articles by quality?

Based on the literature, the author initially set the following hypotheses.

H1: Expertise of a Wikipedia article is positively associated with its Information Quality

H2: Authority of a Wikipedia article is positively associated with its Information Quality

H3: Trustworthiness of a Wikipedia article is positively associated with its Information Quality

## 1.4. Contribution of the study

The major contributions of this study are listed below.

- Introduces a novel theoretical model based on the E-A-T framework to better classify Wikipedia articles by quality.
- Defines Expertise, Authority and Trustworthiness as the key constructs of Wikipedia's IQ.
- Defines a set of IQ dimensions as Informativeness, Readability Understandability, Reliability, Verifiability and Maturity to measure the above IQ constructs.
- Introduces a comprehensive set of 45 IQ attributes (28 new and 17 existing attributes) to measure the above IQ constructs.

These IQ features were extracted from both article content and metadata statistics using an efficient web scraping algorithm. The dataset is comprised of various IQ features extracted from 2000 articles. Following the data analysis, the model was evaluated in various ways to ensure its validity.

## 1.5. Implications of the study

In terms of theoretical implications, the study presents a novel theoretical model based on Google's E-A-T model to assess the IQ of Wikipedia articles. The model is comprising of (1) three precise IQ constructs, (2) a set of IQ dimensions that influence each of the above constructs and (3) 45 IQ attributes to measure the above IQ dimensions.

With regard to practical implications, assessing the IQ of Wikipedia is beneficial for both readers and reviewers. On one hand, this approach help readers to filter reliable and quality information from Wikipedia. This would in turn add more value to Wikipedia being one of the largest free information repositories in the world, but always getting ignored by scholars. On the other hand, the article reviewers, editors and authors could be notified regarding article improvements. Using the proposed approach, an automated method of identifying flaws and sending notifications to the above quality improving parties can be developed. This would provide immediate guidance for the reviewers. Also, this would expedite the process of quality improvements in Wikipedia rather than doing it manually. Furthermore, an important fact to note is, that the proposed approach is not a replacement for the manual reviewing process; however, it can be used as an initial screening and supportive tool during the formal review process.

## 1.6. Limitations of the study

There are a few limitations of this study. First limitation is, this study is based on only the English version of Wikipedia. Since there are over 300 language versions of Wikipedia, similar studies can be conducted for those language versions as well. Almost all the proposed features except language specific features could be directly applied to other languages. The other limitation is domain specificity where the dataset was collected only from six Wikipedia domains and is limited to only 2000 data points. Thus, results could be more generalized by improving the dataset by domain and size.

# CHAPTER 2 LITERATURE REVIEW

This chapter intends to present the findings of the literature survey conducted to establish the baseline of this study. During the survey it was found that, different IQ dimensions have been presented in the literature to measure the IQ of various domains including Wikipedia. An IQ dimension is a set of quality attributes that constitute a single construct of IQ [13] [14]. Accordingly, this chapter first defines the IQ of Wikipedia and then reviews the existing approaches for IQ assessment followed by E-A-T framework.

## 2.1. Information quality of Wikipedia encyclopedia

Ballou and Tayi define data as “the raw material for the information age” [15]. Information is “useful data that have been processed in such a way as to increase the knowledge of the person who uses the data” [16]. Data quality has been defined as, “data that are fit for use by data consumers" [17]. This extends the definition of information quality as satisfying information consumers’ expectations. Gustavsson and Wänström have defined IQ as the "ability to satisfy stated and implied needs of the information consumer” [14]. In other terms, it is the "measure of the value which the information provides to the user of that information" [18]. An encyclopedia is "a reference resource that provides information on all fields of knowledge or addresses a certain branch of knowledge in depth. It is self-contained and explains subjects in greater detail than a dictionary" [19]. Wikipedia’s content is by a large collaborative community comprising of expert users as well as non-expert users. Thus, it is impossible to provide the most accurate and reliable data always in a collaborative encyclopedia. Taking all of these facts into account, the information quality of a community-based encyclopedia may be defined as providing accurate and trustworthy information on many disciplines of knowledge that the layman can readily understand and perceive.

## 2.2. Existing approaches

When looking at the literature, there are three approaches to assessing the IQ of Wikipedia. They are (1) Metadata-based approach which accounts for the article metadata such as number of edits, number of editors and number of references, etc.; (2) Content-based approach which focuses on internal content related attributes such as word count and (3) Hybrid approach which is a combines the above two approaches. Given below is a detailed review of the above approaches.

### 2.2.1. Metadata-based approach

The term 'metadata' refers to the data about data. Accordingly, the metadata-based approach considers how the meta information about the Wikipedia articles such as the number of editors, their coordination behavior, reputation and expertise impact on article quality. For example, authors in [20]–[22] have focused on the number of the editors, coordination of the editors and edit history. Kittur and Kraut have found how IQ is impacted by the number of editors and their coordination behavior [21]. Their results conclude that the number of articles does not solely improve article quality and quality improves only when appropriate coordination techniques are used. Authors in [23] have observed a strong correlation between the unique editors, edit count and aticle quality. Further, they have used the article talk pages to observe the coordination behavior of the editors. Further, the quality of the editors has been considered an IQ indicator in [24]. Another work in [25] examines the editor's local contribution inequality and the global inequality contribution in assessing the IQ of Wikipedia. The local inequality refers to the editors who contribute to a particular article and global inequality refers to their contribution across all the articles on Wikipedia. They also have found that the editor coordination positively impacts of IQ of Wikipedia articles. Moreover, Liu and Ram have studied how article quality might get impacted as a result of the types of editors [26]. For an instance, high quality artiles have been edited by experienced editors where low quality articles have been edited by non-experienced editors.

Some studies such as [22] and [27] have studied about reputation-based features such as authors' and editors' reputation. Certain studies have used Wikipedia articles' revision history to assess th IQ. The study done by Zeng et al. [28] is an example study for this. They have used a revision history-based approach in assessing the IQ. Revision is referred to as an author editing an article. Further, in the study [29], the authors have assessed the article quality using references. This can be considered as one of the pioneered work to measure the IQ of Wikipedia using references.

However, there are some limitations of this approach. Under this approach, it requires a large set of metadata about the articles to assess the IQ properly. On the other hand, the IQ prediction is not straightforward since it relies on many external facts such as coordination and reputation of editors rather than focusing on article content.

### 2.2.2. Content-based approach

This approach relies on the article content related features and therefore is a more reliable and straightforward IQ assessment approach. For example, Blumenstock et al. [30] have presented word count as an IQ measurement. They have found that word count is a precise indicator of quality which fairly discriminates FA articles from the pool. They have used word count 2000 as the cut-off mark for FA such that if the word count is greater than 2000 the article is FA otherwise non-FA. Their experimental results have shown that even though the metric is simple, it has outperformed most of the complex models presented prior to their work. Later in 2010, Lipka and Stein [31] used the writing style of the articles as an IQ indicator. They have identified the writing style using the character-trigram feature. They also have found that Blumenstock's cutoff 2000 only works fine for unbalanced datasets with less number of lengthy articles. Wöhner and Peters have used the article life cycle to discriminate high quality and low quality articles [32]. The two measurements they have introduced using the article life cycle are persistent and transient contributions. Transient contribution means the number of words that have been modified during a particular time duration. Since vandalized contents are reverted quickly, such issues can be easily detected using transient contributions. Persistent contributions are the edits that can remain to be the prescribed duration of time. Those edits remain for so long since they are effective edits. 11 other sub metrics have been presented using these two metrics. They have evaluated these metrics using classification accuracy for the two quality classes Good and FA articles and Articles for deletion.

Furthermore, in studies such as [33], [34], [35], authors have presented the features like completeness, informativeness, number of headings, number of images, references and readability scores. Moreover, studies like [36] have found that the content of web pages positively impacts towards the user's web quality experience.

### 2.2.3. Hybrid approach

The hybrid approach combines both metadata-based and content-based approaches. Thus it uses both metadata related features and content related features. Some hybrid studies given in the literature are discussed below. The study by Stvilia *et al.* [13] is the pioneering research work which uses a hybrid approach. They have proposed seven metrics with nineteen content and metadata related features extracted from the articles. They have come up with seven metrics by subjecting the collected article data into an Exploratory Factor Analysis (EFA). The seven metrics are Authority/Reputation, Completeness, Complexity, Informativeness, Consistency, Currency and Volatility. Much later research has used the model proposed in [13] as the base model. The methodology adopted in this study also follows this approach. Dalip *et al.* [37] have also used the hybrid approach and they propose four types of features as the text related features, review features, network features and network features.

Later in 2013, the study [33] introduced a set of 17 hybrid features and proposed five IQ metrics. They have named their quality model as an actionable quality model. They have defined the term 'actionable quality model' as a model that has features that directly improve article quality. This model was also based on [13] model. A notable point about the features of this work is that it does not require high computational power to analyse the large amount of historical data such as revision histories. Hence this work also inspired our work to extract features that are readily available and easily extracted without much computational overhead. In 2015, the same authors proposed a model comprising 11 structural features as article length in bytes, number of references, number of links to other articles, number of citation templates, number of non-citation templates, number of categories linked in the article text, number of images/article length, information noise score, has an infobox template, number of level 2 section headings, number of level 3+ section headings. Further, authors in [38] have proposed a model to classify articles by quality using 11 features extracted from article metadata and content related features along with readability metrics. The proposed model in this study also aligns with the hybrid approach. The required features were collected from both article metadata and content features.

## 2.3. E-A-T model- the theoretical framework

The proposed framework was derived based on Google's E-A-T framework. This section gives an overview of this framework. After the algorithm update of Google in 2018, the E-A-T framework gained much attention in the community especially in the Search Engine Optimization (SEO) domain. The model comprises three constructs as Expertise, Authority and Trustworthiness. Expertise has been defined as the "knowledge and skills of the Main Content (MC) creator". Authority has been defined as "authoritativeness of the creator of the MC, the MC itself, and the website" while Trustworthiness is defined as "trustworthiness of the creator of the MC, the MC itself, and the website" [39]. In a nutshell, E-A-T is a feature that marksa a web page is high-quality. Given below is quoted from Google which highlights what they mean by E-A-T "High E-A-T news articles should be produced with journalistic professionalism – they should contain factually accurate content presented in a way that helps users achieve a better understanding of events. High E-A-T news sources typically have published established editorial policies and robust review processes". Google uses E-A-T to determine what content is of high quality and should be ranked higher in their SERP. They have incorporated this guideline to their algorithm. Therefore, it is clear that using this framework, Google attempts to deliver highest quality search experience and information to online users. Following this framework, many companies have been trying to improve the visibility of their websites in the top results of SERPs since 2018. A good example is, healthline.com, which saw an upward spike from SEO visibility 1 500 000 to 1 882 and draxe.com, which lost over half its SEO Visibility 390 984 to 177832 in just two weeks of time (29/07/2018 to 2018/08/12) [40]. Since Wikipedia is also an online community-based platform, the author believes the E-A-T framework would be more appealing to describe the information quality of Wikipedia as well. Therefore, the author critically reviewed the literature focusing on the E-A-T framework to understand the IQ determinant factors with regard to the three major constructs Expertise, Authority and Trustworthiness. To the best of the knowledge of the author, none of the previous studies has conducted a systematic review.

### 2.3.1. Expertise

An "expert Wikipedia article" exclusively provides knowledge and information and consists of contributors with relevant expertise. Prior studies have attempted to assess IQ through expertise. For an instance, some previous research [21], [26] [41], [41] have analysed how different editor characteristics, such as editors' experience, concentration, coordination, and coordination patterns, affect the quality of Wikipedia articles. Studies such as [42] and [43] have adopted networking concepts to identify the behaviour of editors. Studies [30] and [44] argue that article length, despite its simplicity, is an accurate measure of article quality. Some studies focus on the article based on the life cycle [32]. Lipka and Stein [31] have emphasized the writing style of the articles in assessing Wikipedia's' IQ. The frameworks introduced by [33] included a separate index for informativeness. The term "Informativeness" could be described as the ability to provide useful information and comprehensive content. Readability has been a major concern in the studies like [45], [46]. Further, the easiness to read a given text can be measured using readability scores. The most commonly used readability merics include Flesch Kincaid Grade Level, Gunning Fog Score, SMOG Index, Coleman Liau Index and Automated Readability Index [13], [33], [37], [38]. Thus, in this study also author paid attemtion towards considering these metrics. Moreover, information presented in an encyclopedia should be easily understood by a layperson without much effort [13]. In addition to studies related to Wikipedia, studies such as [47] have also used reading ability as an indicator to assess the quality of web content. All of these studies have shown that most of the expertise related features are capable of enhancing the IQ of Wikipedia.

### 2.3.2. Authority

When a source of knowledge is frequently quoted by others, it becomes more reputable and authoritative than its vertical hand [48]. When a source is more powerful it shows a higher quality. The authority of a Wikipedia page can be influenced by the contributors' authority, the content, and the website itself. Some previous research has attempted to assess IQ based on the authors' reputation. Studies [20] [22] have found, for instance, that articles with high ratings are written and edited by people with high reputations. Likewise, according to Nemoto et al. [27], articles initiated by high-profile authors are quickly found to be of high quality. On other hand, there is a lack of studies focusing on content authority and website authority on Wikipedia's IQ evaluation. Only the study in [49] follow article rankings on search engines for IQ assessment. According to studies in [22], [27], Authority enhances IQ on Wikipedia.

#### 2.3.3. Trustworthiness

Reliability means the ability to provide accurate and reliable information. Reliability and trustworthiness are important factors in a collaborative context. Furthermore, a trustworthy Wikipedia article contains information that is truly accurate and verifiable. Zeng et al. [28] have found a method based on revision history to evaluate the reliability of an article. They developed a Bayesian network trust model based on information about the article revision history. Authors in [50] created the Online Objective Revision Evaluation Service (ORES). The revision ID of a Wikipedia article is the only input parameter for ORES. The authors of [12] the article argue that the content of the article should always be verified. Nicholson et al. [29] attempts were made to assess the credibility of Wikipedia articles through references. Lewandowski et al. [51] have conducted a successful study on the popularity and credibility of modelling. In hybrid frames introduced by [33] separate indicators are used to indicate the reliability of articles. According to these studies, [12] [29], [50] Trustworthiness improves Wikipedia's IQ.

Accordingly, the author identified seven IQ dimensions from the literature which directly influence the three main IQ constructs as Informativeness, Understandability, Readability, Authority, Reliability, Maturity and Verifiability. Subsequently, the author identified 17 IQ attributes from the literature that can be used to measure the IQ of Wikipedia. Then this list was further enriched by adding 28 new IQ attributes by carefully analysing the Wikipedia articles, their metadata statistic and the article content (see Table 2.1). The special feature of all these articles is that these features are readily available and can be directly extracted from the article content and metadata without much computational overhead.

Table 2.1 IQ attributes

| No | Information Quality measurement | Description | Source |
|---|---|---|---|
| 1 | Info Noise | This is calculated as, *1- (The size of the term/token vector after stemming and stopping) / (document size before processing)* [52] | Article |
| 2 | Total Num. Edits | Total number of edits in an article [52] | Metadata |
| 3 | Num. Unique Editors | The number of distinct editors per article [52] | Metadata |
| 4 | Diversity | This is calculated using, *Num. Unique Editors / Total Num. Edits* [52] | Metadata |
| 5 | Num. Images | Number of images placed within the text of the article [52] | Article |
| 6 | Num. Internal Broken Links | Number of links whose destinations are broken or unavailable [52] | Article |
| 7 | Num. Internal Links | Indicates the number of links from one Wikipedia article to another page in Wikipedia Community itself [52] | Article |
| 8 | Num. Reverted Edits | Number of edits that have been reverted to original/ previous version [52] | Metadata |
| 9 | Median Revert Time | In minutes [52] | Metadata |
| 10 | Age | Article age in days [52] | Metadata |
| 11 | Currency | The time between the dump date and the date of the last update of the article in days [52] | Metadata |
| 12 | Connectivity | Number of Articles connected to a particular article through common editors [52] | Article |
| 13 | Num. External Links | Indicates the number of links from a Wikipedia page to external sources outside Wikipedia [52] | Article |
| 14 | Num. Registered User Edits | Indicates the number of edits of a page made by registered users [52] | Metadata |
| 15 | Num. Anonymous User Edits | Indicates the number of edits of a page made by anonymous users [52] | Metadata |
| 16 | Flesch Kincaid Grade Level | This score is a modification of the Flesch reading score [53], which estimates the difficulty of reading and comprehending an article [54] | Article |
| 17 | Gunning Fog Score | This score can measure the simplicity of the text, which is correlated with the reading level [55] | Article |
| 18 | SMOG Index | This score estimates the reading grade that readers had to | Article |

| | | | |
|---|---|---|---|
| | | attain to interpret the text [56] | |
| 19 | Coleman Liau Index | This index estimates the reading level needed to understand an article [57] | Article |
| 20 | Automated Readability Index | This indicator approximates the age needed to understand an article [44] [58] | Article |
| 21 | Total Num. Headings | The number of headings in a Wikipedia page [38] | Article |
| 22 | Num. Level-1 Headings | Indicates the number of Level-1 subheadings [38] | Article |
| 23 | Num. Level-2 Headings | Indicates the number of Level-2 subheadings [38] | Article |
| 24 | Page Authority | Indicates the authority of the content of a single Wikipedia page over the web. This is measured by how well a specific page is ranked in SERP. This is has a logarithmic scale that ranges from 1 to 100 [59] | Article |
| 25 | Num. Linking Domains | This score measures the number of unique external links to a website. In case of a website has more than one link from the same website, it is considered as one linking root domain [59] | Article |
| 26 | Num. Inbound Links | The inbound links/ backlinks refer to the hyperlinks from another web page to a page on the particular website [59] | Article |
| 27 | Num. Ranking Keywords | The number of keywords of a particular website includes among the top 50 keywords used by Google's ranking [59] | Article |
| 28 | Num. Words | Indicates the number of words of an article [60] | Article |
| 29 | Num. Sentences | Indicates the number of sentences in an article [60] | Article |
| 30 | Num. Complex Words | Indicates the number of complex words in an article [60] | Article |
| 31 | Complex Words Percentage | Indicates the percentage of complex words over the total number of words in an article [60] | Article |
| 32 | Average Words Per Sentence | Indicates the average words per sentence in an article [60] | Article |
| 33 | Average Syllables Per Word | Indicates the average syllables per word in an article [60] | Article |
| 43 | Minor Edits | This indicates the number of edits that the editor believes require no review and could never be the subject of a dispute [61] | Metadata |
| 35 | Bot Edits | Indicates the edits made by automated software tools for a page [13] | Metadata |

| | | | |
|---|---|---|---|
| 36 | Semi-Automated Edits | Indicates the edits made by semi-automated software tools for a page [63] | Metadata |
| 37 | Edits Made By Top 10% Editors | This indicates the number of edits made by the set of editors ranked among the top 10% of the editors [64] | Metadata |
| 38 | Page Size | This measure is calculated bytes, including the entire page's size, including text, images, maps, etc. [65] | Article |
| 39 | Num. Copyrighted Images | Certain images in Wikipedia are copyrighted, while some images are not. Images with copyright status are an indication of IQ [46]. Thus, this measure indicates the number of copyrighted images in an article. | Article |
| 40 | Num. Characters | Number of all the characters (numbers, letters and symbols etc.) included in a page [38] | Article |
| 41 | Num. References | Indicates the number of references included in an article [38] | Article |
| 42 | Num. Page Watchers | For any particular page, it is possible to discover how many users have it on their watchlist. When a page is included in one's watch list, he will be able to track any changes done to that page [66] | Metadata |
| 43 | Num. Page Views in 60 Days | Number of views for a page within the last 60 days [67] | Metadata |
| 44 | Num. Edits Made In Past 30 Days | Number of edits made to a page within the last 30 days [68] | Metadata |
| 45 | Average Time Between Edits | Average time between edits of the page [68] | Metadata |

## 2.4. Limitations of the existing work

According to the above review, several limitations of the existing work need to be discussed. Firstly, compared to Expertise and Trustworthiness, much attention to Authority towards IQ has not been paid off. It is clear that a high web presence leads to a high reputation because other experts are citing a web page/website. Analogously, a good reputation improves the quality of the content [69]. Further, when the authority construct was removed from the proposed model, a significant accuracy drop was observed (reduced the accuracy by 16%). Accordingly, the Authority should be considered as a major construct to measure the IQ of Wikipedia articles. Secondly, there is a lack of studies on how maturity and verifiability affect IQ [12]. Finally, even if there are models and frameworks that have been proposed to assess IQ, they have not been able to provide appropriate performance for classifying articles into corresponding classes according to their IQ  [13], [38], [41]. According to all these facts, it is clearly seen that Expertise, Authority and Trustworthiness can be used to precisely describe the IQ of Wikipedia. Therefore, Google's E-A-T model can be brought to assess Wikipedia's IQ. Further, we identified that, Informativeness, understandability and Readability influence Expertise and Maturity, Verifiability and Reliability influence Trustworthiness. Furthermore, this study identified 45 various IQ attributes that can be used to measure these dimensions. The theoretical framework and the operationalization of the study variables will be discussed in the Methodology section.

## CHAPTER 3 RESEARCH METHODOLOGY

This chapter discusses the research methodology of the study. First, it describes the conceptual framework and the derived hypotheses. Then it explains the operationalization of the study variables and finally, the research design and hypotheses development are discussed in detail.

### 3.1. Conceptual framework

Aligning with the literature, the study proposes the following conceptual framework given in Figure 3.1. This framework is developed on top the Google's E-A-T model. Accordingly, the conceptual framework comprises of three IQ constructs, a set of IQ dimensions that influence each construct and 45 IQ attributes which can be used to measure each dimension. The three main IQ constructs are Expertise, Authority and Trustworthiness. The dimensions of Informativeness, Readability and Understandability influence the construct of Expertise. The dimensions of Verifiability, Trustworthiness and Maturity influence Reliability.

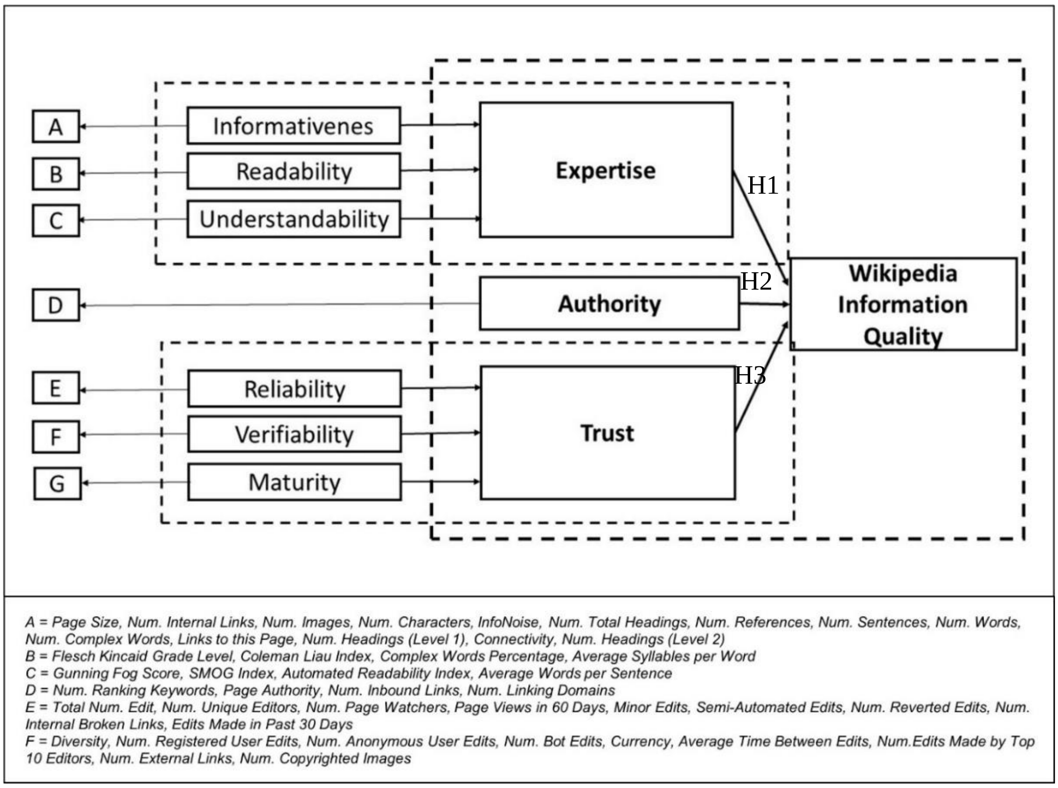



Figure 3.1 Conceptual framework

## 3.2. Operationalization of the study variables

In order to establish the model, first, it is required to test the hypotheses that were developed with the assistance of literature. Accordingly, the IQ becomes the dependent variable of the model. It is categorical and has two values as FA which are of high quality and non-FA which are of not high class (all the other clases except FA should fall under this class). The independent variables are the three IQ constructs Expertise, Authority and Trustworthiness. They are numerical and how the numerical values were derived will be discussed in Section 4. The influencing IQ dimensions for the three constructs and the operationalization of the IQ dimensions were done using an EFA analysis as discussed in [13]. Aligning with the empirical evidence and the EFA results, the following operationalization table (Table 3.1) was derived to operationalize the study variables.

Table 3.1 Operationalization of the study variables

| Type of the variable | IQ construct | IQ dimension | IQ attribute | Data type |
|---|---|---|---|---|
| Dependent | Expertise | Informativeness | Page Size | Numerical |
| | | | Num. Internal Links | Numerical |
| | | | Num. Images | Numerical |
| | | | Num. Characters | Numerical |
| | | | InfoNoise | Numerical |
| | | | Num. Total Headings | Numerical |
| | | | Num. References | Numerical |
| | | | Num. Sentences | Numerical |
| | | | Num. Words | Numerical |
| | | | Num. Complex Words | Numerical |
| | | | Links to This Page | Numerical |
| | | | Connectivity | Numerical |
| | | | Num. Level-1 Headings | Numerical |
| | | | Num. Level-2 Headings | Numerical |
| | | Readability | Flesch Kincaid Grade Level | Numerical |
| | | | Coleman Liau Index | Numerical |
| | | | Complex Words Percentage | Numerical |
| | | | Average Syllables Per Word | Numerical |
| | | Understandability | Gunning Fog Score | Numerical |
| | | | SMOG Index | Numerical |
| | | | Automated Readability Index | Numerical |
| | | | Average Words Per Sentence | Numerical |
| | Authority | Authority | Page Authority | Numerical |
| | | | Num. Linking Domains | Numerical |
| | | | Num. Inbound Links | Numerical |
| | | | Num. Ranking Keywords | Numerical |
| | Trustworthiness | Reliability | Total Num. Edits | Numerical |
| | | | Num. Unique Editors | Numerical |
| | | | Num. Page Watchers | Numerical |

| | | | | |
|---|---|---|---|---|
| | | | Page Views in 60 Days | Numerical |
| | | | Minor Edits | Numerical |
| | | | Semi-Automated Edits | Numerical |
| | | | Num. Reverted Edits | Numerical |
| | | | Num. Internal Broken Links | Numerical |
| | | | Edits Made In Past 30 Days | Numerical |
| | | Verifiability | Diversity | Numerical |
| | | | Num. Registered User Edits | Numerical |
| | | | Num. Anonymous User Edits | Numerical |
| | | | Num. Bot Edits | Numerical |
| | | | Currency | Numerical |
| | | | Average Time Between Edits | Numerical |
| | | | Num. Edits Made By Top 10 Editors | Numerical |
| | | | Num. External Links | Numerical |
| | | | Num. Copyrighted Images | Numerical |
| | | Maturity | Median Revert Time | Numerical |
| | | | Age | Numerical |
| Independent | IQ of Wikipedia | | IQ quality class of the article | Categorical (FA and non-FA) |

## 3.3. Research design

This is performed as quantitative research. A statistical analysis was conducted to analyze the data. The IQ attributes described in Section 2.4 were extracted from Wikipedia articles' content and the metadata statistics page. Accordingly, the data collection technique is secondary. Avoiding most of the drawbacks seen in previous studies about manual feature extraction [38], this study adopted an automated data collection technique which is less time consuming. With the help of their web URLs, an algorithm was implemented with Selenium 3.14 to quickly and inexpensively scrape the relevant data from Wikipedia pages. Selenium is a free cross-browser testing and validation tool for web applications. [70]. The sample code snippets are given in Appendix A. The readability related matrices were extracted from WebFX [71] by sending the article's URL to the tool. Similarly, the Authority related attributes were extracted from Moz [72]. All the other attributes were extracted from the article content and its metadata statistics page.

The study population is the articles from the Wikipedia repository. Wikipedia repository has categorized its article into various projects known as WikiProjects. Each of these WikiProjects covers a specific area of knowledge e.g., Politics. Accordingly, the required data were collected from 2000 articles from six WikiProjects Medicine [73], Politics [74], Sports [75], History [76], Science [77] and, Biographies [78]. These WikiProjects were specifically chosen because they are considered to be the most saturated and active WikiProjects and included a large number of FA articles. Even though many WikiProjects include enough non-FA articles, most of the time they contain a smaller number of FA articles. Since the dataset should be balanced such that data should be extracted from 1000 FA and 1000 non-FA articles, the above mentioned WikiProjects were chosen. Then the 45 IQ attributes described in Section 2.4 were extracted from these 2000 Wikipedia articles. Based on the studies such as [33] [13] and community recommendations for FA articles' quality, the authors decided to consider FA as the basis for assessing the IQ of Wikipedia.

## 3.4. Development of hypotheses

As discussed in Section 2.4, considering the association of Expertise, Authority and Trustworthiness towards IQ of Wikipedia, the following are the postulated hypotheses for the study.

H1: Expertise of a Wikipedia article is positively associated with its Information Quality

H2: Authority of a Wikipedia article is positively associated with its Information Quality

H3: Trustworthiness of a Wikipedia article is positively associated with its Information Quality

# CHAPTER 4 DATA ANALYSIS

This section discusses in detail the data analysis process followed in the study. Mainly two EFA analyses were conducted followed by a regression analysis. The initially established hypotheses were tested using the regression analysis.

## 4.1. Data preparation for analysis

In the first phase, the dataset was subjected for cleaning purposes with IBM SPSS 26.0 software to eliminate missing values and outliers. The missing data can is defined as one or more of the values for variables that are not available for the data analysis. Little's MCAR test was used to alayse the missing value pattern (Chi-Square = 1013.518, DF = 40, Sig = 0.000). Accordng to the test results, missing values have been randomly distributed. Thus, series mean was used to replace the missing values. Subsequently, outlier detection was done. Outliers reflect a significantly different features than the majority of the data points. The detected small number of outliers were also replaced with series mean in order to avoid loss of data points. Afterwards, the data were normalized with log transformation i.e., log10 arithmatic transformation.

## 4.2. Exploratory factor analysis

Then EFA was used to determine the underlying pattern of the data. Accordingly, two EFAs were conducted to identify the variable groupings for the 45 IQ attributes and to identify which IQ dimensions most influence the three IQ constructs: Expertise, Authority and Trustworthiness.

### 4.2.1. 1st order IQ functions

As a first phase, an EFA was undertaken for the 45 IQ variables to determine the underlying structure. Seven components were proposed by the research. The factor loadings are given in Table 4.1.

Table 4.1 Factor loadings of 1st EFA

| Indicator | Components | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Page Size | -0.016 | 0.120 | 0.009 | -0.023 | -0.016 | 0.055 | -0.034 |
| Total Num. Edits | 0.114 | -0.054 | -0.005 | -0.018 | 0.022 | 0.013 | -0.007 |
| Num. Unique Editors | 0.100 | -0.034 | -0.013 | -0.020 | 0.025 | -0.029 | 0.086 |
| Diversity | 0.003 | 0.034 | -0.018 | 0.005 | 0.005 | -0.029 | 0.526 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Num. Page Watchers | 0.176 | -0.171 | -0.014 | -0.009 | 0.054 | 0.022 | -0.103 |
| Num. Page Views in 60 Days | 0.167 | -0.166 | -0.022 | 0.006 | 0.048 | -0.041 | -0.049 |
| Num. Registered User Edits | 0.126 | -0.088 | 0.012 | 0.002 | 0.004 | 0.082 | -0.184 |
| Minor Edits | 0.107 | -0.048 | -0.008 | -0.019 | 0.034 | -0.002 | 0.026 |
| Num. Anonymous User Edits | 0.080 | -0.010 | -0.020 | -0.035 | 0.024 | -0.050 | 0.144 |
| Bot Edits | 0.054 | 0.029 | -0.008 | -0.019 | 0.011 | -0.052 | 0.135 |
| Semi-Automated Edits | 0.119 | -0.062 | -0.016 | -0.025 | 0.039 | -0.016 | 0.057 |
| Num. Reverted Edits | 0.129 | -0.078 | -0.016 | -0.025 | 0.043 | -0.014 | 0.043 |
| Median Revert Time | -0.015 | 0.085 | -0.032 | 0.015 | 0.023 | 0.561 | -0.087 |
| Age | -0.017 | 0.077 | -0.039 | 0.026 | 0.034 | 0.522 | 0.012 |
| Currency | -0.089 | 0.156 | -0.039 | 0.036 | 0.050 | -0.206 | 0.464 |
| Average Time Between Edits | 0.003 | 0.029 | -0.019 | 0.019 | 0.025 | 0.138 | 0.360 |
| Edits Made By Top 10% Editors | 0.058 | 0.013 | 0.008 | -0.012 | -0.005 | 0.060 | -0.085 |
| Connectivity | -0.047 | 0.016 | -0.002 | -0.014 | 0.003 | 0.009 | -0.018 |
| Num. Copyrighted Images | -0.027 | 0.016 | -0.002 | -0.014 | 0.003 | 0.009 | 0.451 |
| Num. Internal Broken Links | 0.119 | -0.099 | 0.005 | 0.024 | 0.043 | -0.045 | -0.024 |
| Num. Internal Links | -0.046 | 0.164 | 0.091 | 0.002 | -0.163 | 0.009 | -0.018 |
| Num. External Links | 0.082 | -0.021 | 0.018 | -0.040 | -0.040 | 0.052 | -0.144 |
| Num. Images | -0.037 | 0.120 | 0.058 | 0.017 | -0.056 | 0.002 | 0.007 |
| Num. Characters | -0.070 | 0.182 | -0.009 | -0.014 | 0.031 | 0.048 | 0.076 |
| Info Noise | -0.070 | 0.181 | -0.013 | -0.014 | 0.031 | 0.047 | 0.079 |
| Total Num. Headings | -0.063 | 0.170 | 0.029 | -0.018 | -0.037 | 0.010 | 0.013 |
| Num. References | -0.032 | 0.141 | 0.001 | -0.018 | -0.030 | 0.070 | -0.008 |
| Flesch Kincaid Grade Level | 0.010 | 0.006 | 0.213 | 0.009 | 0.053 | -0.013 | -0.080 |
| Gunning Fog Score | 0.038 | -0.045 | -0.014 | -0.006 | 0.393 | 0.068 | 0.176 |
| SMOG Index | 0.027 | -0.053 | 0.021 | -0.013 | 0.352 | -0.011 | -0.117 |
| Coleman Liau Index | 0.001 | 0.026 | 0.185 | 0.026 | -0.031 | 0.004 | 0.085 |
| Automated Readability Index | 0.011 | -0.006 | 0.145 | 0.029 | 0.181 | -0.002 | 0.056 |
| Num. Sentences | -0.032 | 0.151 | 0.030 | -0.012 | -0.112 | 0.052 | -0.010 |
| Num. Words | -0.041 | 0.154 | 0.001 | -0.014 | -0.026 | 0.056 | 0.018 |
| Num. Complex Words | -0.039 | 0.155 | 0.037 | -0.014 | -0.029 | 0.055 | -0.010 |
| Complex Words Percentage | -0.028 | 0.069 | 0.260 | 0.012 | -0.131 | -0.061 | -0.117 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Average Words Per Sentence | 0.015 | -0.057 | -0.119 | -0.004 | 0.344 | -0.010 | -0.069 |
| Average Syllables Per Word | -0.012 | 0.048 | 0.253 | 0.011 | -0.126 | -0.042 | -0.080 |
| Page Authority | -0.068 | 0.042 | 0.012 | 0.253 | -0.022 | -0.046 | 0.055 |
| Num. Linking Domains | -0.029 | -0.031 | 0.025 | 0.349 | -0.011 | 0.033 | 0.026 |
| Num. Inbound Links | -0.024 | -0.027 | 0.025 | 0.307 | -0.011 | 0.031 | 0.028 |
| Num. Ranking Keywords | 0.007 | -0.069 | 0.019 | 0.349 | 0.011 | 0.032 | 0.017 |
| Num. Edits Made In Past 30 Days | 0.119 | -0.118 | -0.003 | 0.004 | 0.040 | -0.068 | -0.044 |
| Num. Level-1 Headings | -0.041 | 0.114 | 0.001 | -0.014 | -0.026 | 0.056 | 0.018 |
| Num. Level-2 Headings | 0.001 | 0.114 | 0.005 | 0.026 | -0.031 | 0.004 | 0.085 |
| *Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.* | | | | | | | |

The following seven 1st order IQ functions were developed to calculate values for each of the IQ dimensions depending on the nature of the suggested variable groupings and empirical findings (refer to equations (1)-(7)). The researchers kept the component score coefficients when designing the functions because the IQ measures were not standardized and used diverse scales at the start [13].

$$Informativeness = 0.12 * Page\ Size + 0.164 * Num.Internal\ Links + 0.12 * Num.Images + 0.182 * Num.Characters + 0.181 * InfoNoise + 0.17 * Num.Total\ Headings + 0.141 * Num.References + 0.151 * Num.Sentences + 0.154 * Num.Words + 0.155 * Num.Complex\ Words + 0.016 * Links\ to\ this\ Page + 0.016 * Connectivity + 0.11 * Num.Headings\ (Level\ 1) + 0.11 * Num.Headings\ (Level\ 2) \quad (1)$$

$$Readability = 0.213 * Flesch\ Kincaid\ Grade\ Level + 0.185 * Coleman\ Liau\ Index + 0.26 * Complex\ Words\ Percentage + 0.253 * Average\ Syllables\ per\ Word \quad (2)$$

$$Understandability = 0.393 * Gunning\ Fog\ Score + 0.352 * SMOG\ Index + 0.181 * Automated\ Readability\ Index + 0.344 * Average\ Words\ per\ Sentence \quad (3)$$

$$Authority = 0.253 * Page\ Authority + 0.349 * Num. Linking\ Domains + 0.307 * Num. Inbound\ Links + 0.349 * Num. Ranking\ Keywords \quad (4)$$

$$Reliability = 0.114 * Total\ Num. Edits + 0.1 * Num. Unique\ Editors + 0.176 * Num. Page\ Watchers + 0.167 * Page\ Views\ in\ 60\ Days + 0.107 * Minor\ Edits + 0.119 * Semi-Automated\ Edits + 0.129 * Num. Reverted\ Edits + 0.119 * Num. Internal\ Broken\ Links + 0.119 * Edits\ Made\ In\ Past\ 30\ Days \quad (5)$$

$$Verifiability = 0.526 * Diversity - 0.184 * Num. Registered\ User\ Edits + 0.144 * Num. Anonymous\ User\ Edits + 0.135 * Num. Bot\ Edits + 0.464 * Currency + 0.36 * Average\ Time\ Between\ Edits - 0.085 * Num. Edits\ Made\ by\ Top\ 10\ Editors - 0.144 * Num. External\ Links + 0.45 * Num. Copyrighted\ Images \quad (6)$$

$$Maturity = 0.561 * Median\ Revert\ Time + 0.522 * Age \quad (7)$$

### 4.2.2. 2nd order IQ functions

Next, the dataset with values for all the above IQ dimensions was fed to the second EFA to identify the IQ dimensions that influence each E-A-T construct. The analysis exactly suggested three components as the authors expected (see Figure 3.1). Below Table 4.2 represent the factor loadings of the second EFA analysis.

Table 4.2 Factor loadings of 2nd EFA

| Indicator | Component | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| Reliability | -0.092 | 0.362 | 0.059 |
| Informativeness | 0.144 | 0.01 | 0.255 |
| Readability | 0.032 | -0.252 | 0.654 |
| Authority | 0.370 | -0.199 | -0.033 |
| Understandability | -0.083 | 0.308 | 0.557 |
| Maturity | -0.503 | 0.369 | 0.061 |

| Verifiability | 0.320 | -0.846 | 0.071 |
|---|---|---|---|
| Extraction Method: Principal Component Analysis.<br>Rotation Method: Varimax with Kaiser Normalization. | | | |

Expertise, Authority, and Trustworthiness were constructed by studying the nature of factor loading and results from the literature. Each component was measured using the following 2nd order IQ functions (see equations (8) through (10). The extraction approach for both EFAs was Principal Component Analysis (PCA), and the rotation method was Varimax with Kaiser Normalization.

$$Expertise = 0.255 * Informativeness + 0.654 * Readability + 0.557 * Understandability \quad (8)$$

$$Authority = 0.37 * Authority \quad (9)$$

$$Trustworthiness = 0.362 * Reliability + 0.369 * Maturity - 0.846 * Verifiability \quad (10)$$

Since these metrics are subjected to regression analysis, the data was analyzed for the four parametric assumptions after finding the variable grouping by the first EFA. The results of the aparametric tests are discusses in detail in the next section 4.3. The descriptive statistics of the IQ constructs and IQ dimensions are given in below Table 4.3.

Table 4.3 Descriptive statistics of IQ constructs and IQ dimensions

| IQ dimension | Mean | SD | Skewness | Kurtosis | VIF | IQ constructs | Mean | SD | Skewness | Kurtosis | VIF |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Informativeness | 3.87 | 0.91 | 0.05 | -1.37 | 1.32 | Expertise | 2.08 | 0.28 | -0.21 | -1.03 | 1.67 |
| Readability | 0.77 | 0.07 | 0.11 | 2.83 | 1.53 | | | | | | |
| Understandability | 0.98 | 0.14 | -0.97 | 2.84 | 1.89 | | | | | | |
| Authority | 2.36 | 0.84 | 0.53 | -0.26 | 1.57 | Authority | 2.97 | 0.31 | -0.53 | -0.26 | 1.39 |
| Reliability | 1.76 | 0.73 | 0.39 | -1.07 | 1.25 | Trustworthiness | 3.23 | 0.46 | 0.10 | 0.06 | 1.32 |
| Verifiability | 1.65 | 0.46 | -0.56 | 1.37 | 1.81 | | | | | | |
| Maturity | 1.06 | 0.90 | -033 | 0.06 | 2.18 | | | | | | |

## 4.3. Testing for multivariate assumptions

Following the EFA, the four parametric assumptions (i.e. . (1) Normality, (2) Linearity, (3) Multicollinearity, and (4) Homoscedasticity) were tested in order to conduct the multiple linear regression analysis. Testing for parametric assumptions is important before conducting any multivariate analysis.

### 4.3.1. Normality

Normality assesses the form of each study variable's data distribution and its resemblance to the normal distribution. Skewness and Kurtosis were employed to check the dataset's normalcy (Table 4.3). The presence of values above or below 0 in a distribution indicates that the distribution is normal [79].

### 4.3.2. Linearity

In most multivariate analysis, linearity is a basic assumption. The linear relationship between the study variables is determined by linearity. Scatter plots were used to test linearity in this investigation. All of the correlations between the research variables were found to be linear. As per the scatter plots in Figure 4.1, (a) depicts the scantter plot of Expertise with IQ, (b) depicts the scantter plot of Authority with IQ and (c) depicts the scantter plot of Trustworthiness with IQ.

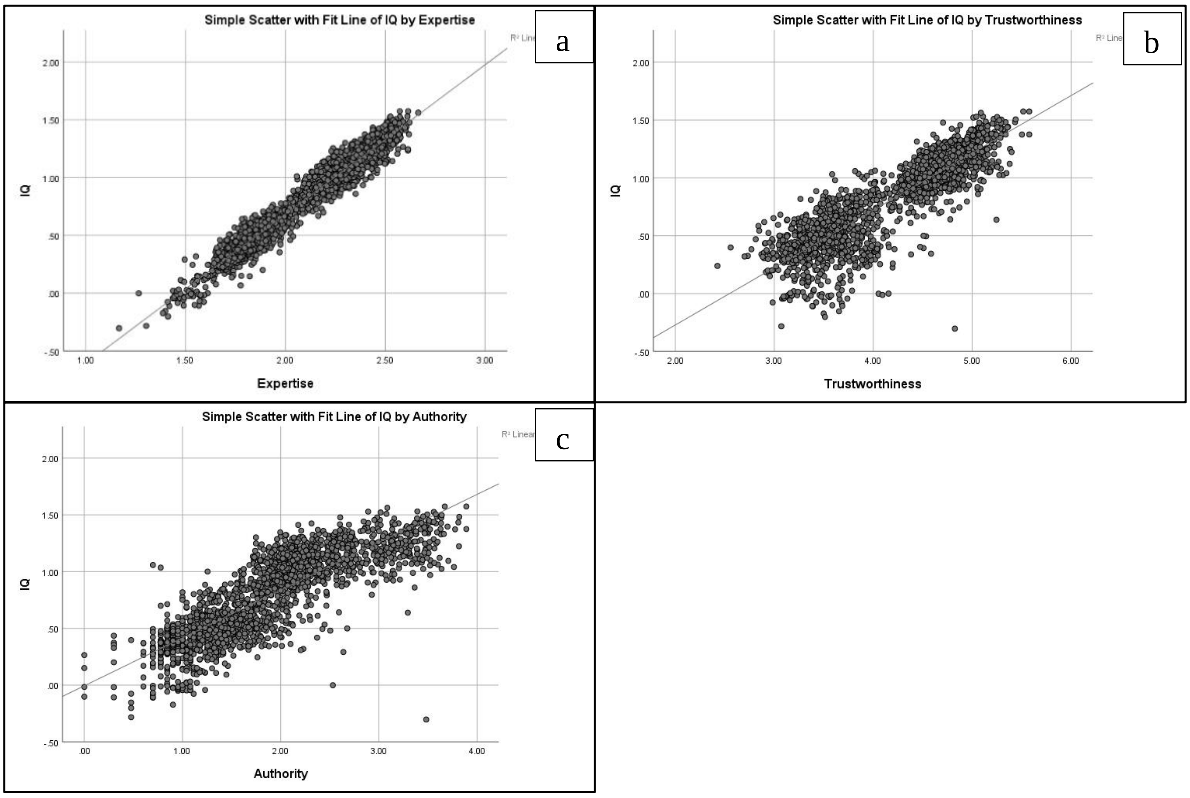



Figure 4.1 Scatter plots drawn for the three independent variables with IQ

### 4.3.3. Multicollinearity

Multicollinearity occur as a result of the effect created when two or more independent variables are combined, intervined and overlapped together [80]. The correlation matrix was used as the primary multicollinearity test. Multicollinearity among independent variables exists when all pairwise correlation values for all independent variables exceed 0.80 or 0.90 [79]. The correlation values of all independent variable pairs in this investigation were less than 0.80 (see below Table 4.5). It means that the independent variable pairings have no significant multicollinearity. For this study, the Pearson correlation coefficient was utilized.

Table 4.4 Pair wise Pearson correlation of independent variables

| Variable | Expertise | Authority | Trustworthiness |
|---|---|---|---|
| Expertise | 1 | 0.531** | 0.511** |
| Authority | | 1 | 0.313** |
| Trustworthiness | | | 1 |
| ** P-value <0.01 (2-tailed) | | | |

Moreover, multicollinearity was tested with VIF. As per [79], VIF values are less than 5, which is a sign of no multicollinearity. The VIF of the research variables ranges from 1.32 – 1.67 (< 3), indicating that there is no substantial multicollinearity (Table 4.6).

Table 4.5 VIF of the independent variables

| Independent variable | VIF |
|---|---|
| Expertise | 1.67 |
| Authority | 1.39 |
| Trustworthiness | 1.32 |

### 4.3.4. Homoscedasticity

Homoscedasticity is described as dependent variables with identical variance across a range of independent variables [80]. The regression line method was used to assess the relationships between variables (See Figure 4.2). It was discovered that all of the dimensions had considerable homoscedasticity when the residual plots were examined since IQ is almost identidcally vary across the ranges of Expertise (a). Authority (b) and Trustworthiness (c).

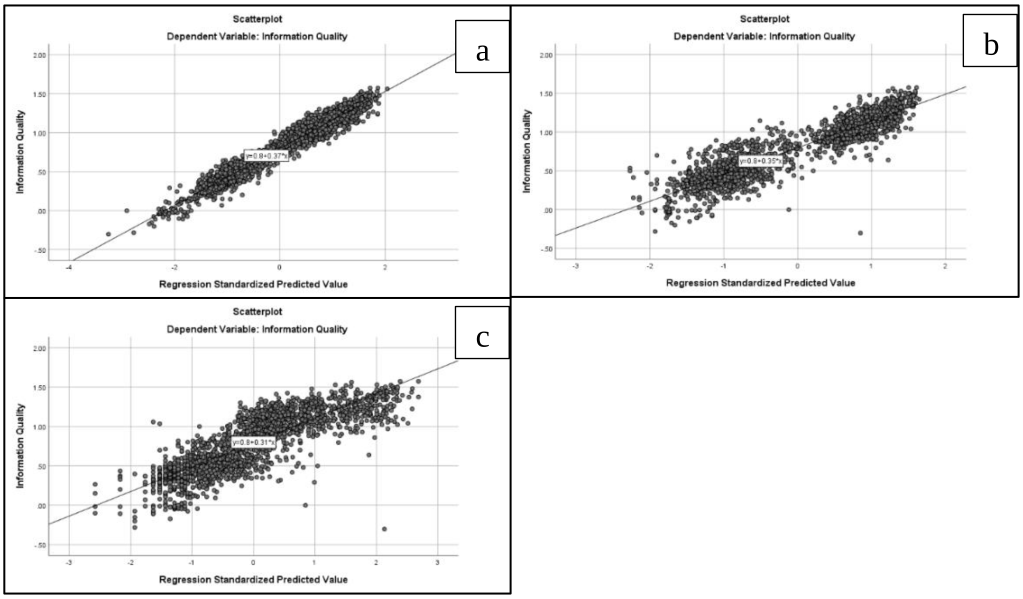



Figure 4.2 Residual plots of four dimensions Expertise (a) Authority (b) Trustworthiness (c) with Information Quality

## 4.4. Regression analysis and hypotheses testing

The relationship between each of the quality constructs Expertise, Authority, and Trustworthiness and Wikipedia's IQ was then examined using a multiple linear regression analysis. A hypothesis test was used to verify the observations. The three IQ constructs Expertise, Authority, and Trustworthiness were the independent variables, and they were measured using the aforementioned 1$^{st}$ and 2$^{nd}$ order IQ functions. The article's relevant quality grading class determined the dependent variable's category value, which has just two alternatives: FA (Featured Articles) and non-FA (Non-Featured Articles). Consequently, the following multiple linear regression model was derived (Equation 11). Further, in order to improve the effectiveness of the regression analysis and mitigate the overfitting while reducing the bias, the dataset was cross-validated using SPSS bootstrapping (refer Table 4.7).

$$Information\ Quality = 1.191 * Expertise + 0.281 * Authority + 0.165 * Trustworthiness - 2.139 \quad (11)$$

Table 4.6 Cross validation results

| Model | B | Bootstrap | | | | |
|---|---|---|---|---|---|---|
| | | Bias | Std. error | Sig. (2-tailed) | Confidence interval | |
| | | | | | Lower | Upper |
| (Constant) | -2.139 | -0.003 | 0.054 | 0.001 | -2.235 | -2.028 |
| Expertise | 1.191 | 0.001 | 0.026 | 0.000 | 1.132 | 1.240 |
| Authority | 0.281 | 0.001 | 0.020 | 0.000 | 0.239 | 0.319 |
| Ttrustworthiness | 0.165 | 0.000 | 0.015 | 0.000 | 0.139 | 0.192 |
| Sampling method – simple<br>Number of samples – 1000<br>Confidence interval 95.0%<br>Confidence interval type - Bias-corrected and accelerated (BCa) | | | | | | |

The regression analysis findings are shown in Table 4.7 below. As R's value is 0.886 and $R^2$ is 0.750, it can be stated that the proposed model is exhibiting a significant explanatory power. This suggests that Expertise, Authority and, Trustworthiness fairly predict IQ. The results indicated that Wikipedia articles' Expertise is positively associated with its IQ (ß = 0.673, p<0.000). Similar to this, a Wikipedia articles' Authority and Trustworthiness are both positively associated with IQ (ß = 0.174 and 0.154, respectively, p<0.000). The study's findings therefore confirm that Expertise, Authority, and Trustworthiness improve the IQ of Wikipedia articles.

Table 4.7 Results of the regression analysis

| Hypotheses | ß | Result |
|---|---|---|
| H1 | 0.673*** | Supported |
| H2 | 0.175*** | Supported |
| H3 | 0.154*** | Supported |
| N = 2000; * $p<0.05$; ** $p<0.01$; *** $p<0.001$ (2-tailed); R = 0.886; $R^2$ =0.750.; Adjusted $R^2$ = 0.747 | | |

# CHAPTER 5 EVALUATION

This chapter discuss about the evaluaton results. The suggested model was extensively evaluated and validated in three ways i.e., (1) Classification performance, (2) Clustering performance and (3) Interrater-reliability test.

### 5.1.1. Classification performance

Expertise, Authority, and Trustworthiness were the independent variables, whereas IQ was the dependent variable. The Decision Tree classifier was used by Stvilla et al. [13] to test their model. The Decision Tree could be examined to see how specific aspects influence the construction of complicated models. As a result, the researchers used the same classifier to evaluate the classification performance of the proposed model in terms of accuracy and F1-measure. The F1-measure combines precision and recall into a single statistic by taking into account the presence of a harmonic mean. As a result, the proposed model obtained a 95% accuracy rating and a 94% F1-measure rating.

Next, proposed model's classification performance was compared to that of the preceding three well-known models. Model 1 was proposed by Stvilla et al. [13] in 2005 and is considered as one of the pioneering study in this knowledge area [13]. This model was chosen for evaluation due to following two reasons: (1) since most of the later work have used this model as the baseline model [33], (2) Author wanted to systematically analyze the significant landmarks about this knowledge area from 2005 to 2021. This study is the pioneering work of hybrid approach of quality assessment for Wikipedia. Therefore, this model was choen for the comparison of the proposed models. Model 2 was proposed by Warncke-Wang et al. [33] in 2013, while Model 3 is a deep learning-based model proposed by Shen et al. in 2017. Shen et al. defined quality assessment as a classification problem and provided a model based on a collection of hand-engineered features utilizing a Long Short Term Memory (LSTM). Even if several models have been presented after 2017 towards end of 2021, author did not chose those models due to following reasons (1) the deployed features are almost same in the selected three models e.g., studies in [81] [82](2) less performance of the proposed models [82] compared to the chosen models. Model 1, Model 2, and Model 3 accuracies were 73 %, 68 %, and 83 %, respectively, according to the findings, with F1-measures of 72 %, 59 %, and 80 %, respectively (refer to Table 4.8). The findings demonstrate that the proposed model improves classification performance significantly. As a result, the suggested methodology produced a finer IQ rating for Wikipedia than other existing methods.

Table 5.1 Comparing the proposed Model's classification performance with Models 1, 2 and 3

| | Model 1 [13] | Model 2 [33] | Model 3 [38] | Proposed model |
|---|---|---|---|---|
| IQ metrics / features | Authority/Reputation<br>Completeness<br>Complexity<br>Informativeness<br>Consistency<br>Currency<br>Volatility | Completeness<br>Informativeness<br>ArticleLength<br>NumHeadings<br>NumRefs/Length | Structural features<br>Readability scores | Expertise<br>Authority<br>Trustworthiness |
| Accuracy | 73% | 68% | 83% | **95%** |
| F1-measure | 72% | 59% | 80% | **94%** |

The model was also tested with four other machine learning methods using cross-validation: Logistic Regression, Support Vector Machine (SVM), Nave Bayes, and KNN, and received accuracy of 96%, 95%, and 94%, respectively. 94%, 96%, 92%, and 94% were the F1-measures, respectively (Table 4.9)

Table 5.2 Comparing the proposed models' classification performance using four existing ML models

| Model | Accuracy | F1 measure |
|---|---|---|
| Logistic Regression | 96% | 94% |
| SVM | 95% | 96% |
| Naïve Bayes | 95% | 92% |
| KNN | 94% | 94% |

### 5.1.2. Clustering performance

Next, model's clustering performance evaluated and compared with previous models. Stvilia et al. [13] used K-means clustering to evaluate the power of their IQ measures in terms of how well they sorted the articles into the correct groupings. As a result, the authors chose to use the same clustering algorithm. Using equation (11), a single numeric IQ value per article was calculated based on the derived values for the first and second order IQ measures. The derived IQ values were then clustered using K-means and the IQ classes to cluster comparisons were observed. Accordingly, the proposed model correctly grouped 93% of the articles into the correct classes. The clustering was repeated with Models 1, 2, and 3, and the results were compared to the proposed model in Table 4.10.

Table 5.3 Comparing the proposed Model's clustering performance with Models 1, 2 and 3

| Classification Type | Model 1 [13] | Model 2 [33] | Model 3 [38] | Proposed model |
|---|---|---|---|---|
| FA correctly classified | 31.4% | 30.3% | 39.6% | **49%** |
| FA incorrectly classified | 15.6% | 18% | 16.7% | 6.4% |
| Non-FA correctly classified | 34.4% | 32% | 33.3% | **43.6%** |
| Non-FA incorrectly classified | 18.6% | 19.7% | 10.4% | 1% |
| Total | 2000 | | | |

According to above Table 4.7, it can be seen that Models 1, 2, and 3 have correctly grouped only 66%, 62%, and 73% of the articles, respectively. Therefore, it can be concluded that the proposed model provides significantly more accurate insight into Wikipedia's IQ assessment than existing models.

### 5.1.3. Inter-rater reliability test

The third evaluation method used by the authors was Fleiss-Davies Kappa [83]. It determines how well a group of raters agree (inter-rater reliability). For the method, three raters were picked who had a prior understanding of IQ and Wikipedia. To begin, they were given 45 IQ qualities and a set of IQ dimensions to group each of the attributes into one of the IQ dimensions, yielding an 80% Kappa value. Second, they were asked to categorize the IQ qualities into three constructs: Expertise, Authority, and Trustworthiness, yielding an 87% Kappa value. The degree of agreement between raters is indicated by Kappa values [74]. Kappa values greater than 0.75 implies strong agreement beyond chance. Since both groups had outstanding rater agreement, with an average Kappa value of 84%, indicating that the proposed model is valid.

# CHAPTER 6 DISCUSSION

This section discusses about the findings of this study in detail, including the implications and suggestions for future research work

## 6.1. Discussion

According to the literature review presented in Chapter 2, it was identified that the three constructs of Google's E-A-T framework i.e., Expertise, Authority and Trustworthiness can be utilized to measure the IQ of Wikipedia articles precisely. The author systematically reviewed the literature under this taxonomy and came up with a conceptual framework to assess the IQ of Wikipedia. In order to validate the framework by answering the established research questions, three hypotheses were populated as H1, H2 and H3.

The RQ1 was 'What are the Information Quality determinants of Wikipedia?'. Under RQ1, three sub research questions were developed as RQ1-a, RQ1-b and RQ1-c. RQ1-a is 'What is the relationship between the Expertise of a Wikipedia article and its Information Quality?' In many studies including [21], [25] [84] [42] [43] [85] authors have tried to assess Expertise of Wikipedia articles in various aspects such as informativeness expertise of editors, concentration and coordination of editors and readability etc. They have also found that these aspects improve the IQ of Wikipedia articles. Out of these aspects, the analysis of this study suggested that Informativeness, Understandability and Readability are the most influencing IQ dimensions for the construct Expertise. Informativeness of a Wikipedia article was defined in this study as, the ability to provide useful information and comprehensive content. Understandability was defined as how well the content is presented so that the user is capable of easily grasping the content. Readability is defined as the amount of effort the reader should exert to read a piece of information. According to the literature, all the aspects of Expertise, improve the IQ of Wikipedia. Accordingly, H1 was populated to answer the RQ1-a. In Chapter 4, based on the regression analysis, H1 was accepted. This was further verified using three evaluation approaches i.e., classification and clustering performance and inter-rater reliability test. Hence, it can be concluded that Expertise in a Wikipedia article is positively associated with its Information Quality.

This implies that Expertise can be considered as one of the most precise constructs to describe the IQ on Wikipedia articles.

Next RQ1-b is ‘What is the relationship between the Authority of a Wikipedia article and its Information Quality?’. Authority was defined as a Wikipedia article getting more powerful as a result of other verticals are citing and rereferring its contents. Studies in [20] [22] [27] [49] present that the Authority of the content improves the IQ of Wikipedia. The H2 was set to observe the association of Authority and IQ. According to the results of regression analysis, H2 was also supported. A positive relationship was observed between Authority and Expertise. Same as discussed in Expertise, it was validated in various means that Authority improves the IQ of Wikipedia articles including empirical findings, study results and the three-phase evaluation approaches. Therefore, Authority also can be used as a key construct to define the IQ of Wikipedia.

RQ1-c is about ‘What is the relationship between Trust of a Wikipedia article and its Information Quality?’. H3 was populated to address this research question. The term Trustworthiness of a Wikipedia article was defined as providing factually accurate and verifiable information. According to many studies such as [28], [50] [28] [29] [51] Trustworthiness is capable of improving IQ. This was found to be true from the results of this study as well since H3 was supported in the regression analysis. Further, it was found that reliability, verifiability and maturity are the most influencing IQ dimensions for Trustworthiness. The validity of this statement was further evaluated, and the evaluation results also supported the claim. Hence, it can be concluded that Trustworthiness can be considered as the other key construct for assessing the IQ of Wikipedia.

Moreover, according to the literature review, there were several limitations of the existing work related to the IQ assessment of Wikipedia. Firstly, compared to Expertise and Trustworthiness, much attention to Authority towards IQ has not been paid off. It is clear that a high web presence leads to a high reputation because other experts are citing a web page/website. Analogously, a good reputation improves the quality of the content [51], [69]. Towards bridging this gap, this study extensively studies how Authority positively affects the IQ of Wikipedia. In the proposed model, it was observed that a significant accuracy drops when the Authority construct was removed (reducing the accuracy by 16%). Accordingly, the author claims that authority should be considered as a major construct to measure the IQ of Wikipedia articles. Secondly, there is a lack of studies on how maturity and verifiability affect IQ [12]. In the proposed theoretical framework, this study introduced maturity and verifiability among the key dimensions that influence the construct of Trustworthiness. Finally, even if there are models and frameworks that have been proposed to assess IQ, they have failed to provide satisfactory classification performance for classifying articles into corresponding classes according to their IQ [13], [38], [41]. Filling this gap, the proposed model outperformed the existing frameworks by producing 95%, 94% and 84% of classification accuracy, clustering accuracy and inter-rater reliability respectively.

## 6.2. Implications of the study

There are several important implications of this study that need to be discussed. In terms of theoretical contribution, this study introduces a theoretical model based on Google's E-A-T model to assess the IQ of Wikipedia articles comprising of three IQ constructs as Expertise, Authority and Trustworthiness, a set of IQ dimensions that influence the above IQ constructs and 45 IQ attributes to measure the above dimensions. The model was extensively validated using a three-phase evaluation approach and was compared with three previous well-known models. The evaluation results found that the proposed model outperformed all the other compared models yielding 95% of classification accuracy, 93% of clustering accuracy and 84% of inter-rater agreement. Further, with the intension of disseminating this knowledge with scholary community, author has published three research articles. Details about these publications are given in Appendix B.

There are important practical implications of this study. The findings of this study help users identify high-quality Wikipedia articles.  Rather than dismissing Wikipedia as an unreliable source of knowledge in the big scope, this approach may be used to extract high-quality, reliable material from Wikipedia, the world's biggest publicly accessible online information repository. On the other hand, the study proposes an automated technique to identify article defects based on the suggested model, which would give instant assistance for reviewers, writers, and editors to adopt quality changes in the

Wikipedia community, which requires a lot of time and effort. Furthermore, this methodology should be utilized as an initial screening tool and a supporting tool throughout the formal review process, not as a replacement for the present manual reviewing procedure.

### 6.3. Suggestions for Further Research

First, the dataset was limited to six domains: Medicine, Politics, Sports, History, Science and Biographies and comprises 2000 data points. The study results can be further generalized by improving the dataset by size and domains. The second limitation is, the study is based on only English language. There are over 300 language versions of Wikipedia. Thus, similar studies could be conducted for those languages as well. However, since the language component was specifically developed for English language Wikipedia, it would not be directly used for other languages. Thus, except the language component, other components can be readily used with any other langage Wikipedia. Thirdly, the proposed E-A-T based model could be used to assess the IQ of other collaborative platforms by modifying and adding new IQ attributes using hybrid approach.

# CHAPTER 7 CONCLUSION

Wikipedia being one the largest collaborative encylopedia suffers from significant issues in its IQ. Even if previous studies have tried to look into this problem, a concrete and genralisable framework was lacking. Further, the accuracy of those approaches was not satisfactory. Therefore, this study tried to develop an IQ assessment framework for English language version of Wikipedia. For this, a comprehensive literature review was conducted and various IQ dimensions and IQ attributes were identified. Out of these dimensions, author identified Expertise, Authority and Trustworthiness can be considered as key constructs for Wikipedia’s IQ. Based on the empirical findings, the study proposed a theoretical model based on Google’s E-A-T model to assess the Wikipedia article's IQ. Furthermore, a set of IQ dimensions that influence the above three IQ constructs, as well as 45 IQ attributes to measure the IQ dimensions, were identified. The IQ attributes were automatically and inexpensively extracted from the content and metadata statistics of Wikipedia articles using a Selenium 3.14 web automation script. A sample of 2000 articles comprising 1000 Featured Articles (FA) and 1000 non-FA articles from six WikiProjects was used for the data analysis. It was observed and comprehensively validated that, the proposed model provides better IQ evaluation for Wikipedia than the state-of-the-art approaches. It directs users to steer into high quality information and reviewers to figure out flaws and arrange necessary quality improvements. Further, it is worth to mention that, the hybrid approach of article of feature extraction which involves both content and meta-data statistics immensely help to to construct and evaluate the IQ variance of any online collaborative resource with low cost and scalabale way. Finally, the findings of this study provide good insights towards assessing the IQ of Wikipedia articles and efficient knowledge collaboration in the Wikipedia community.

## REFERENCES


[1] “Wikipedia - Wikipedia.” https://en.wikipedia.org/wiki/Wikipedia (accessed May 26, 2022).

[2] “Global top websites by unique visits 2021 | Statista.” https://www.statista.com/statistics/1201889/most-visited-websites-worldwide-unique-visits/ (accessed Aug. 21, 2021).

[3] “Wikipedia Seigenthaler biography incident - Wikipedia.” https://en.wikipedia.org/wiki/Wikipedia_Seigenthaler_biography_incident (accessed Aug. 21, 2021).

[4] H. Meishar-Tal, “Teachers’ use of Wikipedia with their Students,” *Australian Journal of Teacher Education*, vol. 40, no. 12, p. 40, 2015, doi: 10.14221/ajte.2015v40n12.9.

[5] “School officials unite in banning Wikipedia | The Seattle Times.” https://www.seattletimes.com/life/lifestyle/school-officials-unite-in-banning-wikipedia/ (accessed May 04, 2021).

[6] P. Konieczny, “Rethinking Wikipedia for the Classroom,” *Contexts*, vol. 13, no. 1, pp. 80–83, Feb. 2014, doi: 10.1177/1536504214522017.

[7] “Wikipedia:Content assessment - Wikipedia.” https://en.wikipedia.org/wiki/Wikipedia:Content_assessment (accessed May 27, 2022).

[8] S. Bykau, F. Korn, D. Srivastava, and Y. Velegrakis, “Fine-grained controversy detection in Wikipedia,” *Proceedings - International Conference on Data Engineering*, vol. 2015-May, pp. 1573–1584, May 2015, doi: 10.1109/ICDE.2015.7113426.

[9] Q. Dang, C. Ignat, Q. Dang, C. I. Measuring, C. Edited, and Q. Dang, “Measuring quality of collaboratively edited documents: the case of Wikipedia,” in *IEEE 2nd International Conference on Collaboration and Internet Computing (CIC)*, 2016, pp. 266–275. doi: 10.1109/CIC.2016.044.

[10] D. H. Dalip, M. A. Gonçalves, M. Cristo, and P. Calado, “A general multiview framework for assessing the quality of collaboratively created content on web 2.0,” *J Assoc Inf Sci Technol*, vol. 68, no. 2, pp. 286–308, Feb. 2017, doi: 10.1002/ASI.23650.

[11] K. D. C. G. Kapugama, S. A. S. Lorensuhewa, and M. A. L. Kalyani, “Enhancing Wikipedia search results using Text Mining,” in *16th International Conference on Advances in ICT for Emerging Regions (ICTer)*, 2016, pp. 168–175. doi: 10.1109/ICTER.2016.7829915.

[12] M. Anderka, B. Stein, and N. Lipka, “Predicting quality flaws in user-generated content: The case of wikipedia,” in *SIGIR’12 - Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2012, pp. 981–990. doi: 10.1145/2348283.2348413.

[13] B. Stvilia, M. B. Twidale, L. C. Smith, and L. Gasser, “Assessing Information Quality of a Community-Based Encyclopedia,” in *Proceedings of the International Conference on Information Quality - ICIQ*, 2005, pp. 442–454.

[14] M. Gustavsson and C. Wänström, “Assessing information quality in manufacturing planning and control processes,” *International Journal of Quality and Reliability Management*, vol. 26, no. 4, pp. 325–340, Apr. 2009, doi: 10.1108/02656710910950333.

[15] D. P. Ballou and G. K. Tayi, “Enhancing data quality in data warehouse environments,” *Commun ACM*, vol. 42, no. 1, pp. 73–78, 1999, doi: 10.1145/291469.291471.

[16] J. A. Hoffer, M. B. Prescott, and F. R. McFadden, *Modern database management*. Pearson/ Prentice Hall, 2005.

[17] R. Y. Wang and D. M. Strong, “Beyond Accuracy: What Data Quality Means to Data Consumers,” *Source: Journal of Management Information Systems*, vol. 12, no. 4, pp. 5–33, 1996.

[18] K. P. Fadahunsi *et al.*, "Protocol for a systematic review and qualitative synthesis of information quality frameworks in eHealth," *BMJ Open*, vol. 9, no. 3, p. e024722, Mar. 2019, doi: 10.1136/BMJOPEN-2018-024722.

[19] *Encyclopædia Britannica*, 15th ed. 2007.

[20] B. Thomas Adler, L. de Alfaro, S. M. Mola-Velasco, P. Rosso, and A. G. West, "Wikipedia Vandalism Detection: Combining Natural Language, Metadata, and Reputation Features," in *International Conference on Intelligent Text Processing and Computational Linguistics*, 2011, vol. 6609, pp. 277–288. doi: 10.1007/978-3-642-19437-5_23.

[21] A. Kittur and R. E. Kraut, "Harnessing the wisdom of crowds in wikipedia: Quality through coordination," *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, pp. 37–46, 2008, doi: 10.1145/1460563.1460572.

[22] K. Stein and C. Hess, "Does it matter who contributes - A study on featured articles in the german wikipedia," in *Proceedings of the Eighteenth ACM Conference on Hypertext and Hypermedia, HT'07*, 2007, pp. 171–174. doi: 10.1145/1286240.1286290.

[23] D. M. Wilkinson and B. A. Huberman, "Cooperation and quality in Wikipedia," *Proceedings of the Conference on Object-Oriented Programming Systems, Languages, and Applications, OOPSLA*, no. October 2007, pp. 157–164, 2007, doi: 10.1145/1296951.1296968.

[24] Y. Suzuki, "Quality assessment of wikipedia articles using h-index," *Journal of Information Processing*, vol. 23, no. 1, pp. 22–30, Jan. 2015, doi: 10.2197/IPSJJIP.23.22.

[25] O. Arazy and O. Nov, "Determinants of wikipedia quality: The roles of global and local contribution inequality," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, 2010, pp. 233–236. doi: 10.1145/1718918.1718963.

[26] J. Liu and S. Ram, "Who Does What: Collaboration Patterns in the Wikipedia and Their Impact on Article Quality," *ACM Trans. Manag. Inform. Syst*, vol. 2, no. 2, 2011, doi: 10.1145/1985347.1985352.

[27] K. Nemoto, P. A. Gloor, and R. Laubacher, "Social capital increases efficiency of collaboration among Wikipedia editors," in *HT 2011 - Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia*, 2011, pp. 231–240. doi: 10.1145/1995966.1995997.

[28] H. Zeng, M. A. Alhossaini, L. Ding, R. Fikes, and D. L. McGuinness, "Computing trust from revision history," in *Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services*, 2006, p. 1. doi: 10.1145/1501434.1501445.

[29] J. M. Nicholson, A. Uppala, M. Sieber, P. Grabitz, M. Mordaunt, and S. Rife, "Measuring the quality of scientific references in Wikipedia: an analysis of more than 115M citations to over 800,000 scientific articles," 2020. doi: 10.1101/2020.04.08.031765.

[30] J. E. Blumenstock, "Size matters: Word count as a measure of quality on Wikipedia," in *Proceeding of the 17th International Conference on World Wide Web 2008, WWW'08*, 2008, pp. 1095–1096. doi: 10.1145/1367497.1367673.

[31] N. Lipka and B. Stein, "Identifying featured articles in Wikipedia: Writing style matters," in *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, 2010, pp. 1147–1148. doi: 10.1145/1772690.1772847.

[32] T. Wöhner and R. Peters, "Assessing the quality of Wikipedia articles with lifecycle based metrics," 2009. doi: 10.1145/1641309.1641333.

[33] M. Warncke-Wang, D. Cosley, and J. Riedl, "Tell me more: An actionable quality model for wikipedia," in *WikiSym '13: Proceedings of the 9th International Symposium on Open Collaboration*, 2013, pp. 1–10. doi: 10.1145/2491055.2491063.

[34] Q. V. Dang and C. L. Ignat, “An end-to-end learning solution for assessing the quality of Wikipedia articles,” Aug. 2017. doi: 10.1145/3125433.3125448.

[35] M. Warncke-Wang, V. R. Ayukaev, B. Hecht, and L. Terveen, “The success and failure of Quality Improvement Projectsin peer production communities,” in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, Feb. 2015, pp. 743–756. doi: 10.1145/2675133.2675241.

[36] L. R. Jimenez, M. Solera, M. Toril, C. Gijon, and P. Casas, “Content Matters: Clustering Web Pages for QoE Analysis With WebCLUST,” *IEEE Access*, vol. 9, pp. 123873–123888, 2021, doi: 10.1109/ACCESS.2021.3110370.

[37] D. H. Dalip, M. A. Gonçalves, M. Cristo, and P. Calado, “Automatic quality assessment of content created collaboratively by web communities: A case study of wikipedia,” in *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries*, 2009, pp. 295–304. doi: 10.1145/1555400.1555449.

[38] A. Shen, J. Qi, and T. Baldwin, “A Hybrid Model for Quality Assessment of Wikipedia Articles,” in *Proceedings of the Australasian Language Technology Association Workshop*, 2017, pp. 43–52.

[39] Google, “Search Quality Evaluator Guidelines,” 2018.

[40] J. Grundmann, “Google E-A-T: How to Improve your E-A-T Ranking Score.” https://blog.searchmetrics.com/us/google-e-a-t-ranking-factor/ (accessed Jul. 10, 2022).

[41] D. M. Wilkinson and B. A. Huberman, “Assessing the value of cooperation in Wikipedia,” *First Monday*, vol. 12, no. 4, 2007, doi: 10.5210/fm.v12i4.1763.

[42] N. T. Korfiatis, M. Poulos, and G. Bokos, “Evaluating authoritative sources using social networks: An insight from Wikipedia,” *Online Information Review*, vol. 30, no. 3, pp. 252–262, 2006, doi: 10.1108/14684520610675780.

[43] G. Wu, M. Harrigan, and P. Cunningham, “Classifying Wikipedia articles using network motif counts and ratios,” in *In Proceedings of the Eighth Annual*

*International Symposium on Wikis and Open Collaboration*, 2012, pp. 1–10. doi: 10.1145/2462932.2462948.

[44] M. A. Suryanto, E. P. Lim, A. Sun, and R. H. L. Chiang, “Quality-aware collaborative Question Answering: Methods and evaluation,” in *Proceedings of the 2nd ACM International Conference on Web Search and Data Mining, WSDM’09*, 2009, pp. 142–151. doi: 10.1145/1498759.1498820.

[45] S. Chinthani, A. Supunmali, and S. Ahangama, “Modelling Wikipedia ’ s Information Quality using Informativeness , Reliability and Authority,” 2021.

[46] Q. V. Dang *et al.*, “Quality Assessment of Wikipedia Articles without Feature Engineering To cite this version : HAL Id : hal-01351226 Quality Assessment of Wikipedia Articles without Feature Engineering,” 2016.

[47] Y. Javed, K. M. Salehin, and M. Shehab, “A Study of South Asian Websites on Privacy Compliance,” *IEEE Access*, vol. 8, pp. 156067–156083, 2020, doi: 10.1109/ACCESS.2020.3019334.

[48] Ian Booth, “Google EAT: How to Create Content That Google Wants - Moz.” https://moz.com/blog/google-e-a-t (accessed Feb. 26, 2021).

[49] D. Lewandowski and U. Spree, “Ranking of Wikipedia articles in search engines revisited: Fair ranking for reasonable quality?,” *Journal of the American Society for Information Science and Technology*, vol. 62, no. 1, pp. 117–132, Jan. 2011, doi: 10.1002/ASI.21423.

[50] A. Sarabadani, A. Halfaker, and D. Taraborelli, “Building automated vandalism detection tools for wikidata,” in *26th International World Wide Web Conference 2017, WWW 2017 Companion*, 2017, pp. 1647–1654. doi: 10.1145/3041021.3053366.

[51] W. Lewoniewski, K. Węcel, and W. Abramowicz, “Relative Quality and Popularity Evaluation of Multilingual Wikipedia Articles,” *Informatics*, vol. 4, no. 4, p. 43, 2017, doi: 10.20944/preprints201709.0130.v1.

[52] B. Stvilia, M. B. Twidale, L. C. Smith, and L. Gasser, “Information quality work organization in Wikipedia,” *Journal of the American Society for Information Science and Technology*, vol. 59, no. 6, pp. 983–1001, 2008, doi: 10.1002/asi.20813.

[53] R. Flesch, “A new readability yardstick,” *Journal of Applied Psychology*, vol. 32, no. 3, pp. 221–233, Jun. 1948, doi: 10.1037/h0057532.

[54] J. Kincaid, R. Fishburne, R. Rogers, and B. Chissom, “Derivation Of New Readability Formulas (Automated Readability Index, Fog Count And Flesch Reading Ease Formula) For Navy Enlisted Personnel,” *Institute for Simulation and Training*, Jan. 1975.

[55] R. Gunning, “The fog index after twenty years,” *Journal of Business Communication*, vol. 6, no. 2, pp. 3–13, Sep. 1969, doi: 10.1177/002194366900600202.

[56] G. H. McLaughlin, “SMOG grading: A new readability formula. - PsycNET,” *Journal of Reading*, vol. 12, no. 8, pp. 639–646, 1969.

[57] M. Coleman and T. L. Liau, “A computer readability formula designed for machine scoring,” *Journal of Applied Psychology*, vol. 60, no. 2, pp. 283–284, Apr. 1975, doi: 10.1037/h0076540.

[58] R. J. Senter and E. A. Smith, “AUTOMATED READABILITY INDEX.” Nov. 01, 1967.

[59] “Learn SEO — Free SEO Learning Center [2021] - Moz.” https://moz.com/learn/seo (accessed Feb. 26, 2021).

[60] “Readable | Free Readability Test Tool.” https://www.webfx.com/tools/read-able/ (accessed Jun. 09, 2021).

[61] “Help:Minor edit - Wikipedia.” https://en.wikipedia.org/wiki/Help:Minor_edit (accessed Jun. 07, 2021).

[62] “Wikipedia:Bots - Wikipedia.” https://en.wikipedia.org/wiki/Wikipedia:Bots (accessed Jun. 07, 2021).

[63] "Automated edits - OpenStreetMap Wiki." https://wiki.openstreetmap.org/wiki/Automated_edits (accessed Jun. 14, 2021).

[64] "Edit count - Wikipedia." https://en.wikipedia.org/wiki/Edit_count (accessed Jun. 14, 2021).

[65] "Help:Page size - Meta." https://meta.wikimedia.org/wiki/Help:Page_size (accessed Jun. 14, 2021).

[66] "Help:Watchlist - Wikipedia." https://en.wikipedia.org/wiki/Help:Watchlist#watchers (accessed Oct. 05, 2021).

[67] "Wikipedia:Pageview statistics - Wikipedia." https://en.wikipedia.org/wiki/Wikipedia:Pageview_statistics (accessed Oct. 05, 2021).

[68] "Help:Editing - Wikipedia." https://en.wikipedia.org/wiki/Help:Editing (accessed Oct. 05, 2021).

[69] C. Sugandhika, S. Ahangama, and S. Ahangama, "Assessing the information quality of Wikipedia articles using web presence," in *International Conference on Innovation and Emerging Technologies*, 2021, p. 37.

[70] "Selenium." Accessed: Sep. 24, 2021. [Online]. Available: https://www.selenium.dev/

[71] "Readability Test | WebFX." https://www.webfx.com/tools/read-able/ (accessed May 27, 2022).

[72] "Learn SEO | Grow Your Knowledge, Advance Your Career - Moz." https://moz.com/learn/seo (accessed May 27, 2022).

[73] "Wikipedia:WikiProject Medicine - Wikipedia." https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Medicine (accessed Jun. 09, 2021).

[74] "Wikipedia:WikiProject Politics - Wikipedia." https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Politics (accessed Sep. 14, 2021).

[75] "Wikipedia:WikiProject Sports - Wikipedia." https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Sports (accessed Sep. 14, 2021).

[76] "Wikipedia:WikiProject History - Wikipedia." https://en.wikipedia.org/wiki/Wikipedia:WikiProject_History (accessed Sep. 14, 2021).

[77] "Wikipedia:WikiProject Science - Wikipedia." https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Science (accessed Sep. 14, 2021).

[78] "Wikipedia:WikiProject Biography/Assessment - Wikipedia." https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Biography/Assessment (accessed Sep. 24, 2021).

[79] A. Field, *Discovering Statistics Using IBM SPSS Statistics* , 4th ed. New Delhi: SAGE Publications India Pvt Ltd, 2009.

[80] B. J. Babin and R. E. Anderson, *on Multivariate Data Analysis Joseph F . Hair Jr . William C . Black Seventh Edition*. 2014.

[81] P. Wang, X. Li, and R. Wu, "A deep learning-based quality assessment model of collaboratively edited documents: A case study of Wikipedia," *Journal of Information Science*, vol. 47, no. 2, pp. 176–191, Apr. 2021, doi: 10.1177/0165551519877646.

[82] E. Bassani and M. Viviani, "Automatically assessing the quality of Wikipedia contents," *Proceedings of the ACM Symposium on Applied Computing*, vol. Part F147772, pp. 804–807, 2019, doi: 10.1145/3297280.3297357.

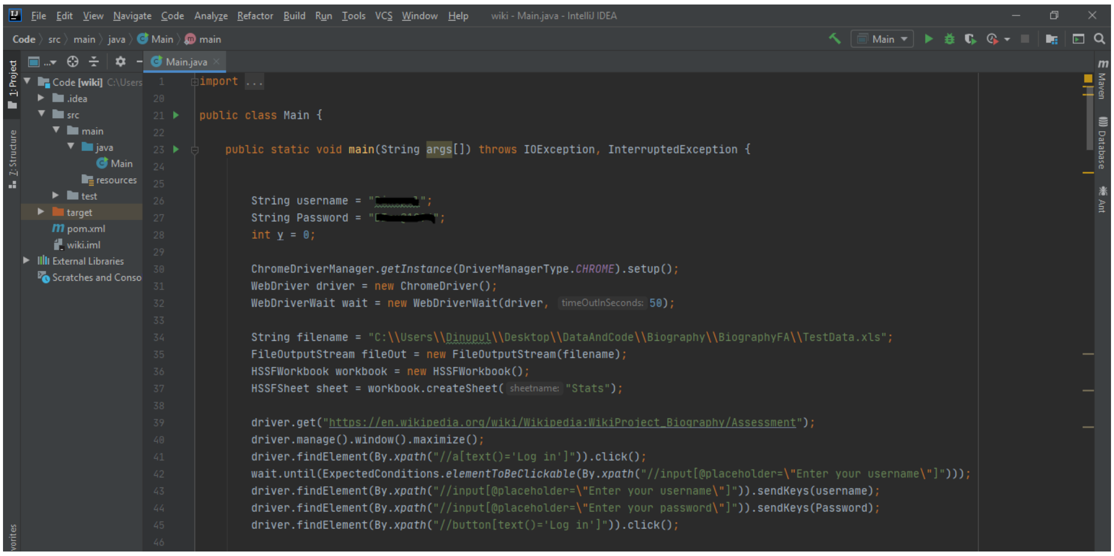
[83] M. Banerjee, M. Capozzoli, L. McSweeney, and D. Sinha, “Beyond kappa: A review of interrater agreement measures,” *Canadian Journal of Statistics*, vol. 27, no. 1, pp. 3–23, Mar. 1999, doi: 10.2307/3315487.

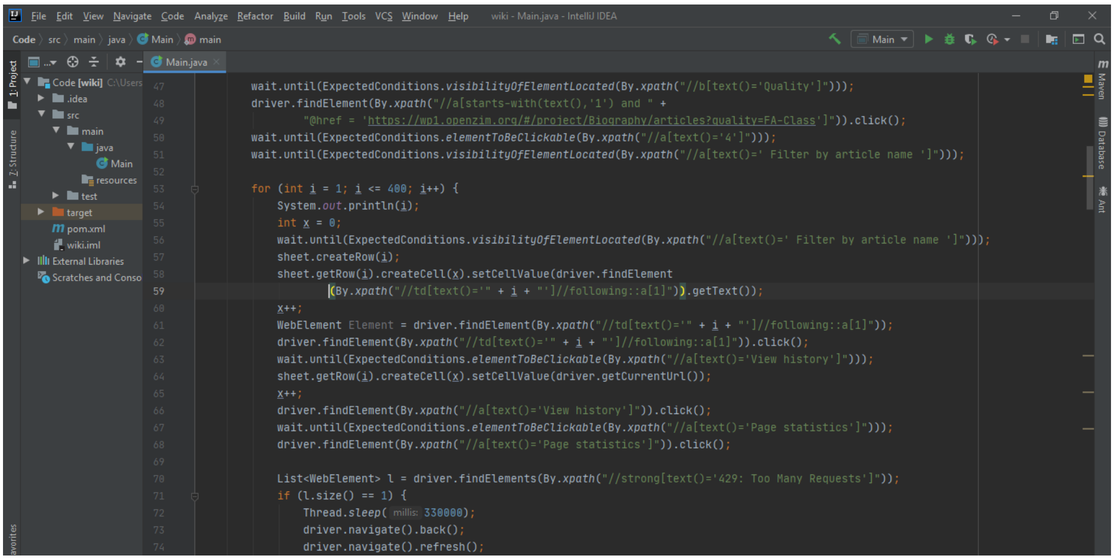
[84] M. Hu, E. P. Lim, A. Sun, H. W. Lauw, and B. Q. Vuong, “Measuring article quality in wikipedia: Models and evaluation,” in *International Conference on Information and Knowledge Management, Proceedings*, 2007, pp. 243–252. doi: 10.1145/1321440.1321476.

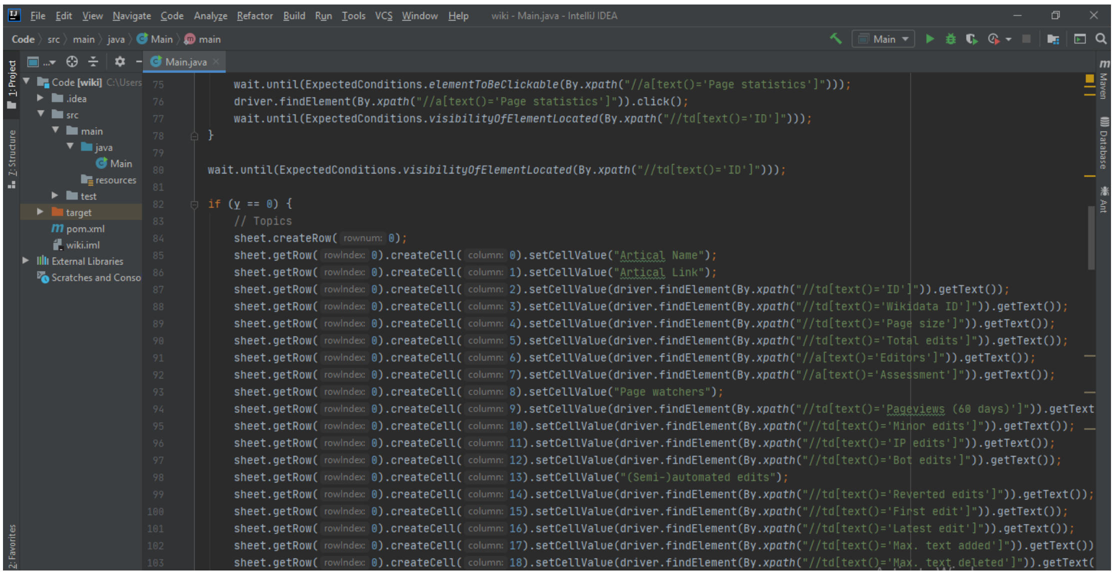
[85] H. Zhang, Y. Ren, and R. Kraut, “Mining and Predicting Temporal Patterns in the Quality Evolution of Wikipedia Articles,” in *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020, pp. 1–10. doi: 10.24251/hicss.2020.485.

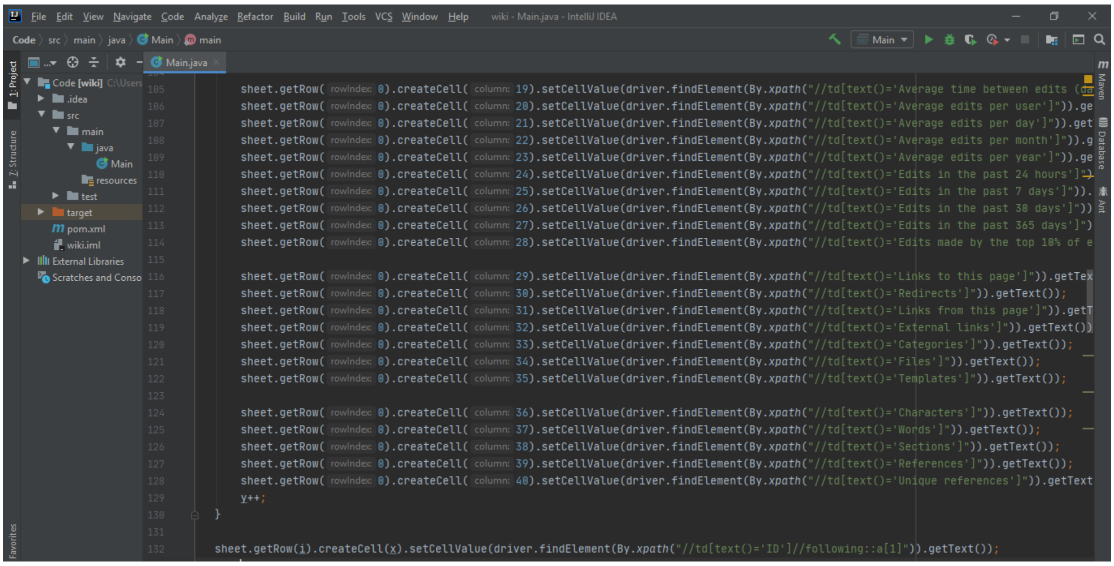
# APPENDIX A

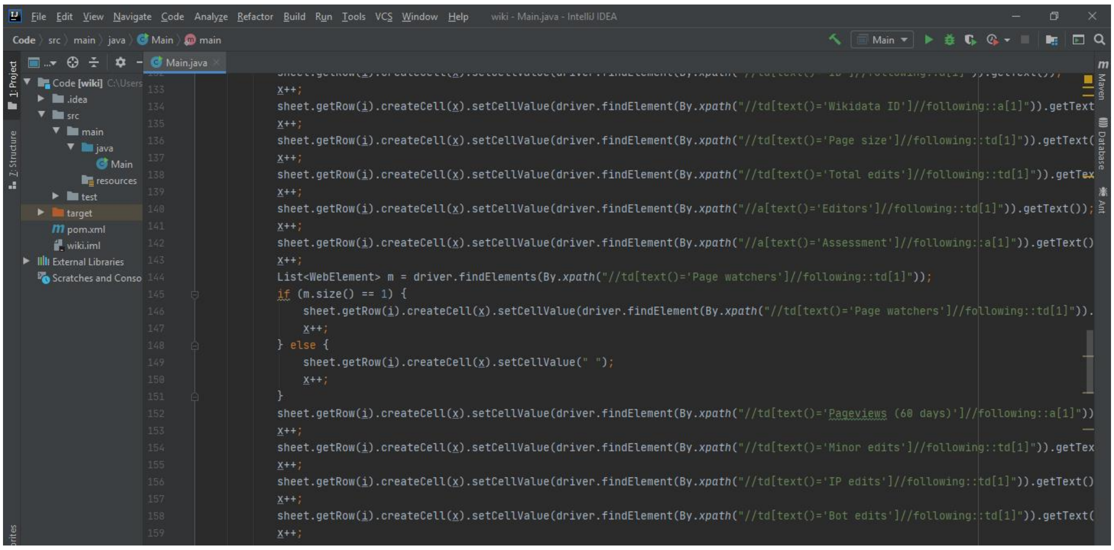
Sample code snippets were written to generate data from Wikipedia articles.

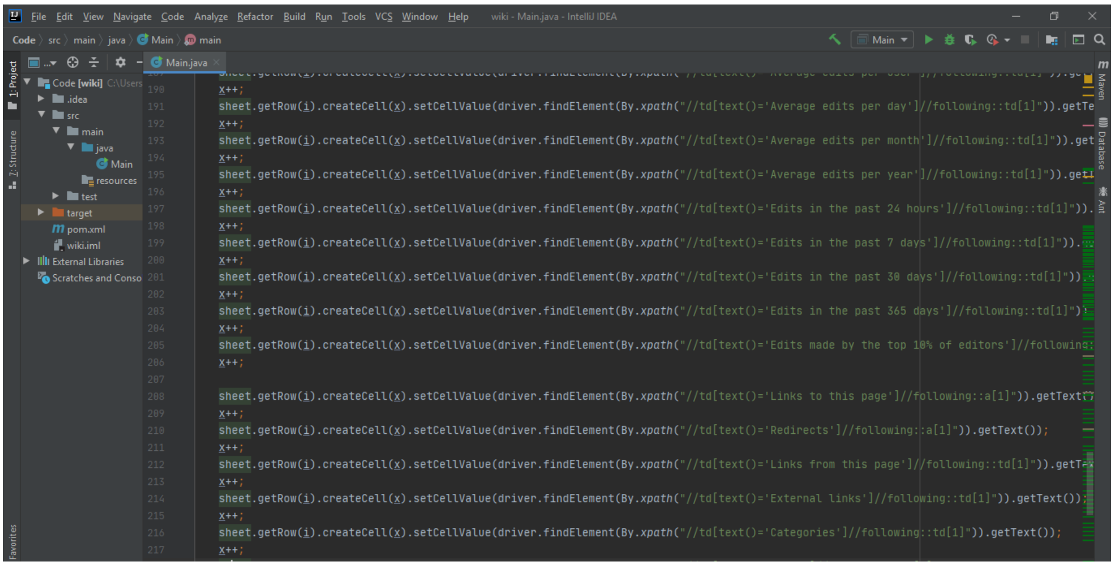


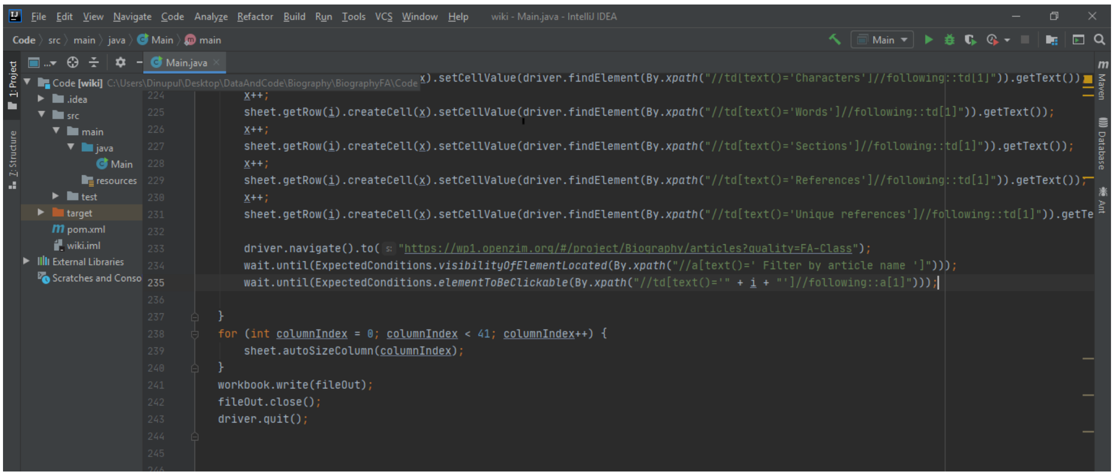




wiki - Main.java - IntelliJ IDEA
Main.java
wait.until(ExpectedConditions.elementToBeClickable(By.xpath("//a[text()='Page statistics']")));
driver.findElement(By.xpath("//a[text()='Page statistics']")).click();
wait.until(ExpectedConditions.visibilityOfElementLocated(By.xpath("//td[text()='ID']")));
}
wait.until(ExpectedConditions.visibilityOfElementLocated(By.xpath("//td[text()='ID']")));
if (y == 0) {
// Topics
sheet.createRow( rownum: 0);
sheet.getRow( rowIndex: 0).createCell( column: 0).setCellValue("Artical Name");
sheet.getRow( rowIndex: 0).createCell( column: 1).setCellValue("Artical Link");
sheet.getRow( rowIndex: 0).createCell( column: 2).setCellValue(driver.findElement(By.xpath("//td[text()='ID']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 3).setCellValue(driver.findElement(By.xpath("//td[text()='Wikidata ID']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 4).setCellValue(driver.findElement(By.xpath("//td[text()='Page size']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 5).setCellValue(driver.findElement(By.xpath("//td[text()='Total edits']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 6).setCellValue(driver.findElement(By.xpath("//a[text()='Editors']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 7).setCellValue(driver.findElement(By.xpath("//a[text()='Assessment']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 8).setCellValue("Page watchers");
sheet.getRow( rowIndex: 0).createCell( column: 9).setCellValue(driver.findElement(By.xpath("//td[text()='Pageviews (60 days)']")).getText
sheet.getRow( rowIndex: 0).createCell( column: 10).setCellValue(driver.findElement(By.xpath("//td[text()='Minor edits']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 11).setCellValue(driver.findElement(By.xpath("//td[text()='IP edits']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 12).setCellValue(driver.findElement(By.xpath("//td[text()='Bot edits']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 13).setCellValue("(Semi-)automated edits");
sheet.getRow( rowIndex: 0).createCell( column: 14).setCellValue(driver.findElement(By.xpath("//td[text()='Reverted edits']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 15).setCellValue(driver.findElement(By.xpath("//td[text()='First edit']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 16).setCellValue(driver.findElement(By.xpath("//td[text()='Latest edit']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 17).setCellValue(driver.findElement(By.xpath("//td[text()='Max. text added']")).getText())
sheet.getRow( rowIndex: 0).createCell( column: 18).setCellValue(driver.findElement(By.xpath("//td[text()='Max. text deleted']")).getText(


wiki - Main.java - IntelliJ IDEA
Main.java
sheet.getRow( rowIndex: 0).createCell( column: 19).setCellValue(driver.findElement(By.xpath("//td[text()='Average time between edits (da
sheet.getRow( rowIndex: 0).createCell( column: 20).setCellValue(driver.findElement(By.xpath("//td[text()='Average edits per user']")).ge
sheet.getRow( rowIndex: 0).createCell( column: 21).setCellValue(driver.findElement(By.xpath("//td[text()='Average edits per day']")).get
sheet.getRow( rowIndex: 0).createCell( column: 22).setCellValue(driver.findElement(By.xpath("//td[text()='Average edits per month']")).g
sheet.getRow( rowIndex: 0).createCell( column: 23).setCellValue(driver.findElement(By.xpath("//td[text()='Average edits per year']")).ge
sheet.getRow( rowIndex: 0).createCell( column: 24).setCellValue(driver.findElement(By.xpath("//td[text()='Edits in the past 24 hours']")
sheet.getRow( rowIndex: 0).createCell( column: 25).setCellValue(driver.findElement(By.xpath("//td[text()='Edits in the past 7 days']")).
sheet.getRow( rowIndex: 0).createCell( column: 26).setCellValue(driver.findElement(By.xpath("//td[text()='Edits in the past 30 days']"))
sheet.getRow( rowIndex: 0).createCell( column: 27).setCellValue(driver.findElement(By.xpath("//td[text()='Edits in the past 365 days']")
sheet.getRow( rowIndex: 0).createCell( column: 28).setCellValue(driver.findElement(By.xpath("//td[text()='Edits made by the top 10% of e
sheet.getRow( rowIndex: 0).createCell( column: 29).setCellValue(driver.findElement(By.xpath("//td[text()='Links to this page']")).getTex
sheet.getRow( rowIndex: 0).createCell( column: 30).setCellValue(driver.findElement(By.xpath("//td[text()='Redirects']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 31).setCellValue(driver.findElement(By.xpath("//td[text()='Links from this page']")).getT
sheet.getRow( rowIndex: 0).createCell( column: 32).setCellValue(driver.findElement(By.xpath("//td[text()='External links']")).getText())
sheet.getRow( rowIndex: 0).createCell( column: 33).setCellValue(driver.findElement(By.xpath("//td[text()='Categories']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 34).setCellValue(driver.findElement(By.xpath("//td[text()='Files']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 35).setCellValue(driver.findElement(By.xpath("//td[text()='Templates']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 36).setCellValue(driver.findElement(By.xpath("//td[text()='Characters']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 37).setCellValue(driver.findElement(By.xpath("//td[text()='Words']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 38).setCellValue(driver.findElement(By.xpath("//td[text()='Sections']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 39).setCellValue(driver.findElement(By.xpath("//td[text()='References']")).getText());
sheet.getRow( rowIndex: 0).createCell( column: 40).setCellValue(driver.findElement(By.xpath("//td[text()='Unique references']")).getText
y++;
}
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='ID']//following::a[1]")).getText());


wiki - Main.java - IntelliJ IDEA
Main.java
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Wikidata ID']//following::a[1]")).getText
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Page size']//following::td[1]")).getText(
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Total edits']//following::td[1]")).getTex
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//a[text()='Editors']//following::td[1]")).getText());
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//a[text()='Assessment']//following::a[1]")).getText()
x++;
List<WebElement> m = driver.findElements(By.xpath("//td[text()='Page watchers']//following::td[1]"));
if (m.size() == 1) {
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Page watchers']//following::td[1]")).
x++;
} else {
sheet.getRow(i).createCell(x).setCellValue(" ");
x++;
}
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Pageviews (60 days)']//following::a[1]"))
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Minor edits']//following::td[1]")).getTex
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='IP edits']//following::td[1]")).getText()
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Bot edits']//following::td[1]")).getText(
x++;


File Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help wiki - Main.java - IntelliJ IDEA
Code src main java Main main
Main.java
Code [wiki]
.idea
src
main
java
Main
resources
test
target
pom.xml
wiki.iml
External Libraries
Scratches and Conso
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Average edits per day']//following::td[1]")).getTe
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Average edits per month']//following::td[1]")).get
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Average edits per year']//following::td[1]")).get
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Edits in the past 24 hours']//following::td[1]")).
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Edits in the past 7 days']//following::td[1]")).
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Edits in the past 30 days']//following::td[1]"))
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Edits in the past 365 days']//following::td[1]")
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Edits made by the top 10% of editors']//following
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Links to this page']//following::a[1]")).getText(
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Redirects']//following::a[1]")).getText());
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Links from this page']//following::td[1]")).getTe
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='External links']//following::td[1]")).getText());
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Categories']//following::td[1]")).getText());
x++;


File Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help wiki - Main.java - IntelliJ IDEA
Code src main java Main main
Main.java
Code [wiki] C:\Users\Dinupul\Desktop\DataAndCode\Biography\BiographyFA\Code
.idea
src
main
java
Main
resources
test
target
pom.xml
wiki.iml
External Libraries
Scratches and Conso
x).setCellValue(driver.findElement(By.xpath("//td[text()='Characters']//following::td[1]")).getText())
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Words']//following::td[1]")).getText());
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Sections']//following::td[1]")).getText());
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='References']//following::td[1]")).getText());
x++;
sheet.getRow(i).createCell(x).setCellValue(driver.findElement(By.xpath("//td[text()='Unique references']//following::td[1]")).getTe
driver.navigate().to("https://wp1.openzim.org/#/project/Biography/articles?quality=FA-Class");
wait.until(ExpectedConditions.visibilityOfElementLocated(By.xpath("//a[text()=' Filter by article name ']")));
wait.until(ExpectedConditions.elementToBeClickable(By.xpath("//td[text()='" + i + "']//following::a[1]")));
}
for (int columnIndex = 0; columnIndex < 41; columnIndex++) {
sheet.autoSizeColumn(columnIndex);
}
workbook.write(fileOut);
fileOut.close();
driver.quit();

# APPENDIX B

Publications of the author about this study are listed below.

1. **Journal paper**

C. Sugandhika and S. Ahangama, "Assessing Information Quality of Wikipedia Articles Through Google's E-A-T Model," in *IEEE Access*, vol. 10, pp. 52196-52209, 2022, doi: 10.1109/ACCESS.2022.3172962.

2. **Conference paper**

C. Sugandhika, S. Ahangama and S. Ahangama, "Modelling Wikipedia's Information Quality using Informativeness, Reliability and Authority," *3rd International Conference on Advancements in Computing (ICAC)*, 2021, pp. 169-174, doi: 10.1109/ICAC54203.2021.9671092.

Awards - Overall Best Research Paper of the Conference, Best Research Paper of the Track General

3. **Abstract**

C. Sugandhika, S. Ahangama and S. Ahangama, "Assessing the information quality of Wikipedia articles using web presence", *Proceedings of International Conference on Innovation and Emerging Technologies*., pp. 37, 2021.

Awards - Outstanding Oral Presentation-1st Place, Best Oral Presentation-Postgraduate under the track, Artificial Intelligence, Machine Learning, and Data Science