

## IDENTIFYING THE ETHICAL ISSUES IN TWITTER: A KNOWLEDGE ACQUISITION FOR ONTOLOGY

*Mohamad Hafizuddin Mohamed Najid<sup>1\*</sup>, Zahidah Zulkifli<sup>2</sup>, Roslina Othman<sup>3</sup>, Rohaiza Rokis<sup>4</sup>, Ashraf Ali Salahuddin<sup>5</sup>*

<sup>1,3</sup>Department of Library and Information Science, Kulliyah of Information and Communication Technology, International Islamic University Malaysia, Malaysia

<sup>2</sup>Department of Information System, Kulliyah of Information and Communication Technology, International Islamic University Malaysia, Malaysia

<sup>4</sup>Department of Sociology and Anthropology, Kulliyah of Islamic Revealed Knowledge and Human Sciences, International Islamic University Malaysia, Malaysia

<sup>5</sup>Humanitarian Crisis Management Programme, BRAC, Marine Drive Road, Cox's Bazar, Bangladesh

Email: hafizuddinajid@gmail.com<sup>1\*</sup>(corresponding author), zahidahz@iium.edu.my<sup>2</sup>, roslina@iium.edu.my<sup>3</sup>, rohaiza@iium.edu.my<sup>4</sup>, muhammad.ashraf@brac.net<sup>5</sup>

DOI: <https://doi.org/10.22452/mjcs.sp2021no2.7>

### ABSTRACT

*Social media is an open platform to communicate, share and exchange information freely. This uncontrolled exchanged information carries out both negative and positive impacts in others' lives. In this regard, this study aims to identify ethical issues on this information in line with Ibn Khaldun's ethical considerations. Out of many other social networking sites, Twitter has been identified as one of the most popular microblogging social networking platforms. Using a simple algorithm in R programming and 43 keywords based on Ibn Khaldun's thoughts, 1075 public tweets have been extracted from Twitter as a sample of ethical issues. The sentiment analysis in Parallel Dots was performed on the collected tweets, and it was discovered that 700 of the tweets are positive statements, 229 are neutral statements, and 146 are negative statements. Having done the validation process on these sentiments, the study proposed these identified ethical issues from tweets as a domain in developing ontology relationships with Ibn Khaldun's thoughts. In this process, further study can be carried on with wider data from various sources beyond the limitation of this study. Thus, a semantic database could serve as a guideline for SNS ethical issues based on Ibn Khaldun's thoughts.*

**Keywords:** *Ethics, Ethical Issues, Ibn Khaldun's Thoughts, Parallel Dots, R-Programming, Sentiment Analysis, Social Networking Sites, Twitter*

### 1.0 INTRODUCTION

Social media has evolved not only as a personal communication medium but also, as a medium for people to express their opinions on goods and services as well as political and general events. Because of its broad and widespread popularity, a massive amount of user reviews or opinions are produced and exchanged daily [1]. With over 4.2 billion active social media users connecting to the internet in January 2021 [2], the volume of data and metadata is incredible. By now, technology development has intensified and a world without the internet is unimaginable. The internet, which connects billions of people worldwide, is a core pillar of the global information society. This massive amount of data is created by many sources on the internet, with social media users leading the way in terms of the amounts of data being generated. Social media has changed how people share their opinions on products and services, as well as their sentiments and views and the platforms on which they do so. Now that social networking sites (SNS), blogs, online forums have come into play and thus they generate massive amounts of data that when analyzed can be useful for business, technologies, global issues and so on [3].

Humans are generally social and the SNS platforms offer diverse avenues for people to remain in touch all over the world. Social media research is considered to include investigations into user behavior and opinions or utilization at all the SNS stages that display user contributions or user activities. As user comments have become an essential part of many online environments, the boundary between the web and the SNS is disappearing [3]. Mining users' opinions from social media can be refined in several ways. Analyzing unstructured data is a difficult

task in and of itself, and extracting valuable information from it is a big challenge. In this article, an open-source approach is presented in which the tweets have been collected through Twitter API (Application Programming Interface) and then pre-processed by utilizing a statistical tool, R programming. These powerful tools and software will assist in handling and extraction of millions of tweets as part of a knowledge acquisition process. Several measures should be taken. R is an open-source method for analyzing online reviews and doing text mining [1]. All scripting languages in R have the frameworks necessary for reproducible analysis, and open-source software allows the interested researcher to execute the scripts needed to reproduce findings without incurring any costs, provided adequate hardware is available. Twitter data are obtained using the R package. Using keyword search queries, the package allows authentication and access to Twitter posts. R makes it possible to collect targeted portions of Twitter data sample for social media mining. R can read a wide range of advanced formats in addition to standard data types and files from traditional statistical software packages. R for example will read relational databases, Hadoop, and other web formats like Twitter [4].

The research objective is to identify the ethical issues in SNS and relate them with Ibn Khaldun's thoughts. In the knowledge acquisition process, tweets were extracted using R, and sentiment analysis was performed to seek ethical issues, which resulted in negative sentiments. This approach is to address the research questions of what the ethical issues from tweets on Twitter that can be examined. Later, in the following step, the tweets will be linked to Ibn Khaldun's thoughts as a solution to the identified ethical issues. Specifically, the paper explores the role of Twitter based on textual communications using social media analytics in identifying the ethical issues, namely, sentiment analysis, based on tweets written in Bahasa Malaysia (i.e., the national language, herein referred to as the Malay language). The rest of the paper is structured as follows: Section 2.0 provides the literature review of the study focusing on SNS, Ibn Khaldun's thoughts, R-Programming and sentiment analysis. This is followed by the research methodology in Section 3.0. The results are presented and discussed in the subsequent section, before concluding the paper in Section 5.0.

## **2.0 LITERATURE REVIEW**

### **2.1 Social Networking Sites and Ethics**

SNS are the communication platforms where people electronically connect with others and matters. Various platforms on the internet get people socially engaged with their surroundings. Their main role is to facilitate relational connections such as LinkedIn advocates social relation with professional lives, MySpace started to promote musicians and connect their fans, Twitter is used as a simple way of communication between individuals and public figures, etc. [5], [6].

However, day by day, the focus and objectives of SNS are being changed or evolved with other concerns. Some are featuring e-commerce, some creating entrepreneurs, and some of them are helping business commercials to profile links with their existing and future customers. And Facebook is a good example in this regard [7]. Likewise, when it comes to the concern about business, following business ethics becomes mandatory for the betterment of customers. So, as the business merges with technology in SNS, users' privacy and transparency became the most highlighted concern in research such as the study of [3], [5], [6].

In addition to that, [6] has come up with few other contemporary ethical concerns in SNS. In his view, civil discourse, freedom and democracy in the public arena must be brought into the ethical discussion as privacy, identity, community and friendship on SNS involve these cases. Apart from that, he also brings an interesting philosophical debate on whether or not the SNS helps us to lead to a good life by growing a broader intellectual virtue or hinders our development. According to the analysis of Edward Spencer as cited in [6], the contemporary ethical concern is confined to interpersonal issues (violation of privacy and copyright issues, etc.). If SNS needs to be used for a more harmonious and good life, the ethical values should expand their focus beyond its present concerns to a more universal ethical question of prudential wisdom. In his word, "this concern about prudential wisdom and the good life is part of a growing philosophical interest in using the resources of classical virtue ethics to evaluate the impact of SNS and related technologies".

### **2.2 Islamic Ethics and Ibn Khaldun's Thoughts**

In line with that, Islamic ethics have a slight similarity in this context. Islam clarifies that ethics should be considered as manners towards a person and manners with God. Ethics, in other words, knowledge of morals or moral values that become a guide for individuals or communities. In such a discussion, [8] stated that current technology is synonymous with our daily life. As a result, Islamic ethics can be disseminated through SNS.

On this account, Ibn Khaldun, a prominent scholar in sociology, has explicitly discussed how Muslim ummah should lead their lives in this world. In his book *Muqaddimah*, he emphasizes that knowledge, values and ethics should be followed in every aspect of life. Following that theory of Ibn Khaldun, Abdullah [9] studied the youth human capital development in accordance with knowledge and ethics. She said that if the young generations neglect ethics and values, many chaotic problems will be swept over their minds. Thus, the civilization will face moral decay, ethical problems and many kinds of degradation of character. Therefore, considering Ibn Khaldun's thoughts, the author proposes that young generations be educated about society's negativity and positivity so that they might obtain what is good and avoid what is harmful.

This is how this study has been influenced as well. The SNS is full of information with good and bad. That can influence the generation either way. In this case, the tweets are retrieved to be analyzed and connected in accordance with Ibn Khaldun's thoughts. In this article, The *Muqaddimah: An introduction to history Ibn Khaldun, 1332-1406* [17] book has been used in the understanding of Ibn Khaldun's thoughts. This book was translated into English by Franz Rosenthal, who is well-known for his huge contribution to the translation version and has been cited by many scholars in their research papers.

### 2.3 Twitter

Thousands of users exchange data in every moment on various SNS such as Facebook, Instagram, and Twitter are providing the opportunity to benefit suppliers. Twitter claims to have the application of a social network that centers upon capable systems and micro-blogging. It offers an easy and friendly interface to chat and stay connected with other users. [10].

Likewise, the posts on Twitter can be found using a particular keyword or topic. It follows traditions that recognize it from other sorts of literary information. Twitter names, hashtags, and notices give a clear way to comprehend individuals, subjects and sort out substances. The tweets in its system are related to unmistakable tweet qualities that can be retrieved through its API or other tools. To gather specific information, advance pre-processing can be done by utilizing @twitter\_names and #hashtags. This method is clearer and accumulates better results narrowing down the contents than any other method [10].

For example, hashtags assist users to get information according to the culture or trends. Users make hashtags using its symbol (#) before a word to find out a certain topic or to spread out that specific topic. Then the system tries to retrieve available posts related to those hashtags [11]. Thus, Twitter also allows users to find any specific tweets in real-time so that the users can focus on the desired topic they wanted to get using hashtags. In addition to that, a new feature is also there to personalize the search results, so the retrieved tweets do not match with other users' tweets.

Moreover, its system uses inverted indexes to categorize the search results in rank. This helps users to identify the required tweets from the ranking list immediately since the most matched one with the query is the first one in the list. These data structures are incredibly beneficial [9]. At the same time, Twitter is distinctive from other SNS in some significant ways. Its capacity to cultivate second-order associations is one point of distinction. These are noteworthy since they bring data to people from other users with whom they share less.

Like other SNS, Twitter also encourages self-organization. Its users can effectively utilize the platform to accumulate knowledge, make comments, and hold up open complaints. If the users find the tweets worthwhile, they can demonstrate their level of acceptance that it is accurate or correct [4]. Besides, Twitter is one of the user-friendly and convenient SNS that allows easy access to data for social media research. As a result, the larger part of research on SNS is found on Twitter data despite being the number of Twitter users is less than other social media users. So, this paper has focused on tweets from Twitter to do the identification of SNS ethical issues based on Ibn Khaldun's thoughts so that a subsequent study may come up with domain ontology based on the results.

### 2.4 R-Programming

The R-programming language is used for information control, recreation, calculation and reality shows. This language can analyze data very efficiently and it is designed with array and matrix processing operators. It can present data in illustrations and data demonstration. On the other hand, R is comparable to the S programming language developed by John Chambers and colleagues at Bell Research facilities. So, the R language is open-source software and adapted from the S programming language. The commercial version of R is based on S plus.

However, R has many advantages such as multi-platform compatibility so that it can run on different operating systems. Project R was started in 1993 for statistical computing, bioinformatics and graphics [12]. Since then, its demand has persistently increased over the years. R has many more features that can be utilized as required [13]. In 2020, this language became the sixth most popular programming language in IEEE Spectrum ranking. Hence, the method of this study has considered R as a tool to extract tweets from Twitter in order to identify the ethical issues based on Ibn Khaldun's thoughts.

## 2.5 Sentiment Analysis and ParallelDots

Over the years, social media became one of the most accessible media to communicate. It is not only used as a personal communication medium, but it also became the source of communication about products, reviews, political views, social problems, and general events among people [1]. In this regard, service providers take this opportunity to use these shared data over SNS for various purposes; social worker needs these data to identify social issues, commercial vendors need to find their targeted customers analyzing the needs of users in their shared texts and companies need to get direct feedback of the consumers mining their opinions and reviews about their products.

However, due to this huge amount of data, it is not easy to mine their texts to generate sentiment analysis. There are tools to do this task using artificial intelligence (AI). For example, [2] has analyzed the views of people in eight different categories of emotion using R-programming (anger, trust, fear, anticipation, disgust, sadness, joy, surprise) and two different sentiments (namely positive and negative) from the emotion lexicon EmoLex. They have done this process to identify the people's reaction to the Barcelona terror attack. Throughout this sentiment analysis on Twitter, they tried to figure out whether people showed support for the victims in these inhuman acts or showed aggression against the terrorists.

ParallelDots is one of the AI-based sentiment analysis platforms in natural language processing (NLP). Unlike the R language, ParallelDots does not need any programming or algorithm to run the analysis. Its NLP is trained on a billion documents and proved to be accurate in sentiment analysis and emotion detection. ParallelDots interface provides 10 different features for NLP such as semantic analysis, sarcasm analysis, text classification, named entity recognition, intent analysis, etc. Besides, the API of this platform is available in fifteen different languages. However, the whole platform is not free, only limited characters are free to analyze [14]. Despite this challenge, ParallelDots has been used in this study to sort out the tweets according to the sentiments: positive, negative and neutral.

## 3.0 METHODOLOGY

Data Extraction is the process of gathering and measuring information on targeted variables. Accurate data collection is essential in maintaining the integrity of research. This section elaborated the knowledge acquisition process in discovering the potential sources in retrieving the tweets on Twitter. Search API is the official authentication API for Twitter, which returns the tweets that match the given string and writes them into the object. The extraction of tweets data from the tweets on any topic using the developing hashtag “#” keyword. In this paper, the keyword was generated from Ibn Khaldun's thoughts from his prominent book name *al-Muqaddimah*. This book has been used in getting an understanding of Ibn Khaldun's thoughts and in identifying the reference of description for each Ibn Khaldun's thoughts as listed.

The document analysis process involves skimming, reading, and interpretation in refining ideas, identifying concepts and pinpointing the significance of themes (Fig. 1). The *Muqaddimah* was analyzed together with data from other sources to distinguish the themes. Several previous research papers have also been reviewed in giving supporting ideas and giving an understanding of what has been discussed in that book within that topic area. The thoughts of Ibn Khaldun's have also been analyzed from work by Abdullah [9] and Salleh [15] as the supporting data. These Ibn Khaldun's thoughts have been identified and reviewed related to the tweets retrieved from Twitter and compared with what has been found in the *Muqaddimah* by arranging them in the same themes. The method of evaluating documents in such a way to produce experimental knowledge and developed an understanding of Ibn Khaldun's discourse. The initial coding of the content of the documents was related to humans and civilization based on the aims of the research.



Fig. 1: Identification of Ibn Khaldun's Thoughts

The data was checked and rechecked the codes and concepts to organize ideas and the theme concepts that appeared to cluster together. To be sure the diverse sources of data gave a complete picture of the representing Ibn Khaldun's thoughts, the triangulation of data sources countered threats to trustworthiness and researcher bias. The researcher needs to determine not only the existence and accessibility but also the authenticity and usefulness of the proposed Ibn Khaldun's thoughts by presenting and verifying them with experts. There are 10 thoughts of Ibn Khaldun which were justified from his discourse related to humans and civilization for this research. Table 1 provides Ibn Khaldun's thoughts as contained in the *Muqaddimah*. Also, from this process, the thoughts of each Ibn Khaldun's were then extended in searching for their relevant subtopic to support the main thoughts.

Table 1: Ibn Khaldun's Thoughts and Subtopic

Ibn Khaldun's Thoughts	Subtopic
1. Worship God	Creed, Happiness
2. Social Integration or Solidarity	Bad or Negative Acquaintances, Thinking Ability, Weak Law Enforcement
3. Good or Childhood Education	Spiritual or Emotional Attachment, Development of Intellect, Value of Experience, Grounded Religious Knowledge
4. Teaching al-Quran	Early Age Learning and Memorization Skills, The Truth Seeking, Tawhidic Application
5. Good Ethics	Righteous Manners or <i>Adab</i> , Ethical Issues and Problems, <i>Ta'dib</i> in Education
6. Great Scholars	Piety and Sincerity, Pursuant of Scientific Knowledge, Holistic Governance
7. Awareness Society	Divinity, Kinship, Quality of Human Resources, Group Solidarity
8. Good Governance or <i>Khalifah</i>	<i>Khalifah</i> Governance System, Cooperation, Leadership, Propaganda
9. Economy	Fair Businesses, Monetary Policy, Wealth Protection, Wealth Distribution
10. Human Civilization	Establishment and Spreading of Islam, Civilized Nation, Prosperity, Prosperous Society

Then, from 10 Ibn Khaldun's thoughts justified from his discourse related to human and civilization, 43 synonyms have been generated and verified by searching through the Institute of Language and Literature (DBP) thesaurus to be generated as a keyword for searching as a hashtag in Twitter (Fig. 2).

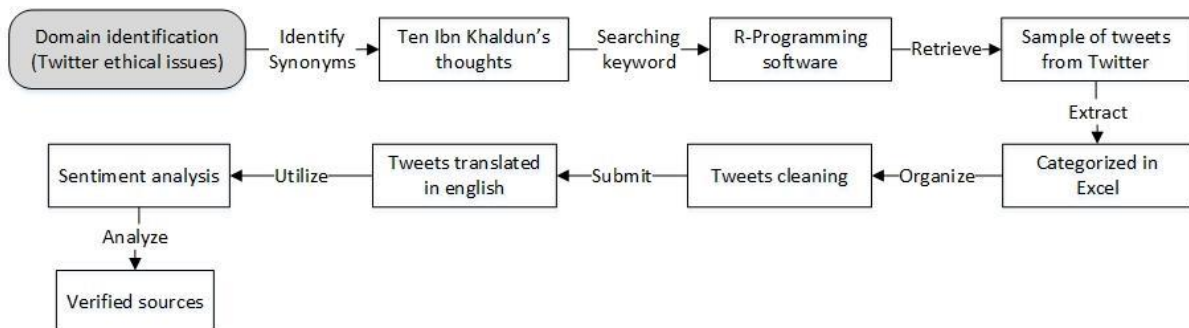


Fig. 2: Identification of Twitter Ethical Issues

Each of the 10 thoughts of Ibn Khaldun's was searched through DBP's thesaurus in identifying a keyword to be searched as hashtags on Twitter. Then, from the search results of each thought, the list of proposing synonymous words with another word will be used for the next search in getting the variety of synonym words related to those thoughts as shown in (Fig. 3). This provided a variety of keywords to be performed as a search query related to

each of his thoughts. The figure below shows the searching of Ibn Khaldun’s thoughts which is ‘*Ekonomi*’ (Economy), then ‘*Perniagaan*’ (Trade) word from the list of proposing synonymous words has been used to get another search process. Table 2 provides Ibn Khaldun’s thoughts as keyword searching results from the DBP thesaurus website. All searches of the keywords had to be in the Malay language (Malay keywords) because respondents for tweets data collection were only Malaysian Twitter users.



Fig. 3: Searching Using Ibn Khaldun’s Thoughts

Table 2. Keyword Searching Results from DBP

Ibn Khaldun’s Thoughts	DBP (Thesaurus)	Definitions
1- Worship God	Ibadat (Worship)	Amal (charity), Amalan (practice), Bakti (devotion), Khidmat (service)
	Sembah (Worship)	Menyembah (worshiping), Sembahan (worship), Penyembah (worshiper)
	Puja (Hymn)	Memuja (adore)
	Abdi (Slave)	Hamba (slave)
	Doa (Prayer)	Permohonan (application), Permintaan (request)
2- Social Integration / Solidarity	Sosial (Social)	Kemasyarakatan (society)
	Sahabat (Friend)	Teman (acquaintance), Sekutu (ally)
	Komunikasi (Communication)	Perantaraan (intermediation), Penyebaran Maklumat (dissemination of information), Penyampaian Maklumat (delivery of information)
	Perpaduan (Unity)	Permuafakatan (consensus), Kesepakatan (agreement), Keharmonian (harmony), Muhibbah (goodwill), Integrasi (integration)
3- Good / Childhood Education	Berpasukan (Team)	Berkumpulan (grouped), Beregu (doubled), Berkawanan (swarmed)
	Pendidikan (Education)	Mendidik (educate), Berpendidikan (educated), Pendidik (educator)
	Belajar (Learn)	Mengajar (teaching), Belajar (learned)
	Kursus (Courses)	Pelajaran (lessons), Latihan (exercises)
4- Teaching al-Quran	Kanak-kanak (Children)	Kekanak-kanakan (childish)
	Budak (Child)	Anak (son)
	Mengajar (Teaching)	Pengajaran (teaching), Pembelajaran (learning)
	Hafal (Memorize)	Ingat (remember), Lancar (smoothly)
5- Good Ethics	Musabaqah (Competition)	Tilawah (recitations), Pertandingan (contests)
	Keadilan (Justice)	Saksama (fair), Sama rata (equal), Sama berat (equal weight), Tidak berat sebelah (impartial), Seimbang (balanced)
	Etika (Etiquette)	Prinsip (principles), Standard (standards)

	Moral (Moral)	Akhlaq (morals), Budi pekerti (manners), Kesopanan (courtesy), Kesantunan (decency), Ketertiban (order)
	Perangai (Conduct)	Kelakuan (conduct), Tingkah laku (behavior), Perlakuan (treatment), Perwatakan (character)
6- Great Scholars	Hebat (Great)	Unggul (superior), Agung (reliable), Terkenal (renowned), Ternama (famous), Terbilang (distinguished)
	Cemerlang (Excellent)	reliable, Luar biasa (extraordinary), Gemilang (glorious), (Kelas satu) first-class, Terkemuka (leading), Ternama (reputable)
	Unggul (Superior)	Ideal (ideal), Teristimewa (special), Terutama (especially), Ulung (outstanding)
	Intelektual (Intellectual)	Cendekiawan (scholars)
	Ahli Falsafah (Philosopher)	Pakar (expert), Jauhari (jeweller), Pandai (clever), Tukang (craftsman), Juru (interpreter), Cerdik pandai (ingenious)
7- Awareness Society	Waspada (Cautious)	Berjaga-jaga (be alert), Berhati-hati (be careful), Beringat-ingat (be mindful), Bersiap-sedia (be prepared), Bersedia (be ready)
	Awas (Caution)	Berawas-awas (watch out), Mengawasi (watch over), Pengawasan (supervise)
	Masyarakat (Community)	Umat manusia (mankind), Kelompok manusia (human group), Bangsa (nation), Kaum (race), Puak (tribe), Khalayak ramai (crowd), Penduduk (population)
8- Good Governance / Khalifah	Tadbir (Administration)	Mentadbir (administer), Pentadbiran (administration), Pentadbir (administrator)
	Urus (Manage)	Berurus (deal), Menguruskan (manage), Pengurusan (management), Pengurus (manager)
	Laksana (Feasible)	Melaksanakan (execute), Pelaksanaan (execution), Pelaksana (executor)
	Kerajaan (Government)	Perintah (order), Pemerintah (government)
	Politik (Politic)	Kenegaraan (nationhood), Kerajaan (governance), Pentadbiran negara (state administration)
	Raja (King)	Perintah pemerintah (government orders)
	Ketua (Master)	Pengetua (principal), Pemimpin (leader), Khalifah (caliph), Pengarah (director), Penghulu (chief)
9- Economy	Pemimpin (Leader)	Ketua kerajaan (head of government), Perdana Menteri (prime minister), Gabenor (governor), Menteri besar (chief minister), Menteri (minister)
	Ekonomi (Economy)	Belanjawan (budget), Hal ehwal kewangan (financial affairs)
	Perniagaan (Business)	Peniaga (merchant)
10- Human Civilization	Perdagangan (Trade)	Urusan jual-beli (dealings), Eksport import (export-import), Penggalasan (bearing)
	Tamadun (Civilization)	Budaya (culture), Cara hidup (way of life), Kemajuan (progress), Kemodenan (modernity)
	Cara Hidup (Life style)	Amalan: adat (practices: customs), Kebiasaan (habits), Kelaziman (norms)

In collecting the tweets from Twitter, the validated 43 keywords which have been identified in the process before will be used as a hashtag to search through Twitter by using R-programming software (Fig. 4). This was to search a dataset in the form of public comments on Twitter obtained by crawling. Posted tweets collected, were then finalized to ensure they represent the issues that are related to Ibn Khaldun's thoughts. All the tweets have been collected then be saved in the Excel file to be sorted and matched with each of Ibn Khaldun's thoughts.

Before collecting the data from Twitter, a secure connection was created between the R console and Twitter. The initial step starts from the connection framework R process to the Twitter API where to be connected there needs to be an authenticated API token first. We must authenticate with Twitter by using consumer Key, Consumer

Secret, Access Token, and Access Secret. The process of registration was done at the Twitter developer’s page at <http://dev.twitter.com> by filling in the application’s details to get the key and access token secret (Fig. 2). R authentication on Twitter is the process of connecting R with Twitter by entering all four API tokens into the script window successfully connected in the R Console window. Thus, the authentication in the R framework with the API token has been successfully carried out and then can be crawled Twitter data according to what the user wants.

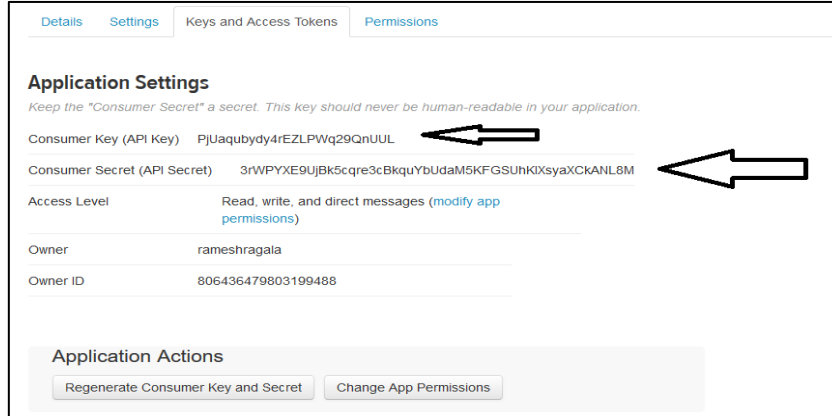


Fig. 4: Screenshot of Twitter Developer Homepage

After the authentication is successful, the Twitter data crawling process is carried out with the verified keyword searched on Twitter by using R-programming to retrieve the tweets through a hashtag. It can be used to retrieve a data set from Twitter for text mining purposes to gather relevant tweets related to the topics by formalizing the query language. When identifying the relevant dataset, the tweets posted were limited within Malaysia. Language filtering was done to extract data that were posted in the bilingual text.

The analysis was carried out using a simple algorithm that was built on R-programming (Fig. 5) by doing Tweet’s classification based on the keyword that had been constructed. Most usefulness of the API is supported, with an inclination towards API calls that are more valuable in data analysis instead of day-by-day interaction. The method we use in data collection is observation, that is by watching, observing and collecting community Twitter data through the Twitter REST API and crawling those Tweets data using R-programming. The data is taken from documents that contain Tweets collected from 1 December 2017 until 31 May 2018 to collect Tweets from users with specific keywords.

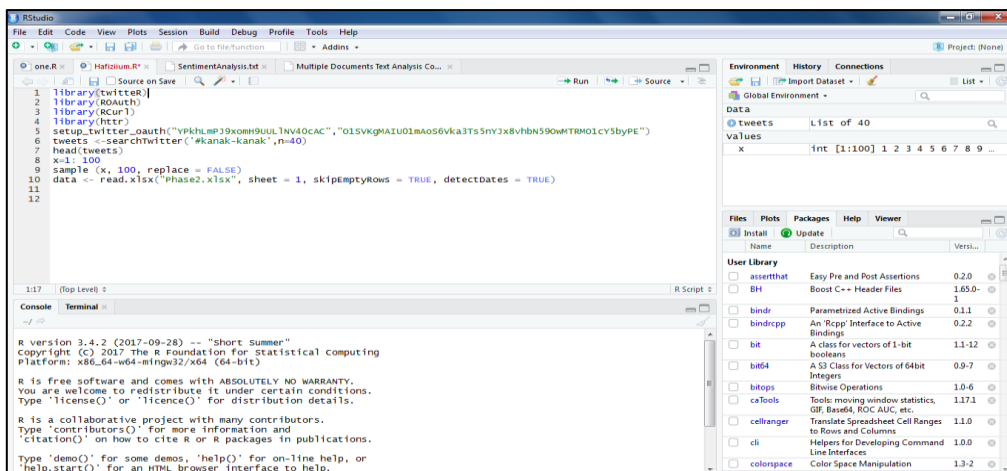


Fig. 5: Screenshot of R-Programming Console Window

R-programming was used to retrieve the tweets from real users on Twitter by using hashtags. Figure 6 displays the structure of statistical computation algorithm structure in R windows.



```

library(twitteR)
library(ROAuth)
library(RCurl)
library(httr)
setup_twitter_oauth("YPkhLmPJ9xomH9UULINV40cAC","O1SVKgMAIU01mAoS6Vka3Ts5nYJ
x8vhbN59OwMTRMO1cY5byPE")
tweets <- searchTwitter("#Perniagaan=40")
head(tweets)
x=1: 100
sample(x, 100, replace = FALSE)
data <- read.xlsx("Phase2.xlsx", sheet = 1, skipEmptyRows = TRUE, detectDates = TRUE)
    
```

Fig. 6: R-Programming Windows

#### 4.0 RESULTS

Data Extraction process through the search API which is the official authentication API for Twitter returns the tweets that match the given string and writes them into the object. A total of 1075 public’s tweets were collected relating to the 43-hashtag keyword in the knowledge acquisition process. After the conversion of tweets data into a data frame, we will save it to a file with .csv in Microsoft Excel. All the retrieved data was taken which will be used as a sample representing the keyword and population. Data that has been obtained from Twitter crawling results is mostly made up of comments and opinions from users. Extracted tweets will be parsed into several categories in the Excel file based on their hashtag keywords search query followed by ten Ibn Khaldun’s thoughts as shown in Fig. 7.

Then, a data cleaning process will be done as is aimed at cleaning data from noises. The data gathered is not pure (Table 3). It contains hashtags, URLs (Uniform Resource Locators), abbreviations, punctuation, stop words, removing spaces, retweets, links and other symbols which don’t express any emotions. It also contains a lot of noise, such as spam and other bot activity, emphasizing the need of preprocessing them before doing any classifications or regression [16]. The tweet is needed to be cleaned to perform better analysis. Some tweets which have been collected might contain many characters that are not easy to read and understood. Data cleaning refers to identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty or coarse data. After performing data cleaning (Table 4), the data is transformed into the required format. Extracted tweets data will be parsed into several categories in the Excel file based on their hashtag keywords search query followed by ten Ibn Khaldun’s thoughts for the next phase of development.

1	Ibn Khaldun's Thoughts	Keywords	Tweets Data
2	Worship Allah	doa	#Traffic standstill ... so the #doa will go this way ... May Allah ease my journey home this...
3			regrann from @alizakim - The #doa #dua of Prophet Yunus (Alaihi Sallam) acknowledging that...
4			#doa keep it up #Allah knowz whatz da #best #doa jgak
5			Amin. :) RT @QuranHadis: O Allah, I am coarse, so make me soft. I am stingy, so make me generous & I am weak, so make me strong
6			performed #Sahabat by #ThreeTimesBrag with my bro @zaidywalawei at Central Market
7	Social Integration / Solidarity	sahabat	Be professional #sahabat don't have to run - - ggagagggag
8			Yeay!!! It's #sahabat
9			It's will forever remains.... In #heart #friendship #sahabat no more bad feeling!
10			#Menteri #Pendidikan bercakap kat sidang dewan rakyat kat #parlimen harini ... tetiba terdetik...
11	Good / Childhood Education	pendidikan	#Pendidikan Islam tidak dipinggir @MuhyiddinYassin" - STEM/TVET and agama, which one does Malaysia want?
12			Lemahnya sist. #Pendidikan #Negara sgt ketara sifatnya apabila wktu #pilihanraya. Masy x mampu tok trima pbezaan pndgn. Beza kita ban & label.
13			Nyte sygg ! Sweet dreams ! #perangai budak dah mengantuk
14	Good Ethics	perangai	#Adriana ngn ayh ngoh rihuhs same #perangai je... ishh
15			@NazmiRahman23: #perangai #taksenonoh #kentut #katkedai #makan #hiks"allah terasa !! Bhahahaha
16			@tankengliang: BN youth chief @khairyk showing off his dancing skills with @petrajayamp
17			#sahabat #hebat @ustazdondaniyal #baik #taat #ceramah best
18	Great Scholars	hebat	faeizmusa @PaanJr #sebab #aku #hebat hahahaha
19			Tahniah Kpd DS @NajibRazak & #Kerajaan #Malaysia Atas Pelancaran #GST - Moga Keputusan Yg Dibuat Bijak Utk Next
20			lain cakap lain buat??? Dah jadi #Kerajaan macam ni ka jadinya?? ayyaaarkk.. dua2 mahu sembelih
21	Good Governance / Politics / Khalifah	kerajaan	#what #government #kerajaan #funny hahahahaha!! #fuck

Fig. 7: Sample of Twitter Data Categorization in Excel

Table 3: Sample of Original Tweets Data from Users Before the Cleaning Process

<b>Users Tweets from Twitter</b>
"Batak boy. Batak boy." Sorry, it was not me.
Focus #HariIni8 @501Awani: 1. A mother and her boyfriend were arrested for the death of her six-year-old child. 2. Believed to contain worms, the Ministry of Health orders for a recall on sardine products from China. 3. Harumanis innovation has improved the economy of farmers in Perlis
Please do not call me "akak (elder sister)". I am still your basic happy little child.
SchaMcDonalds Disgusted but still wanting it, do you not? *Laugh out loud* Do not be dirty minded.
Want to contribute to your beloved #Malaysia? Want greater #Unity? Join #JPAM (Malaysia Civil Defence Force) Volunteers. Visit...
#1News: Britain's Court of Appeal dismissed the appeal from #AlfieEvans' parents to continue life support and take him to Italy for treatment. The 23-month-old child has been in a coma for a year.
A 20% discount offer for full payment of PTPTN (National Higher Education Fund Corporation) loans in one receipt. Offer will end 31st March 2015.
The walk with @zurianichan and @musthaqimz was really enjoyable. Thank you for the pizza treat.
yazleenyazli: Oh, wow! Is this true? Haha.
FakhiraAzis: "And now, children. The magic show!"
Tun Mahathir was asked by children, "Why are you still working?"

Table 4: Sample of Tweets Posts from Users After the Cleaning Process

<b>Users Tweets from Twitter</b>
Batak boy Batak boy Sorry it was not me
Focus HariIni8 at 501Awani 1. A mother and her boyfriend were arrested for the death of her six year old child 2. Believed to contain worms the Ministry of Health orders for a recall on sardine products from China 3. Harumanis innovation has improved the economy of farmers in Perlis
Please do not call me akak I am still your basic happy little child
SchaMcDonalds Disgusted but still wanting it do you not Laugh out loud Do not be dirty minded
Want to contribute to your beloved Malaysia Want greater Unity Join JPAM Volunteers Visit
1News Britain's Court of Appeal dismissed the appeal from Alfie Evans parents to continue life support and take him to Italy for treatment the 23 month old child has been in a coma for a year
A 20 per cent discount offer for full payment of PTPTN loans in one receipt Offer will end 31st March 2015
The walk with zurianichan and musthaqimz was really enjoyable Thank you for the pizza treat
yazleenyazli Oh wow Is this true Haha
FakhiraAzis and now children the magic show
Tun Mahathir was asked by children Why are you still working

Tweets collected were from within Malaysian users only. The challenge here is that informal writings were mostly seen on SNS platforms. Users tend to not follow syntactic formats that are grammatically correct in writing and they are often messy and ambiguous. They moreover illustrated that although Malay is the dominant language and dominated Twitter in Malaysia, some of the messages on Twitter also were from those who spoke English. Most of the tweets in this data set are in Malaysia, and most of the users within the data set are tweeting in Malay. Since most of the respondents wrote their tweets in the Malay language and some of them in improper English language, all the tweets were translated into correct English. The translations were then submitted and proofread by an expert translator from The Malaysian Institute of Translation and Books (ITBM) so they were in line with expectations. Translated documents would assist in the development of concepts in ontology through the formatting of the text. Table 5 depicted the sample of the results from the translation process for tweets data collection.

Table 5: Sample of Tweets Translation

Tweets in Malay	Tweets in English
Selamat hari kanak kanak dari kami bekas kanak kanak	Happy Children's Day by us the former children
Andai waktunya telah tiba, sedetik pun tak bisa di tunggu lagi	When the time has come not a second can wait
Sgtt merindui suasana ini	I miss this atmosphere very much
Dah nak akhir zaman beringat lah semua	It is the end of times may we be reminded
lupakan minat seketika fokus untuk masa depan aku perlu berjaya dalam hidup	Forget about momentary pleasures; focus on the future I need to be successful in life
Belajor..demi diri sendiri,demi keluarga,demi syarikat,demi Allah!!	Learn for the sake of my own, my family, the company for Allah
Melentur buluh biar dari rebungunya.	Strike the iron while it is hot
Kita berubat sampai sihat	We seek medical help until we regain health
budak pailang mungkin tak perlu kita bercakap kalau kita tak memahami	Gangster boy there may be no need for us to talk if we do not understand each other
lain mcm permainan budak skrg halussss	Children today play a different kind of game Very subtle
Kawan-kawan kata kamu sombong, maka ada yang tak kena pada diri kamu	Friends say you are arrogant then something is wrong with you
Tweet lah saya..fb skng dah mengarut karoot komedia laa	Just Tweet me. Facebook is now full of nonsense
kalau mcm tu baik aku buat google tu jd sahabat aku.. Thanks	If that is the case I should make Google my friend Thank you
Perpaduan dari mereka yang bertanggungjawab menjaga perpaduan	Union from those responsible for maintaining unity
Allah menilai setiap amal ibadat yg kita lakukan sehari-hari...	Allah evaluates every act of worship we do every day
Kerja itu ibadat. Tugas itu ibadat. Buat kerana Allah swt	Work is worship Duty is worship For the sake of Allah swt
Bekas hakim dan isteri dera kanak-kanak dipenjara setahun	Former judge and wife convicted of child abuse and sentenced to a year in prison

After the translations have been verified by experts, sentiment analysis processes will be run in identifying the ethical issues on tweets data collection from Twitter. It is a study that analyses users' opinions, sentiments, evaluations, attitudes, and emotions regarding entities such as products, services, individuals, issues, events, topics, organizations, and their attributes. Sentiment analysis is known as opinion mining and focuses on identifying users' opinions that express positive, neutral or negative sentiments [14]. Sentiment classification will recognize opinions in a text and categorize them as positive, negative, or neutral, based on the emotions expressed by users and appearing in a given text.

The sentiment analysis tool will assist in effectively analyzing text conversations and evaluating the tone, intent, and emotion behind each message. To maximize the analysis process in identifying the sentiment analysis, language barriers must be eliminated because some of the systems only support certain languages. In this case, the Parallel Dots application works in fourteen different languages except for the Malay language. Due to this, tweets had to be translated into English to be run and identify the sentiment analysis. Parallel Dots gives capable DIY (Do It Yourself) tools and customized solutions to turn the organization into an AI powerhouse [14]. It gives AI-driven visual analytics and content analytics APIs. Parallel Dots sentiment analysis is able to help in analyzing contents or open supposition of certain points over social media. By analyzing tweet statements from chosen users on Twitter, the purpose of intelligent content analytics can be utilized to recognize the positive, neutral or negative sentiments within the tweets. The tweets posts which have been translated by an expert translator were arranged in Excel sheets and run-on Parallel Dots (Fig. 8). Through Excel add-in, it made a difference in analyzing unstructured texts inside the spreadsheet by utilizing the form of API functions provided for case 'parallel dots sentiment for sentiment analysis' to perform the analysis.

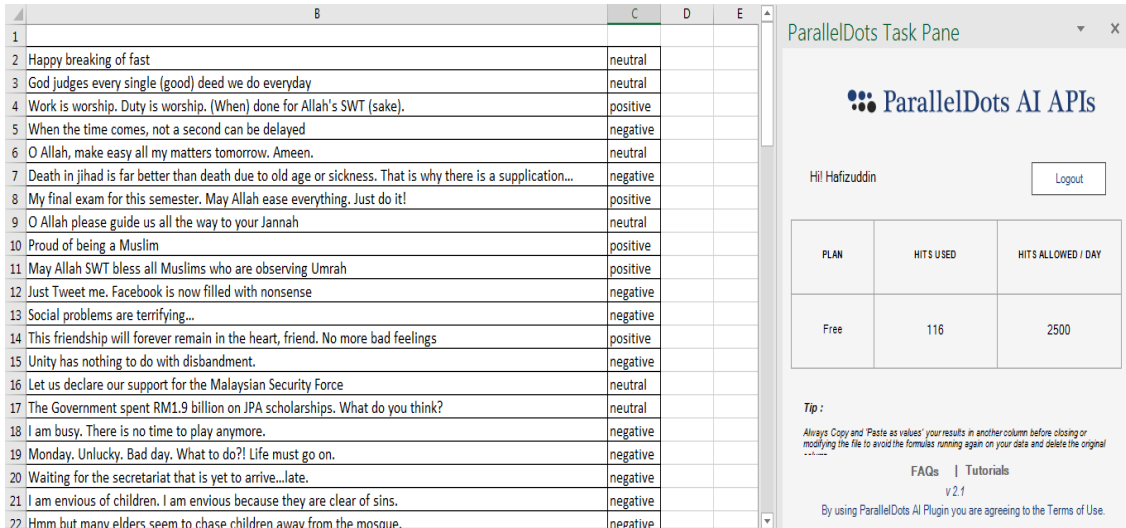


Fig. 8: Screenshot of Parallel Dots Homepage in Excel Sheets

Table 6: Sentiment Analysis Results Run in Parallel Dots

Tweets from Twitter	Parallel Dots
May Allah SWT bless all Muslims who are observing Umrah	positive
Just Tweet me Facebook is now filled with nonsense	negative
Social problems are terrifying	negative
This friendship will forever remain in the heart, friend. No more bad feelings	positive
Let us declare our support for the Malaysian Security Force	neutral
Whenever it is time for fun, there is always something blocking the way	negative
Monday Unlucky Bad day What to do Life must go on	negative
UNICEF (United Nations Children’s Emergency Fund) is also helping Syrian refugee children so please do our part to help them	neutral
The children are having fun	positive
We are no longer children who have no obligation and responsibility	negative
I read my old childish posts on Facebook where it is just so mortifying at all levels that I had to delete most of it	negative
Even if we are blocked and sabotaged we will still fight	negative
God-willing we will win	positive
You need to have courage to try something new	neutral
Have a blessed day	positive
Do not judge me by my past I do not live there anymore	negative

The text analysis brings out the exact sentiment analysis from tweets data. Table 6 illustrated the sample of sentiment analysis results from Parallel Dots software. The sentiment analysis in Parallel Dots was run on the collected 1075 public’s tweets and identified that 700 of the tweets are positive statements, 229 of the tweets are neutral statements and 146 of the tweets are negative statements. The final results were checked and verified by internal experts first and then external experts to ensure that the final output is in line with the aims of this research.

This research focused on negative tweets as their objective was ‘To identify the ethical issues in SNS related to Ibn Khaldun’s thoughts’ through the knowledge acquisition process (Table 7). The positive and neutral tweets contribute to the good supporting points of Ibn Khaldun’s thoughts and highlight the good deed within the SNS users. It becomes the additional information to contribute to the good tweets from users and support as a solution for the identified ethical issues in tweets. From the tweets collection, users’ statements may be related to Ibn Khaldun thoughts in understanding the concept of human behavior and sociology based on ten Ibn Khaldun’s thoughts. These also help in determining the nature and characteristics of Ibn Khaldun’s theory of social analyses and the science of civilization to acquire a better understanding of it. In the future, Twitter users will benefit from automatically discovering solutions to ethical issues identified in tweets using the search box by pasting the tweets. This concept may also be used by Malaysian authorities in communications and multimedia by investigating the usefulness of thoughts with current concerns in SNS highlighted in their report to be given to Malaysians as a new

initiative. It may also be a new prospect to bring forward by collaborating with top local technology and service providers to deliver an awareness and advocacy campaign to communities about being ethical in cyberspace.

The *Muqaddimah* book contains valuable information in explaining how each of these sources of social and civilization theory gives a solution to issues encountered in SNS. Ibn Khaldun's work represents a good choice for study because his work is also a development of the science of society. The simple algorithm in R-programming is capable of facilitating the process of extracting data from Twitter, such that it is easily known to the public tweets that exist in a given period and this can be used as reference material for the further development process.

Table 7: Sample of Ethical Issues from Tweets in Twitter

<b>Negative Tweets from Twitter</b>
Just Tweet me Facebook is now full of nonsense
The social phenomenon is terrifying Sigh
Monday Unlucky Bad day What to do Life must go on
Do you understand that I am drawing a nerve
After some inspection and observation you are truly awful
Alright done Now we will be making some trouble while looking for a clause
We can never behave around each other Will we ever grow up
Killer subjects tomorrow Kill me right now
Abdul Hadi Awang and Najib Razak should know better than to misuse Islam for their benefits Do we want to slander Islam just to cover their corruption
Even if we are blocked and sabotaged we will still fight
Hypocrites and munafiq of PAS (Malaysian Islamic Party) need to come to their senses Do not be deceived by PAS misusing religion
O PAS voters Please do not embarrass and make a fool of yourself with PAS play on religion
Only THIEVES like Najib Razak want to stay in power oppress and persecute the people
So this is the RASCAL and TRAITOR of the people the Malays who betrayed Malays
Such a shame Shame on you SPR Even the Prime Minister of Malaysia disagrees with the decision
The world lately has become strange The tyrant is acquitted while the honorable are convicted Where is justice
Water problem in Pahang Is this not Najib Razak state legislative Najib is such a HYPOCRITE
When Nik Aziz is betrayed by a devil named
BloodyIzz There are truant and immoral girls everywhere
Feel the hatred towards someone because he does not have the decency to cover the disgrace of others
Eugh shameless Claim yourself government and others as the opposition
A Year Five boy bullying a Year Three boy

## 5.0 CONCLUSION

Ethics is important in life. It is a light that guides us in the right direction and saves us from not being lost. Ibn Khaldun has understood the importance of ethics decades ago and his thoughts and concerns regarding this are still relevant. Many Muslim scholars like him have also given deep concern about the ethics such as Al-Ghazali, Abdullah Nasih 'Ulwan, etc. Ethics or *Akhlaq* in Arabic is a person's character reflecting both sides, good or bad. An ethically good person lives a life without harming himself/herself or anyone else. People around him/her feel safe and get the vibe of his/her good behavior. This is how the wisdom of a good person spreads around and benefits others. Wisdom is the combination of knowledge and virtue. This knowledge has to be pure in a sense that does not contaminate with polluted information. The knowledge that helps people identify what is right and wrong is free of pollution.

Accordingly, SNS in this era is flooded with information. Generations now and ahead are dependent on it. It is very common to be lost in this flood if no firm guidance or support system is not found. For example, the study of tweets in this study; it is seen how a 280 characters' simple tweet could carry a negative and positive impact on people's lives. To make those more understandable, this study aimed to generate an ethical consideration analyzing those collected tweets and come up with sentiment analysis that shows what type of tweets are negative and what is positive according to Ibn Khaldun's thoughts. If ethics are not followed in twitter like other SNS, it causes issues like cyberbullying, scams, privacy theft, and sexting; even it leads to committing suicidal cases

sometimes. Therefore, identification of ethical issues from tweets could be useful so that good and bad can be distinguished.

Similarly, if an ontology platform is built with the results of this and similar studies, sentiment analysis based on ethical issues on SNS can play a more vital role for the betterment of people. Ontology can build up a semantic co-relation with the ethical concept and the sentiment analysis. On this account, this paper suggests a further study on semantic analysis of Ibn Khaldun's thoughts with the results found after sentiment analysis on tweets. In this regard, protégé can be used to build up the domain models for a collaborative view of ethics in SNS. The collected tweets will be presented as the ethical issues related to the SNS user behaviour and further contributed to developing Ibn Khaldun's model using the ontology concepts. The developed ontology can help to provide an understanding of the structure and design the ontology development of the ethical issues from the negative statements in Twitter related to Ibn Khaldun's thoughts as a solution. The ontology model could be applied in a knowledge environment by exploiting the tweets which could be linked to the concepts of human sociology that are related to the SNS user behaviour.

## 6.0 ACKNOWLEDGMENT

This research was supported by the Fundamental Research Grant Scheme (FRGS/1/2018/ICT04/UIAM/02/2) Research Project 19-037-0645 provided by the Ministry of Education Malaysia (MOE). The authors also would like to acknowledge International Islamic University Malaysia and the Islamic Science University of Malaysia for their support.

## REFERENCES

- [1] E. M. G. Younis, "Sentiment Analysis and Text Mining for Social Media Microblogs using Open-Source Tools: An Empirical Study Article. International," *Int. J. Comput. Appl.*, vol. 112, no. February, pp. 44–48, 2005.
- [2] S. Singh and S. S. Choudhary, "Social Media Data Analysis: Twitter Sentimental Analysis Using R Language," *Int. J. Adv. Electron. Comput. Sci.*, vol. 4, no. 11, pp. 13–17, 2017.
- [3] K. Weller, "Accepting the challenges of social media research," *Online Inf. Rev.*, vol. 39, no. 3, pp. 281–289, Jun. 2015, Doi: 10.1108/OIR-03-2015-0069.
- [4] N. Danneman and R. Heimann, *Social Media Mining with R Deploy*. Birmingham: Packt Publishing Ltd, 2014.
- [5] S. Jones, "Doing social network ethics: a critical, interdisciplinary approach," *Inf. Technol. People*, vol. 30, no. 4, pp. 910–926, 2017, Doi: 10.1108/ITP-04-2016-0093.
- [6] S. Vallor, "Social Networking and Ethics (Stanford Encyclopedia of Philosophy)," *Stanford Encycl. Philos.*, pp. 1–16, 2015, Accessed: Jul. 10, 2021. [Online]. Available: <https://plato.stanford.edu/entries/ethics-social-networking/>.
- [7] J. Lichy, J. D. Farquhar, and M. Kachour, "Entrepreneurship via social networks – 'connected woman' in Lebanon," *Qual. Mark. Res.*, 2020, doi: 10.1108/QMR-01-2020-0004.
- [8] M. H. Shuhari, "The Importance of using Current Technology in the Study of Islamic Ethics," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 3, pp. 3945–3949, Jun. 2020, Doi:10.30534/ijtcse/2020/217932020.
- [9] S. Abdullah, "Ibn Khaldun's Discourse on the Importance of Knowledge and Ethics in Youth Human Capital Development," *J. Hum. Dev. Commun.*, vol. 3, no. August 2013, pp. 59–77, 2014.
- [10] W. Wolny, "Knowledge Gained from Twitter Data," vol. 8, pp. 1133–1136, 2016, doi: 10.15439/2016F149.
- [11] H. Buarki and B. Alkhateeb, "Use of hashtags to retrieve information on the web," *Electron. Libr.*, vol. 36, no. 2, pp. 286–304, 2018, doi: 10.1108/EL-01-2017-0011.

- [12] R. Ihaka and R. Gentleman, "R: A Language for Data Analysis and Graphics," *J. Comput. Graph. Stat.*, vol. 5, no. 3, pp. 299–314, 1996, doi: 10.1080/10618600.1996.10474713.
- [13] S. Sulaiman, "Komunikasi Berkesan Di Media Sosial: Pendekatan Ibnu Khaldun," *Proceeding 2nd Int. Conf. Econ. Bank. 2016*, vol. 2016, no. November, pp. 253–261, 2016.
- [14] ParallelDots; World Class NLP APIs for Text Analysis. <https://komprehend.io/>, 2021.
- [15] S. Salleh, "Komunikasi Berkesan di Media Sosial: Pendekatan Ibn Khaldun," in *2nd International Conference on Economics & Banking 2016 (2nd ICEB)*, 2016.
- [16] B. Vimala, K. Mohammed, A. R. Hajar, I. Nazari, "Social Media Analytics Using Sentiment and Content Analyses on the 2018 Malaysia's General Election," *Malaysian Journal of Computer Science*, Vol. 34 (2), 2021.
- [17] I. Khaldun. *The Muqaddimah: an introduction to history (Translated by Franz Rosenthal)*. New Jersey: Princeton University Press, 2005.