TASKFORCE FRAMEWORK IN VALIDATING AUTHENTICITY OF UNIVERSITY INFORMATION OVER THE WIKIPEDIA

Hayati Abd Rahman[1*], Zainab Mohd Noor[2], Amirul Abd Rashid[3], Roziah Mohd Janor[4]

Universiti Teknologi MARA (UiTM), Shah Alam, MALAYSIA
[1*]hayatiar@tmsk.uitm.edu.my, [2]zaina847@uitm.edu.my, [3]amirul2550@uitm.edu.my,
[4]roziahmj@uitm.edu.my

ABSTRACT

Wikipedia is considered a globally primary resource for public review. As it is the largest, free, and most popular general reference work on the World Wide Web, it is easy for web crawlers to do their search indexes and improve the visibility of the page in the digital world. To sustain its top ranking in organic search and display of any search engines, Wikipedia attracts many links and contains a large amount of content on a broad set of topics. Document about Universiti Teknologi MARA (UiTM) started in Wikipedia since 2004, and it has been consistently evolving since then. UiTM has many branch campuses throughout the nation. It is Malaysia's largest institution of higher learning in terms of size and population, with the anchor campus in Shah Alam and one in each state. Whilst the main university is visible digitally, the branch campuses are almost invisible digitally. These may lead to inauthentic content and dispute of affiliation among the UiTM branches. This paper discusses the approach taken to overcome the problem of the authenticity of facts about UiTM. Content analysis is deployed throughout all related campus pages found on Wikipedia. A table of guideline was developed to form an authenticity checklist.

Keywords*: Content Authenticity, Digital Visibility, Search Engines, Web Crawler, Wikipedia.*

1.   Introduction

Wikipedia is one of the most popular websites in the world. It was launched on January 15, 2001 by Jimmy Wales and Larry Sanger (Augustyn *et al*, 2020) initially as an English-language encyclopaedia and was then expanded to various versions of Wikipedia in other languages. Its impact has been both negative as well as positive in the digital environment because of corporate misuse of personal data and its potential for spreading disinformation. However, in improving the visibility of a university around the world, the Wikipedia platform is crucial as it has the potential to channel information about a university to the world.

A study by Amina *et al*. (2020), of 225 students from the students of master programmes at the University of Punjab found that the students have a positive behaviour towards Wikipedia and regarded it as a good source of information. They use Wikipedia for both academic and leisure purposes. Because of its frequent use and conscious attempts by its authors, indirectly the Wikipedia has begun to be accepted by the community, especially among university communities.

This is supported by a study by Meseguer-Artola *et al*. (2020) which cites the Wikipedia as an open educational resource where the frequency of use and importance in

higher education is growing. However, empirical evidence about Wikipedia's contribution to students' academic performance is scant and many higher education's actors express concern regarding its value.

Regarding the authenticity of information, the Wikipedia has taken several approaches. According to Wyatt (2020), the policies and practices enacted by Wikipedia to achieve their standards are rarely unique because they cannot achieve a uniformly high standard. However, discussion pages, several associated paratexts and the statistics demonstrating article popularity are considered.

Wyatt (2020) also investigated Wikipedia articles on historical people for studying link-related temporal features of articles on past people and his findings have addressed issues regarding the characteristics of information about historical people recorded in the English Wikipedia. It seems that the use of signals derived from the hyperlink structure of Wikipedia as well as from article view logs which overlay over temporal dimension are able to understand the relationship between time periods, link structure and article popularity.

According to Bonk *et al.* (2009), the content of web pages can be edited by public users, hence the issue of content quality can be questionable. However, a study by Norlidah *et al.*, (2013) found that there was a significant increase in the contribution of articles on the Wikipedia platform and it was highly relevant in collaborative learning. In terms of content quality assurance that has been argued by Ruprechter *et al.* (2020), long-term improvement of content quality by offering valuable insights about the interplay of article quality, controversies and edit wars, editing behaviour, and network properties via sequence- based edit and network-based article metrics will assist online collaborators to ensure quality and authenticity of information.

Universiti Teknologi MARA (UiTM) is a public university located primarily in Shah Alam, Malaysia. UiTM is Malaysia's largest institution of higher learning in terms of number of students and the largest university system in the country. Besides the main campus in Shah Alam, the university has expanded nationwide with 34 branch campuses, which collectively offers more than 500 academic programmes. Wikipedia plays an important role in enhancing the digital visibility of any organization (Ebrahim, 2016), especially one as large as UiTM. This is because google search has used Wikipedia pages as a reliable source of text. Based on the Wikipedia page history, the content about Universiti Teknologi MARA (UiTM) in Wikipedia was started in 2004 and its content is still updated from various sources. The updating method is done automatically and manually. The automation method is the updating of Wikipedia content using text analysis of authentic article sources available on the Internet while the manual method uses human effort as the author.

The involvement of authors from UiTM's organization itself has contributed a lot to the authenticity of information about UiTM including about the latest governance organization, achievements, and successes. This is made more significant to ensure that the distribution of information and references is relevant in the digital space. The purpose of this paper is to recommend the actions taken by UiTM in ensuring that the organization is digitally renowned with authentic information using Wikipedia.

2. Related Work

The relationship between Wikipedia and Google is very significant where it has greatly improved the information retrieval process on the internet. An experiment done by McMahon *et al.* (2017) has shown that the Google search engine did not perform well for many queries when it was not mutually related with Wikipedia. This shows that the text analysis in

Wikipedia system can provide the best document searching impact for queries. This is called an extensive interdependency between Google and Wikipedia.

Mahji *et al*. (2016) mentioned that content analysis is a set of procedures for collecting and organizing information in a standardized format that allows analysts to make inferences about the characteristics and meaning of written and other recorded material. It is a powerful tool for examining trends, patterns in documents, and authorship pattern. The values and intentions of the authors can be inferred from the data which may reveal underlying themes and associations.

According to Lewoniewski (2019), Wikipedia contains over 50 million articles in over 300 different languages and English language articles is the largest which contains over 5.8 million articles. Wikipedia can be trusted as much as people telling any information to everyone, however it is not a source of information at all. Even though the Wikipedia is an excellent case study on research in the digital age, Hoeven *et al*. (2019) said that digital visibility of anything is the ability to make people's behaviours, decisions, and preferences more visible. For scholars who study organizations and their effects upon society, increased information visibility raises the hope that organizations might become more transparent. However, there is a lack of empirical evidence to support this assumption.

UiTM is one of the organizations that look forward to being visible digitally. It is an effective and fastest method to be known by the world community about the organization. According to Mohamed Hashim et al. (2021), digital transformation in the global higher education will determine the future roadmap in sustainable education management strategy. Indirectly, it could increase its popularity. Therefore, the information about UiTM on the Wikipedia platform need to be monitored from time to time so that they are authentic and updated. Without consistent monitoring, the tendency of misinformation may happen, and it will affect the perception about the organization to the society.

3.   Methodology

The paper discusses the approach taken to standardize the topic and the authenticity of facts about UiTM. Content analysis is deployed throughout all related campus pages found on Wikipedia. Based on UiTM's governance, the university comprises one main campus and 34 state campuses where each branch campus has a standard official organization name (Wikipedia, 2022). This study uses a qualitative method in which all Wikipedia page related to UiTM and an analysis was conducted manually. Each name of the branch campus will be checked against the list of topics on Wikipedia. If there are differences between the topic keywords and the branches' name, it means that the authenticity of Wikipedia content about the topic is in doubt.

A team has been formed to make corrections and confirmations to the contents that are available on the existing Wikipedia page. Each branch / campus has an appointed officer to monitor the content of the page from time to time to ensure the authenticity of the information on Wikipedia. As the main organization, the nomenclature or official name of UiTM in Wikipedia is Universiti Teknologi MARA. UiTM has 13 branches with each campus having its own governance. This study was conducted using the qualitative method that thoroughly studied all Wikipedia pages which are related to UiTM and its contents were analyzed by domain experts.

A total of 35 topics or queries were searched using the Google search engine and search function in Wikipedia. The topic name used for the searching is the affiliation for each campus. From the findings, there are several related pages has been retrieved from Wikipedia pages. The management teams which worked as the domain experts will validate retrieved pages whether it found doubtful or not. For example, for the query titled as "Universiti

Teknologi MARA (UiTM) Cawangan Sarawak", the related pages have been retrieved as shown in the Figure 1 below with the topic name as "Kampus UiTM Samarahan".
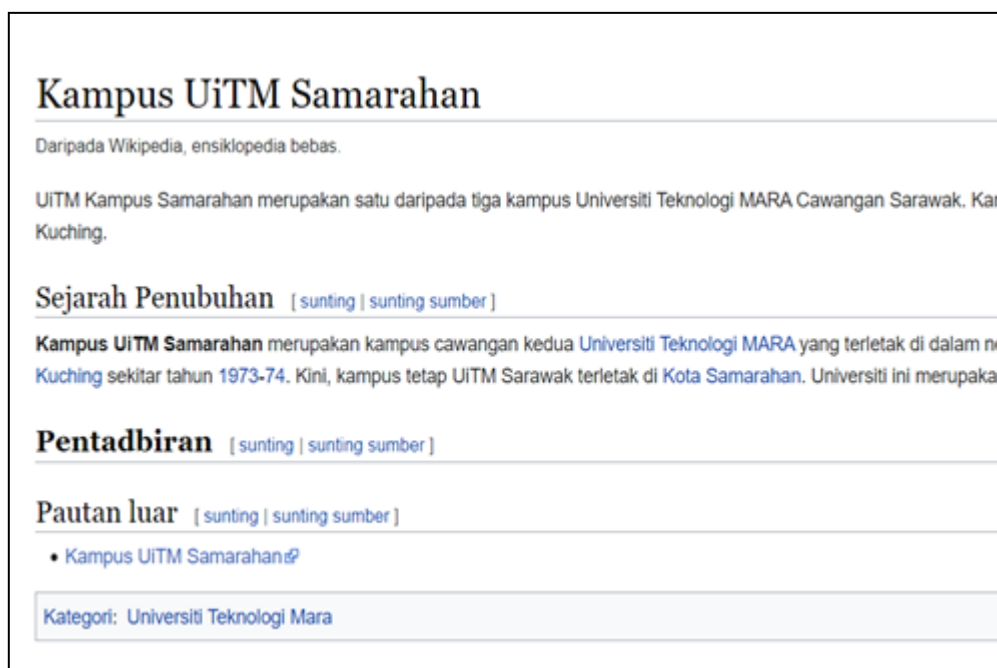

Figure 1. A snippet from the retrieved Wikipedia page.

4.   Analysis

The process of organizing content is the key to identifying the validity and the authenticity of the information about the topic. The results found some of the retrieved pages obtained were developed by the bot function found in Wikipedia. All the queries used the university and affiliation name and the findings have been verified based on the topic's name and the authenticity of its content. Table 1 shows the findings from the search results on the list of topics/keywords used in the search that has been studied since January 2019.

The authenticity of the content has been evaluated by the domain expert from UiTM. The higher percentage indicates the authenticity of the content in Wikipedia pages. From the observation of the Wikipedia's activity log, the contents reviewed made by bot function were based on the authentic website resources. However, there are some contents that were mismatched, and this may lead to misperception and misunderstanding about the topic (for example organization) generally. Therefore, thorough checking is a necessary process by the owner of the topics.

Table 1. Findings from Search Results on the List of Topics/Keywords.

Organisation Name	Retrieved Pages	The authenticity of the content
Universiti Teknologi MARA	Universiti Teknologi MARA	>80%
UiTM Perlis Branch/ UiTM Cawangan Perlis	UiTM Perlis Kampus UiTM Arau	>50% >20%
UiTM Pulau Pinang Branch/ UiTM Cawangan Pulau Pinang	Kampus UiTM Permatang Pauh Kampus UiTM Bertam	>20% Not relevant
UiTM Kedah Branch/ UiTM Cawangan Kedah	Kampus UiTM Kedah	>30%
UiTM Perak Branch/ UiTM Cawangan Perak	Kampus UiTM Sri Iskandar	>20%
UiTM Selangor Branch/ UiTM Cawangan Selangor	Kampus UiTM Puncak Alam Kampus UiTM Puncak Perdana Kampus UiTM Jalan Othman	>20% >20% Not relevant
UiTM Melaka Branch/ UiTM Cawangan Melaka	Kampus UiTM Bandaraya Melaka Kampus UiTM Alor Gajah Kampus UiTM Jasin	>20% Not relevant Not relevant
UiTM Negeri Sembilan Branch/ UiTM Cawangan Negeri Sembilan	Kampus UiTM Seremban 3 Kampus UiTM Kuala Pilah	>20% Not relevant
UiTM Johor Branch/ UiTM Cawangan Johor	Kampus UiTM Segamat Kampus Johor Bahru, UiTM	>20% Not relevant
UiTM Pahang Branch/ UiTM Cawangan Pahang	Kampus UiTM Jengka Kampus UiTM Raub Kampus UiTM Bukit Sekilau	>20% Not relevant Not relevant
UiTM Terengganu Branch/ UiTM Cawangan Terengganu	Kampus UiTM Dungun Kampus UiTM Bukit Besi	>10% Not relevant
UiTM Kelantan Branch/ UiTM Cawangan Kelantan	Kampus UiTM Machang Kampus UiTM Bandar Kota Bharu	Not relevant Not relevant
UiTM Sabah Branch/ UiTM Cawangan Sabah	UiTM Kampus Sabah Kampus UiTM Kota Kinabalu Kampus UiTM Tawau	>20% Not relevant Not relevant
UiTM Sarawak Branch/ UiTM Cawangan Sarawak	Kampus UiTM Samarahan Kampus UiTM Mukah	Not relevant Not relevant

5.   Result and Discussion

A task force framework approach is a management method used to determine the validity of information and identified topics. The first approach is by forming up a team from each branch that represents each topic of their respective organization name. They are responsible for checking the availability of corresponding content which are available from time to time.

Moreover, they must create a new topic with a right affiliation for the organization in the Wikipedia page if the existing topic is invalid.

The second approach is to produce a reminder or a periodic table for monitoring purposes. Since one of the policies of Wikipedia is open basis where anyone can create articles, and any Wikipedia page could be edited by any reviewer, even those who did not have a Wikipedia account, the team need to put effort to monitor the content twice a year. According to Guo and Stevens (2011), all the document collection are easily can be editable, interlinked through hypertext systems for storing and modifying information. This method is to ensure that the content is always relevant in the digital world.

Another approach is by using a hyperlink. Since the Wikipedia page of "Universiti Teknologi MARA" is the main page of all other branches, all the related hyperlink to the branches is stated in the page. Figure 2 below shows the interface page of the main page. The reference of the table is the related hyperlink. There are two types of hyperlinks used, namely as hyperlink to the official website or hyperlinks to Wikipedia pages on topics that have been maintained.

List of UiTM campuses around Malaysia

Campus	Location	Establish	Ref
Main Campus			
Shah Alam	Shah Alam	1967	[1]
UiTM Selangor Branch			
Puncak Alam	Bandar Puncak Alam	2009	[2]
Puncak Perdana	Shah Alam	2005	[3]
Jalan Othman	Petaling Jaya	1956	
Dengkil	Dengkil	2016	[4]
Selayang	Batu Caves	2010	[5]
Sungai Buloh	Sungai Buloh	2012	[6]
UiTM Perak Branch			
Seri Iskandar	Seri Iskandar	1985	[7]
Tapah	Tapah	2010	[8]
Bandaraya Ipoh	Ipoh		
UiTM Melaka Branch			
Alor Gajah	Alor Gajah	1984	[9]
Bandaraya Melaka	Melaka Bandaraya Bersejarah	2007	[10]
Jasin	Merlimau	2011	[11]
UiTM Negeri Sembilan Branch			
Kuala Pilah	Kuala Pilah	1999	[12]
Seremban	Seremban	2014	[13]
Rembau	Rembau	2016	[14]

UiTM Johor Branch			
Segamat	Segamat	1985	[15]
Pasir Gudang	Masai	2013	[16]
UiTM Pahang Branch			
Jengka	Bandar Tun Razak	1993	[17]
Raub	Raub	2015	[18]
Kuantan	Kuantan		
UiTM Terengganu Branch			
Dungun	Dungun	1975	[19]
Kuala Terengganu	Kuala Terengganu	2008	[20]
Bukit Besi	Bukit Besi	2013	[21]
UiTM Kelantan Branch			
Machang	Machang	1985	[22]
Kota Bharu	Kota Bharu	2007	[23]
UiTM Pulau Pinang Branch			
Permatang Pauh	Permatang Pauh	1996	[24]
Bertam	Kepala Batas	2011	[25]
UiTM Perlis Branch			
Arau	Arau	1974	[26]
UiTM Kedah Branch			
Sungai Petani	Merbok	1997	[27]
UiTM Sarawak Branch			
Samarahan	Kota Samarahan	1973	[28]
Samarahan 2	Kota Samarahan	2014	[29]
Mukah	Mukah	2007	[30]

Figure 2. A part of Wikipedia Interface from "Universiti Teknologi MARA" page.

In order to ensure the digital visibility of UiTM through using Wikipedia, the comprehensive content about UiTM in the Wikipedia will bring the visibility impact to UiTM digitally. From the analysis and thorough observation, a standard content arrangement has been proposed. Each representative from the UiTM branches is required to follow the guidelines below which cover the encyclopaedia type of article on a topic. The purpose of the guidelines is to ensure that information is easy to monitor and update.

Appendix A is the findings that have been obtained on 1 June 2020 based on below criteria.

Topic Name:
- Universiti Teknologi MARA (UiTM) Cawangan (……..) or
- Universiti Teknologi MARA (UiTM) Cawangan (……..) Kampus (……)

Body of Content:
- Introduction
- History of establishment
- List of campuses
- Academic programs
- Research
- Reputation / Achievement / Internationalization
- Alumni

6. Conclusion

The main challenge is to address the Wikipedia itself which allows anyone to author all the articles in Wikipedia. Thus, the role of the representatives from related organization is important to ensure the authenticity of content about Universiti Teknologi MARA and its branches remain relevant to ensure the visibility of UiTM in cyberspace.

The use of the Google search engine is very significant in searching for any articles about UiTM on the Internet. The collaboration between the search engine and Wikipedia is a good combination of making a search to be more meaningful. Anyone can gain information by searching via Google, apart from browsing the official website of the organization itself to obtain information about UiTM. Hence, the authenticity, veracity, and relevance of information regarding UiTM and all its branch campuses is of ultimate significance both to the institution and its users.

Acknowledgement

References

Amina, W., Warraich, N.F. & Malik, A. (2020, July). Usage of and learning from Wikipedia: A Study of University Students in Pakistan. *Global Knowledge, Memory and Communication*, 70(3), 282-292.

Augustyn, A., Zeidan, A., Zelazko, A., Eldridge, A., McKenna, A., Tikkanen, A., Schreiber, B. A., Duignan, B., Mahajan, D., Promeet, D., Rodriguez, E., Gregersen, E., Shukla, G., Liesangthem, G., Lotha, G., Young, G., Bolzon, H., Luebering, J.E., Wallenfeldt, J., Hibler, J., Rafferty, J.P., Gupta, K., Rogers, K., Chmielewski, K., Manchanda, K., Chaveriat, L., Dixon, L., Matthias, M., Petruzzello, M., Ray, M., Metych, M., Bauer, P., Riley, P., Bhathya, P., Hollar, S., Singh, S., Gupta, S., Seddon, S. & Setia, V. (2020, October). *The Information Architects of Encyclopaedia Britannica available at https://www.britannica.com/topic/Wikipedia.*

Bonk, C. J., Lee, M. M., Kim, N., & Lin, M.-F. G. (2009, December). The tensions of transformation in three cross-institutional wikibook projects. *The Internet and Higher Education,* 12, 126–135. DOI: 10.1016/j.iheduc.2009.04.002.

Ebrahim, N. A., (2016, August). Contribute to Wikipedia: An approach to Increase Research Visibility on the Web. In *3rd Series of Introductory Workshop On: Strategies to Enhance Research Visibility, Impact & Citations.* DOI: 10.6084/m9.figshare.3750252.v1.

Guo, Z., & Stevens, K. J. (2011). Factors Influencing Perceived Usefulness of Wikis for Group Collaborative Learning by First Year Students. *Australasian Journal of Educational Technology*, 27(2), 221-242.

Hoeven, C.L., Stohl, C., Leonardi, P. & Stohl, M. (2021, August). Assessing Organizational Information Visibility: Development and Validation of the Information Visibility Scale. *Communication Research*, 48(6), 895-927, DOI: 10.1177/0093650219877093.

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

Mahji, S., Jal, C & Maharana, B. (2016, February). *Content Analysis of Journal Articles on Wiki in Science Direct Database*. Library Philosophy and Practice (e-Journal), ISSN: 1522-0222.

McMahon, C., Johnson, I. & Hecht, B., 2017, The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship Between Peer Production Communities and Information Technologies. In *Eleventh International AAAI Conference on Web and Social Media*. (pp.142-151), AAAI Press, Canada.

Meseguer-Artola, A., Rodríguez-Ardura, I., Ammetller, G. & Gilabert, E.R. (2020, Jun). Academic Impact and Perceived Value of Wikipedia as a Primary Learning Resource in Higher Education. *Profesional de la Informacion*, 29 (3), 1-16.

Mohamed Hashim, M. A., Tlemsani, I. & Matthews, R. (2021, September), Higher Education Strategy in Digital Transformation. *Education and Information Technologies,* DOI: 10.1007/s10639-021-10739-1.

Norlidah, A., DeWitt, D. & Siraj, S. (2013, November). A Content Analysis of Wikis in Selected Journals from 2007 to 2012. *Procedia – Social and Behavioral Sciences*, 28 – 36. DOI: 10.1016/j.sbspro.2013.10.303.

Ruprechter, T., Santos, T. & Helic, D. (2020, September). Relating Wikipedia Article Quality to Edit Behavior and Link Structure. *Applied Network Science,* 5(61), DOI: 10.1007/s41109-020-00305-y.

Wikipedia. (2022). Campuses and the University System. [online] Available at: https://en.wikipedia.org/wiki/Universiti_Teknologi_MARA. Accessed 29 Mac 2022.

Wyatt, L. (2020, April). Endless Palimpsest: Wikipedia and the Future's Historian. *Studies in Higher Education*. 45 (5), 963-971.

Appendix A

The content updating progress checklist based on Wikipedia's topic about UiTM Branches on 1 June 2020

No	States in Malaysia	The latest title/ topic in Wikipedia	Language used	Introduction	History of establishment	List of campuses	Academic programs	Research	Reputation / Achievement / Internationalization	Alumni	Topic and Content: Relevancy
1	Kedah	Kampus UiTM Kedah	Malay	√	√	X	√	X	X	X	Irrelevant
2	Perlis	Universiti Teknologi MARA, Perlis	English	√	√	Irrelevant	√	X	X	X	Irrelevant
3	Pulau Pinang	No changes	Malay	X	X	X	X	X	X	X	Irrelevant
4	Perak	UiTM Cawangan Perak	Malay	X	X	√	X	X	X	X	Irrelevant
5	Selangor	No changes	Malay	X	X	X	X	X	X	X	Irrelevant
6	Melaka	Kampus UiTM Bandaraya Melaka Kampus UiTM Alor Gajah	Malay	√	√	X	√	X	X	X	Irrelevant
7	Negeri Sembilan	No changes	Malay	X	X	X	X	X	X	X	Irrelevant
8	Johor	Kampus Johor Bahru, UiTM	Malay	X	X	X	X	X	X	X	Irrelevant
9	Pahang	Universiti Teknologi MARA Cawangan Pahang	Malay	√	√	X	√	X	√	√	Irrelevant
10	Kelantan	Kampus UiTM Machang	Malay	√	√	X	√	X	X	X	Irrelevant
11	Terengganu	Kampus UiTM Dungun	Malay	√	√	X	√	X	X	X	Irrelevant
12	Sabah	Kampus UiTM Kota Kinabalu	Malay	√	√	X	√	X	X	X	Irrelevant
13	Sarawak	No changes	Malay	X	X	X	X	X	X	X	Irrelevant

(Note: √ means completely done)