_____

CONTENTS

_____

COAS

_____

Perceived Influence of Social Media
Usage Among Youth: A Survey

Olalekan Moses Olayemi

Nigerian Institute of Medical Research, Lagos, NIGERIA
Department of Library & Information Technology

Abstract

The use of social media has significantly changed means of communication and reshaped individual habits, especially the younger generation. These changes have vital implications on youth. This study investigated the perceived influence of social media usage among youth in Lagos State, Nigeria. Survey design was adopted for the study. The population of the study were youth and, the sample size was 200 participants. Questionnaire was used to collect the data. The findings revealed that the most common social media accounts used by the respondents were Facebook and Whatsapp. The findings also show that a majority of the youth agreed to the negative influences of social media usage on their behavior and lifestyle. The study concluded that the current findings may help to better understand the influence of social media use on the youth, and hence inform strategies to balance between control and participation as well as curtailed its negative effect.

Keywords: internet, online, youth, social media, social networking sites.

1. Introduction

One of the important developments associated with online internet presence is the advent of social media. The emergence of the internet and, more recently, social media has significantly changed personal habits and reshape lifestyle, particularly among the younger generation (Tarimo, 2016). The addition of smartphones and other online mobile devices such as the tablet, Ipads brought about the added value of mobile internet usage and access to information. Coupled with social media applications, this technology enabled people to use social networking sites on the go, from any location at all the time (Almenayes, 2014).

Annapoorna, Reshma and Sawad (2015) defined social media as websites and applications that enable users to create and share content or to participate in social networking. It is one of the most modern and favorite form of social platform including many features and social characteristics in it. Social media such as Facebook, Twitter, LinkedIn, Instagram, Whatsapp, amongst others has brought a remarkable change in how people access and share information (Fallahi, 2011). It serves as a channel to communicate, texting, images sharing, audio and video sharing, fast publishing and linking with people all over the world. It is also a cheapest fast access to the world, so it is very essential for all age groups. The youth, however, remains the most prominent users of social media. This, of course, is not surprising as the digital age is widely believed to belong to the youth. Despite its benefits and usefulness, social media has its own effects

Correspondence: Olalekan Moses Olayemi, Nigerian Institute of Medical Research, Department of Library & Information Technology, Lagos, NIGERIA. E-mail: lekus2000@yahoo.com.

such as causing moral degeneration among the youth and the society at large. This is alarming and call for concern.

2. Statement of the problem

According to Edeh (2018) social media addiction is becoming an increasing problem that has negative consequences on youth as many have become obsessed with the media. Many young people are found so attach to their smartphone that they hardly sleep over night surfing the internet / social media. Many end up sleeping off with their phone in their hand and when they wake up in the morning, they first check the status/update (Dimka, 2018). In addition, while using some social media sites, youth can easily fall prey to sexting, pornography, cyberbullying, and other anti-social behavior. Other probable negative consequences include a lack of privacy, reduced learning and research abilities, decreased real human contact and writing skills, increased vulnerability to crime, emotional pain and anxiety, and severe isolation amongst other vices which can ruin career and purposeful living (Ononogbu & Chiroma, 2018).

3. Objectives of the study

The general objective of this study is to investigate the perceived influence of social media usage among youth in Lagos State, Nigeria. The specific objectives are:

1. To find out the use of social media among youth;

2. To find out the medium of access to social media;

3. To find out the most commonly used social media among the respondents;

4. To find out their frequency of social media use;

5. To ascertain the perceived benefits and effect of social media usage among the respondents;

6. To ascertain the problems the youth encountered when using social media.

4. Literature review

4.1 *Social media use and youth*

In today's world, people — particularly young people — are continually finding and adapting new ways of communicating electronically to meet their needs. Majority of the younger generation have moved quickly from electronic media like as television viewers and radio listeners to the social media. The number of youths using social media platforms continues to increase and the statistics are overwhelming on how often they do so each day. In Nigeria, citing the Minister of Communication Shittu Adebayo said "about 75 percent of Nigeria's population that use the internet are on social media", of which majority of them are youth (Ogunkola, 2018). By implication, a large part of this generation's social, emotional and even mental development takes place online (Ononogbu & Chiroma, 2018).

Social media has become a popular trend today, especially among young people, and as such many youths are more concerned about the social platforms available on any mobile device. The use of social media helps the youth to stay connected. Social media is readily available everywhere, and at the same time allows connection and information flow to happen quickly

(Tarimo, 2016). For many young people, this is their preferred means of communication and it has been identified as an effective tool. Nevertheless, the negative ways in which youth embrace the use of social media is high and alarming. The amount of time that youth, and even adults, have been spending with these social networks has caused a wave of concerns about their lifestyle and content of these sites, as well as the effects of these on social relationships and belief (Nyland & Near, 2007; Edeh, 2018). The outcome of the over-dependence and misuse of the social media takes so much of the youths' time off their academic work, morals and lifestyle. The younger generation tends to spend more of their time online chatting, posting selfies, keeping up in the latest entertainment news and exhibiting celebrity mania and all this tends to have a negative impact in their reading culture and career (Nkordeh, Olowononi & Ibinabo Bob-Manuel, 2017).

Review of related literature revealed that various studies have been conducted on the use of social media among youth globally. However, there is paucity of empirical studies that specifically address issues relating to the perceived influence of social usage among the youth in Lagos State, Nigeria. Akpan, Akwoano and Senam (2013) studied found out that youth in Uyo, Nigeria registered in at least 46 social networking websites which indicates that they are internet-savvy and is part of the global social networking. They also found that the youth were active users of social networking websites and accessed it for various purposes and gratifications. It also established that the youth were not always fulfilled in their expectations of social networking sites, although they made use of these websites daily. Ezumah (2013) studied the use of social media by college students who are predominantly youth and ranging in age from 18 to 28 years. The findings revealed Facebook emerged as the preferred social media networking site followed by Twitter, while LinkedIn was the least popular site among this group. The findings also revealed so many reasons why they use social media include keeping in touch with friends and family, sharing photos, and entertainment among others.

On the benefits derived from the usage, Dhaha and Igale (2013) studied Facebook usage among Somali youth. The results suggest that Somali youth are motivated by virtual companionship escape, interpersonal habitual entertainment, self-description of own country, self-expression, information seeking, and passing time gratifications. Baglari et al. (2020) explored the social media use among youth in Karnataka, India. The finding revealed that the respondents in the study, opinioned that social media is an effective communication tool that enables users to communicate with other social media users instantly as well as sharing information such as videos and images.

Amofah-Serwaa and Dadzie (2015) explored the implication of social media use on child behaviour in a basic school in Ghana. The outcome of the study indicated that the negative effects of social media observed among the pupils involve distraction of pupils' attention from their studies, frequent use of "pidgin English" as well as unnecessary fashion consciousness. Similarly, Onah and Nche (2014) investigated the moral implication of the phenomenon of social media in Nigeria especially among the youth. They found out that the social media phenomenon is overwhelmingly popular among Nigerian youth as well as one of the major contributors to moral degeneration in the country. Edogor, Jonah and Ojo (2014) investigated the influence of social media on the youths' usage of traditional mass media in Nigeria. The study revealed that social media create wider links enabling youth audiences using them to access traditional mass media online. The study also found that the youth use social media to date, indulge in cybercrimes, discuss national issues like economic, politics, and socio-religious matters; but most of them do not use the media to communicate to their national leaders on aforementioned issues. Umeogu and Ojiakor (2014) sought to ascertain what the internet especially social media has done to and for Nigerian youth. The findings revealed that the internet have negatively affected education, morality and culture of Nigerian youth amongst others.

Ambuchi (2015) studied the negative impact of media on the morality of the youth in Nairobi, Kenya. The findings revealed that the negative effects of the media have impacted the

youth as media access and media use (internet) among the youth is rampant and subsequently consumes a lot of their time. Shu-Sha Angie Guan and Subrahmanyam (2009) conducted a study on the risks and opportunities of internet use on the youth. The results show that online risks such as addiction, cyber bullying, and sexual solicitation are associated with negative consequences among the youth. Shabir et al. (2014) surveyed the impact of social networking sites in the changing mind-set of the youth. The finding indicated that, majority of the respondents agreed that use of social media is deteriorating social norms and its use is affecting negatively on study of youth. More so, the findings revealed that social media promotes unethical pictures, video clips and images among youth, as well as anti-religious post and links that create hatred among peoples of different communities.

5. Theoretical framework

5. 1 *Use and Gratification theory*

This study deployed the uses and gratifications theory propounded by Elihu Katz, Jay Blumler and Micahel Gurevitch in 1974. The use and gratification theory is a well-known approach to understanding of media utilization among individuals. The theory focuses on how users proactively engaged with the use of media that will meet a given need. The theory posits that users are responsible for choosing media to meet their needs. The approach suggests that people use media to fulfil specific purposes (Katz, Blumler & Gurevitch, 1974). The use and gratification theory seeks to explain the reason behind people using a particular media and what motivates people to continually engage in using such media in order to derive gratifications to fulfill those needs. The theory also helps researchers to gain an understanding on why people switch from one media to another.

Unlike earlier theories that suggested that media opinion might shape individual and their belief, the uses and gratification theory suggest that media audiences are not passive, and that people choose media messages to gratify their needs (Adelabu, Sanusi & Esiri, 2015). As such, media consumers decide the extent to which media has an influence on them as well as the idea that users decide media alternatives merely as a means to an end. Consequently, the adoption of the uses and gratification theory is informed by the fact that people consciously choose the medium (social media) that could satisfy their needs. The extent to which a social media platform is used to meet a demand is a measure of its effects and gratification. As such, users / media consumers actively choose specific media content based on their need. If there are any effects, either positive or negative, these are consciously or at least actionably intended. Uses and gratifications theory offers a way to explain the proliferation of social media – why various social media outlets are so popular and commonly used (Matei, 2010). This theory was adopted to explain how and why youth in all countries of the world develop so much habit choice of using social media to the very extent that it has become part of their daily lifestyle.

6. Methodology

This study is quantitative in nature. A survey design was adopted for the study. The population of the study comprised youth that attended the RECALP. RECALP stands for "Relationship, Entrepreneurship, Career, Leadership and Politics." RECALP is a non-denomination and inter-denomination program that gather youth from all works of life to venue to, educate and teach them on various issues such as relationship, entrepreneurship, career, leadership and politics by different renowned facilitators. Purposive sampling technique was used for this study. Questionnaire was the research instrument used for the data collection. To ensure validity of the questionnaire, face and content validity was conducted through a panel of experts who have years of experience research in a related field. The Cronbach's Alpha was used to test the

_____

internal consistency of the instrument. The overall value for the entire instrument yielded 0.84, which indicates that the instrument is very reliable enough for the conduct of the study. A sample size of two hundred (200) participants was drawn for the study. Israel (1992) suggested that nearly sample size will suffice if descriptive statistics are to be used. In a similar study, Dimitrov and Nongkynrih (2017) used the sample size. The collected data were arranged, coded, tabulated using frequency, percentages, mean and standard deviation, and analyzed using SPSS version 21 for Windows.

7. Analysis of results

Out of the two hundred (200) copies of the questionnaires administered to the respondents of the study, one hundred and sixty-eight (168) copies were retrieved, out of which one hundred and fifty-nine (159) representing (79.5%) was found useful for the data analysis.

Table 1. Demographics of the respondents

Gender		Frequency	Percentage (%)
	Male	71	44.7
	Female	88	55.3
	Total	**159**	**100**
Age			
	15-20	41	25.8
	21-25	50	31.4
	25-30	68	42.8
	Total	**159**	**100**
Marital status			
	Single	147	92.5
	Married	11	6.9
	Missing value	1	.6
	Total	**159**	**100**
Educational qualification			
	WASC/SSCE	59	37.1
	NCE/OND	40	25.2
	HND/BSC	41	25.8
	Masters	11	6.9
	Others	8	5.0
	Total	**159**	**100**
Occupation			
	Student	88	55.3
	Business	21	13.2
	Self-employed	22	13.8
	Unemployed	7	4.4
	Civil Servant	7	4.4
	Others	13	8.2
	Missing value	1	.6
	Total	**159**	**100**

From table 1 above, there was a slightly higher number of females 88 (55.3%) in this study compared to males. The respondents of the study represent different age bracket. The highest respondents 68 (42.8%) fell under the age bracket of 25-30, while the lowest category 41 (25.8%) were between the age bracket of 15-20 years. From the analysis, it shows that 147 (92.5%) respondents were single, while 11 (6.9%) respondents were married. This shows that majority of

_____

the respondents are single people who probably like to spend more time to consume social media content.

The highest educational qualification of the respondents revealed that 59 (37.1%) were WASC/SSCE certificates holders. Analysis on the occupation distribution of the respondents shows that a majority 88 (55.3%) were student.

Table 2. Social media account

	Frequency	Percentage (%)
Yes	159	100
No	0	0
Total	**159**	**100**

From the table 2 above, all the respondents 159 (100%) had social media account.

Table 3. Devices used to access social media

Devices	Frequency (n=159)	Percentage (%)
Cellphone	146	91.8
Desktop	24	15.1
Laptops/Ipad	29	18.2
Parents' phones	7	4.4
Tablet	14	8.8

*Multiple responses

The respondents were asked the devices they use to access social media. From the table 3,146 (91.8%) owned a cellphone, laptops/Ipad 29 (18.2%), desktop 24 (15.1%), tablets 14 (8.8%), and parents' phones 7 (4.4%).

Table 4. Types of social media account

Social Media	Frequency (n=159)	Percentage (%)
Facebook	155	97.5
Twitter	65	40.9
Whatsapp	136	85.5
Instagram	102	64.2
LinkedIn	41	25.8

*Multiple responses

From the table 4,155 (97.5%) had social media account with Facebook, 136 (85.5%) had accounts with Whatsapp, 102 (64.2%) respondents had accounts with Instagram, 65 (40.9 %) respondents had accounts with Twitter, and 41 (25.8%) respondents had accounts with LinkedIn.

Table 5. Frequency of Social Media Use

Social media	Very often	Often	Sometimes	Not often / Rarely
Facebook	42(26.4%)	53(33.3%)	34(21.4%)	27(17.0%)
Whatsapp	85(53.5%)	35(22.0%)	16(10.1%)	9(5.7%)
Twitter	13(8.2%)	10(6.3%)	28(17.6%)	30(18.9%)
Instagram	17(10.7%)	24(15.1%)	37(23.3%)	41(25.8%)
LinkedIn	8(5.0%)	4(2.5%)	15(9.4%)	43(27.0%)

The respondents were asked to indicate the frequency with which they used different social media network. The results on table 5 indicate that Whatsapp 85 (53.5%) and Facebook 42 (26.4%) were used more frequently used than the other social media network as compared with Twitter 17 (10.7%) and LinkedIn 8 (5.0%).
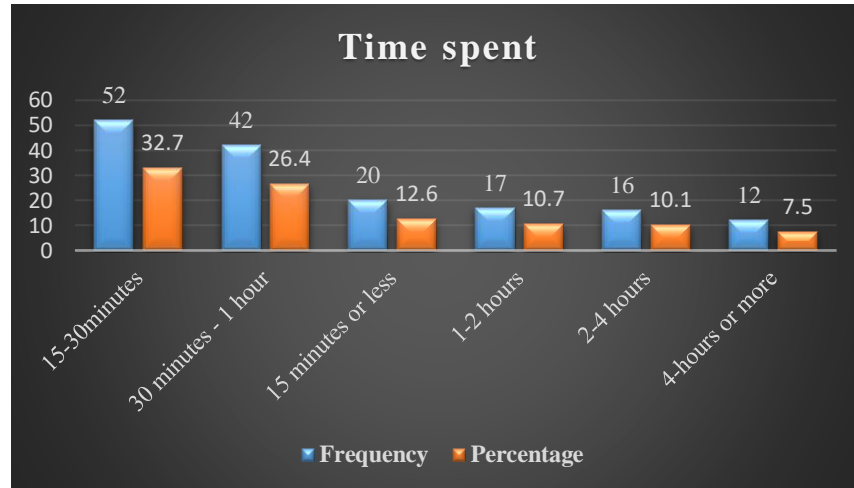
Figure 1. Time spent on social media

The respondents were asked how many minutes /hours they spent on the social media in a typical session. From the figure 1 above, 52 (32.7%) respondents spent 15-30 minutes, 42 (26.4%) spend 30 minutes – 1 hour, 20 (12.6%) respondents spend 15 minutes or less, 17 (10.7%) spend 1-2 hours, 16 (10.1%) spend 2-4 hours, and 12 (7.5%) spend 4 hours or more. The usage be for updating status, finding trending issues, posting and commenting / responding friends' timeline every moments or business purposes.

Table 6. Purpose of using social media

Items	Frequency (n=159)	Percent age (%)	Mean ($\bar{x}$)
For communicating and interaction with friends	143	89.9	1.10
For online learning	137	86.2	1.14
For finding friends online	58	36.5	1.64
For private messaging	74	46.5	1.53
For general networking	76	47.8	1.52
For updating profile information's and uploading photos	66	41.5	1.58
Entertainment	77	48.4	1.52
To promote/doing business	98	61.6	1.34
To while away time	53	33.3	1.67

Table 6 above shows the benefit derived using social media by the respondent's understudy. They include communicating and interacting with friends ($\bar{x}$=1.10), online learning ($\bar{x}$=1.14), for promoting/doing business ($\bar{x}$=1.34), entertainment ($\bar{x}$=1.52), general networking ($\bar{x}$=1.52), and private messaging ($\bar{x}$=1.53), for updating profile information's and uploading photos ($\bar{x}$=1.58), for finding friends online ($\bar{x}$=1.64), and to while away time ($\bar{x}$=1.67).

Table 7. Perceived benefits derived on the use of social media

Items	Frequency (n=159)	Percentage (%)	Mean($\bar{x}$)
It encourages virtual meeting with people	102	64.2	1.36
It increases self-esteem and well-being	76	47.8	1.52
It helps in research and learning	144	90.6	1.09
It helps in sharing religious activities	104	65.4	1.35
It supports collaborative and peer to-peer learning	97	61.0	1.45
It helps in strengthening interpersonal relationships	110	69.2	1.31
It promotes reading and writing web skills	95	59.7	1.40

The respondents were asked the benefit derived using social media. Table 7 reveals that, ($\bar{x}$=1.09) respondents said it helps in research and learning, ($\bar{x}$=1.31) it helps in strengthening interpersonal relationships, ($\bar{x}$=1.35) it helps in sharing religious activities, ($\bar{x}$=1.36) respondents said it encourages virtual meeting with people, ($\bar{x}$=1.45) it supports collaborative and peer to-peer learning, ($\bar{x}$=1.40) it promotes read and write web skills, and ($\bar{x}$=1.52) it increases self-esteem and well-being.

Table 8. Perceived effect of using social media

Items	Frequency (n=159)	Percentage (%)	Mean ($\bar{x}$)
Many young people end up wasting a lot of their time on games and chatting on social networks	146	91.8	1.08
Use of social media is deteriorating our social norms	89	56.0	1.44
Social media promotes unethical/sexual pictures, video clips and images among youth	143	89.9	1.10
Useless information creates ambiguity and confusion in the mind of youth	128	80.5	1.19
Irrelevant and anti-religious post and links create hatred among peoples of different communities	106	66.7	1.33
Social media is becoming a hobby of youth to kill the time	131	82.4	1.18
Mobile phones and computers users have become very addictive	134	84.3	1.16

On the possible perceived effect of using social media, data in table 9, indicates that ($\bar{x}$=1.08) said many young people end up wasting a lot of their time on games and chatting on social networks, ($\bar{x}$=1.10) social media promotes unethical/sexual pictures, video clips and images among youth, ($\bar{x}$=1.16) mobile phones and computers users have become very addictive, ($\bar{x}$=1.18) social media is becoming a hobby of youth to kill the time, ($\bar{x}$=1.19) useless information creates ambiguity and confusion in the mind of youth, ($\bar{x}$=1.3) irrelevant and anti-religious post and links create hatred among peoples of different communities, and ($\bar{x}$=1.44) use of social media is deteriorating our social norms.

Table 9. Problems encountered in using social media

Items	Frequency (n=159)	Percentage (%)
Unwanted friend's request	118	74.2
E-crime, e.g., identity theft, theft of valuable data, interruption of business, financial loss	52	32.7
Unwanted and offensive messages	112	70.4
Sleep deprivation	56	35.6
Defamation of personalities	43	27.0
Unwanted sexual advances, e.g., prostitution, pornography	73	45.9
No privacy	51	31.2
Cyber bullying	40	25.2
Fake identity and profile	84	52.8

With regards to problems encountered in using social media, Table 9 show that 118 (74.2) unwanted friend's request, 112 (70.4%) unwanted and offensive messages, 84 (52.8%) fake identity and profile, 73 (45.9%) unwanted sexual advances, e.g., prostitution, pornography, 56 (35.6%) sleep deprivation, 52 (32.7) e-crime, e.g., identity theft, theft of valuable data, interruption

_____

of business, financial loss, 51 (31.2%) no privacy, 43 (27.0%) defamation of personalities, and 40 (25.2%) cyber bullying.

8. Discussions of findings

The study was anchored on the perceived influence of social media usage among youth in Lagos State, Nigeria. The result indicated that majority of the respondents are single people who falls within the age range identified as millennium generations children that are refers as digital natives (Zickuhr, 2010). By implication, this show that majority of the respondents are people who probably like to spend more time accessing and consuming social media content and digital resources. The findings of study also show that all the respondents confirmed having a social media account, with the majority of them using their cellphone to access it. This may be attributed to the mobility and portability of cellphone. This finding supports the results of the study of Oyewusi and Adamu (2014), who discovered that almost all of their respondents had a mobile phone or a smartphone, which they used for different purposes ranging from voice calls, short text messaging, listening to music and responding to social media messages. The availability and accessibility to mobile phone and internet network service, and cheaper data in the country has increased the widespread usage of social media among youth. Today, access to information via social media is crucial to the youth's everyday lives; as a result, it is no longer a luxury but a necessity serving as a means of accessing and sharing information.

The findings show that the majority of the respondents have Facebook and Whatsapp social media accounts, with Whatsapp being used more frequently than any other type of social media network. This finding is consistent with Markwei and Appiah's (2016) findings, which revealed that Facebook and WhatsApp are the most popular social networking sites among young people. Whatsapp, like Facebook, is quickly becoming one of the most popular social media platforms. According to reports, WhatsApp instant messaging is the most popular messaging program in the world, with 109 nations (or 55.6 percent) using it. It is, without a doubt, the most engaging smartphone app available. Approximately 70 percent of WhatsApp users utilize it on a daily basis (Gadgets360, 2016; Goodwin, 2016). The popularity of WhatsApp usage can be attributed to the fact that it is accessible on any internet-enabled mobile phone, offers instant and unlimited messaging, multimedia, appealing features, video calls, and can be used on a variety of devices and operating systems (IOS, Android, etc.). This outcome also supports the uses and gratification theory underlying presupposition that, media users selectively engage media that possess the features or characteristics that can meet or satisfy their needs at that point in time (Adaja & Ayodele, 2013).

The findings of the study show that 32 percent of the respondents spent about 15-30 minutes per day on social media, and 26 percent spent over 30 minutes -1hour. There is tendency for youth that spent much of their time on social media to start dreadfully becoming addicted; thereby losing concentration on relevant issues pertinent to their lives. The findings revealed that a majority of the youth spend about fifteen to thirty minutes on social media in a typical session. The hours the youth spend on the social media in the study seems to be lower compare to a study by Haruna and Danladi (n.d.). However, this established the popularity of social media among the youth. This concurs with the findings of Onah and Nche (2014) that social media phenomenon is overwhelmingly popular among Nigerian youth. This also agreed with the findings of Akpan, Akwoano and Senam (2013) that the youth visit social media websites daily. Mageto (2017) opined that spending long hours chatting in social networking sites decreases productivity amongst the youth. This in turn causes the youth not to be self-dependent and instead depend on their parents and families for upkeep. The ease with which smartphones are available and accessible may lead to excessive use of social media. It must be noted that social media itself is neutral but multi-tasking interface and interactive applications such as chatting in real time are time-consuming.

_____

The effect of this on the youth could lead to infiltration, flirting, invitation into wrong belief and unwholesome teaching as well as becoming less interested in ethical and moral conduct. According to Nyland and Near (2007), the amount of time that youth spend on these social networks has raised concerns among parents about their children's safety and morals, as they are exposed to many different contents on these sites, such as pornographic content.

From the findings, the young people utilize social media for different purposes. These include for communicating and interacting with friends, online learning, for promoting/doing business, entertainment, general networking, and private messaging, uploading photos and online profiles, and for finding friends online. This finding is in tandem with that of Dhaha and Igale (2013) that youth use social media for virtual companionship escape, interpersonal habitual entertainment, self-description of own country, self-expression, information seeking, and passing time gratifications among others. This suggests that the use of social media reflect the diversity of meeting the information needs and yearning of the youth.

On the perceived benefits derived from using social media, the study also revealed that the youth derived benefits in using social media in this sequence: that it helps in research and learning, strengthening interpersonal relationships, helps in sharing religious activities, encourages virtual meeting with people, supporting collaborative and peer to-peer learning, promote read and write web skills, and it increases self-esteem and well-being. This outcome is in agreement with the findings of Baglari et al. (2020), who reported that youth believed that social media is an effective communication tool that allows users to quickly communicate t with other social media users as well as easily sharing information. On the perceived effect of using social media, the findings of the study revealed that a majority of the respondents of the study agreed that many young people end up wasting a lot of their time on games and chatting on social networks and that social media promotes unethical / sexual pictures, video clips and images among youth. They also complained about social media becoming a hobby of youth to while away time and addiction to mobile phones and computers. The findings is in agreement with the findings of Ambuchi (2015), who reported that the negative effects of the media have impacted the youth as media access and media use (internet) among the youth is rampant and subsequently consumes a significant amount of their time.

In the same vein, it supported the findings of Umeogu and Ojiakor (2014) that internet, particularly social media, had a negative impact on Nigerian youth's education, morality and culture amongst others. Since the advent of the social media in the country, the loss of sense of dignity is on the increase among the youth. It has become an avenue for some youth to share nude or sexual images (Chukwuebuka, 2016). The risks associated with the negative impacts of social media on the younger generation are daily displayed in terms of their changing mental set-up, unhealthy lifestyle and the declining quality of concentration. Concentration and quality time of time seems to be fading away as they barely do without the media on a daily basis. Previously, the addiction to old media (Television and home video) was top-notch among the youth; today it has shifted to addiction to smartphone usage and social media.

Unsolicited friend requests, unwanted and disrespectful messages were identified by the majority of respondents as major problems facing the youth when using social media. This has become a common occurrence in which a fake friend copies the profile information of an already known friend and resends it to impersonate that friend in order to send immoral, frivolous, and seducing photos after the latter accept the request. As a result, the intensity of emotion displayed on social media has the potential to lure young people into the hands of those who wish to constantly thrust their agenda in their faces and wrongly influence their attitudes. Other problem identified by respondents include, unwanted sexual advances, e.g., prostitution, pornography unwanted sexual advances, e.g., prostitution, pornography, sleep deprivation, e-crime, e.g., identity theft, theft of valuable data, interruption of business, defamation of personalities, and cyber bullying. This finding is also in line with the findings of Shu-Sha Angie Guan and

Subrahmanyam (2009) on the risks and opportunities of internet use on the youth. The results show that online risks such as addiction, cyber bullying, and sexual solicitation are associated with negative consequences for youth. Though, many studies have affirmed the positive effect of social media use, the level of the negative influence among the youth have not been adequately reported.

9. Conclusion and recommendations

This study was anchored on the perceived influence of social media usage among youth in Lagos State, Nigeria. The findings of the study indicate that 32.7 percent of the respondents spent about 15-30 minutes a day using social media and (26.4 percent spent over 30 minutes – 1 hour. The youth use social media for different purposes which include communicating and interacting with friends, online learning, for promoting/doing business, entertainment, general networking, and private messaging etc. The study also discovered that the youth derived benefits in using social media in the area of research and learning, strengthening interpersonal relationships, sharing religious activities, encourages virtual meeting with people, supporting collaborative and peer to-peer learning. The negative influence of the media on youth, on the other hand, is causing a lot of time wasting on games and chatting, and that social media promotes unethical / sexual pictures, video clips, and photos. This study concluded that the current findings may aid in a better understanding of the influence of social media use on youth and, hence, inform strategies for balancing control and participation, as well as reducing its negative effect. Based on the findings of the study, it is recommended that, since wasting of time is one of main identified negative effect of social media use among youth, parents and guidance need to indirectly monitor and controlled activities of the children while they are on internet-enabled mobile phone. The opportunities of social media use are endless, but young people should use it more for academic and entrepreneurial purposes than for social interactions. Finally, a nationwide youth orientation and campaign on the ethical use of social media is needed.

This study is not without some limitations. The study design used was survey, hence, difficult to establish cause-effect relationships between the variables of the study. The sample size used for the analysis was relatively small as many of the respondents did not return the questionnaire, while many did not complete them as expected despite the efforts of the researchers and the research assistants. Therefore, generalizing the findings should be done with cautions. Nevertheless, these limitations, this study has several implications. The findings of this study could be used to develop public enlightenment programs aimed at educating the youth on effective ways of using social media. Finally, researchers, parents, guidance and counsellors could find this study findings significant and use the study as references resources in other related studies such as technology and media use.

References

Adaja, A. T., & Ayodele, F. A. (2013). Nigerian youths and social media: Harnessing the potentials for academic excellence. *Kuwait Chapter of Arabian Journal of Business and Management Review*, *2*(5), 65-75.

Adelabu, O., Sanusi, B. O., & Esiri, M. (2015). Smartsphones and communication pattern among students in higher institutions. *Singaporean Journal of Business Economics, and Management Studies*, *3*(12), 56-63.

Akpan, U., Akwoano, E., & Senam, S. (2013). Use of gratification of social networking websites among youth in Uyo, Nigeria. *International Journal of Asian Social Sciences*, *3*(2), 353-369.

Almenayes, J. J. (2014). Religiosity and the perceived consequences of social media usage in a Muslim Country. *Journal of Arts and Humanities*, *3*(5), 108-117.

Ambuchi, A. (2015). The impact of media on the morality of young Adventist church people in the Nairobi Central Church. *Project Documents*, *258*, 1-137.

Amofah-Serwaa, N., & Dadzie, P. S. (2015). Social media use and its implications on child behaviour: a study of a basic school in Ghana. *Int. J. Social Media and Interactive Learning Environments*, *3*(1), 49-62.

Annapoorna, S., Reshma, R., & Sawad, H. (2015). The impact of social media on youth. *International Journal of Innovative Research in Computer and Communication Engineering*, *3*(7), 379-383.

Baglari, H., Sharma M, K., Marimuthu, P., & Suma, N. (2020). Pattern of social media use among youth: implication for social media literacy. *Ment Health Addict Res 5*, 1-5. https://doi.org/10.15761/MHAR.1000189

Chukwuebuka, U. F. (2013). The effect of social media on youth development. *Nigeria Village Square*. Retrieved from http://nigeriavillagesquare.com/forum/threads/the-effect-of-   social-media-on-youth-development.77876/.

Dhaha, I. S. Y., & Igale, A. B. (2013). Facebook usage among Somali youth: a test of uses and gratifications approach. *International Journal of Humanities and Social Science*, *3*(3), 299-313.

Dimitrov, B. G. E., & Nongkynrih, A. (2017). A study on social media and its impact on youth. I*nternational Journal of Creative Research Thoughts*, *025*, 149-158.

Dimka, P. (2018). The effect of social networking/ media on Christian youths. Retrieved from https://realphidims.wordpress.com/

Edeh, M. O. (2018). *The impact of social media on faith and Christianity*. A Paper Delivered at the 13th Annual National Conference of Catholic Law Students Association of Nigeria, Held at Enugu State University of Science and Technology on 18-24 September, 2018.

Edogor, I. O., Jonah, A. A., & Ojo, L. I. (2014). Influence of social media on youths' usage of traditional mass media in Nigeria. *New Media and Mass Communication*, *3*, 55-66.

Ezumah, B. A. (2013). College students' use of social media: site preferences, uses and gratifications theory revisited. *Int J Bus Soc Sci*, *4*, 27-34.

Fallahi, V. (2011). Effects of ICT on the youth: A study about the relationship between internet usage and social isolation among Iranian students. *Procedia Social and Behavioral Sciences*, *15*, 394-398.

Gadgets360 (2016). *WhatsApp in over 109 countries, with 70 million users in India: Study*. Retrieved from http://gadgets.ndtv.com/apps/news/whatsapp-in-over-109-countries-with- 70-million-users-in-india-study-841858.

_____

Goodwin, R. (2016). *Whatsapp vs. Facebook Messenger: One billion users logged in 2016. Know your mobile*. Retrieved from http://www.knowyourmobile.com/mobilephones/whatsapp/21430/whatsapp-vs-facebook-messenger-features-differences-compared-which-is-best.

Guan, S. A., & Subrahmanyam, K. (2009). Youth internet use: risks and opportunities. *Current Opinion in Psychiatry*, *22*, 351-356.

Haruna, A., & Danladi, K (n. d.). Social media, youths and culture change among ABU Zaria undergraduates. *Communication and the New Media in Nigeria*, 176-188.

Israel, G. D. (1992). Determining sample size. Retrieved from https://www.tarleton.edu/academicassessment/documents/samplesize.pdf.

Katz, E., Blumler, J.G., & Gurevitch, M. (1974). Uses and gratifications research. *The Public Opinion Quarterly*, *37*(4). 509-523. Retrieved from https://www.jstor.org/stable/pdf/2747854.pdf.

Mageto, J. (2017). Impact of social media on the youth. Retrieved from https://www.grin.com/document/358350.

Markwei, E. D., & Appiah, D. (2016). The impact of social media on Ghanaian youth: A case study of the Nima and Maamobi communities in Accra, Ghana. *The Journal of Research on Libraries and Young Adults*, *7*(2), 1-26.

Matei, S. A. (2010). What can uses and gratifications theory tell us about social media. Retrieved from https://matei.org/ithink/2010/07/29/what-can-uses-and-gratifications-theory-tell-us-about-social-media/.

Nkordeh, N., Oni, O., Olowononi, F., & Bob-Manuel, I. (2017*). Impact of social media on the reading culture of Nigerian youths; Is the social media platform providing an alternative source of education?* Proceedings of the World Congress on Engineering and Computer Science I WCECS 2017, October 25-27, 2017, San Francisco, USA.

Nyland, B., & Near, N. (2007). *Jesus is my friend: Religiosity as a mediating factor in internet social networking use*. Paper presented at the AEJMC Midwinter Conference, Feb 23-24, 2007 in Reno, Nevada.

Ogunkola, O. (2018). Influence of social media on Nigerian youths. *The Nation Newspaper*. Retrieved from: https://thenationonlineng.net/influence-social-media-nigerian-youths/

Onah, N. G., & Nche, G. C. (2014). The moral implication of social media phenomenon in Nigeria. *Mediterranean Journal of Social Science*, *5*(20), 2231-2237.

Ononogbu, D. C., & Chiroma, N. (2018). Social media and youth ministry in Nigeria: Implications and Christological thrust. *International Journal of Humanities Social Sciences and Education*, *5*(1), 48-54. http://dx.doi.org/10.20431/2349-0381.0501008

Oyewusi, L. M., & Adamu, B. J. (2014). The chemistry of mobile phones: A research report on the extent of usage of the compact technology among students on Nigerian campuses. *World Journal of Education*, *4*(4), 38-42.

Shabir, G., Hameed, Y. M. Y., & Safdar, G. et al. (2014). The impact of social media on youth: A case study of Bahawalpur City. *Asian Journal of Social Sciences & Humanities*, *3*(4), 132-151.

Tarimo, F. (2016). The analysis of the social media among the youth in the Catholic diocese of Moshi and the implications for youth faith sharing. *Masters Theses*. Paper 15.

Umeogu, B., & Ojiakor, I. (2014). The internet communication and the moral degradation of the Nigerian Youth. *International Journal of Computer and Information Technology*, *3*(2), 450-463.

Zickuhr, K. (2010). Generations online in 2010. *Pew Research Center Internet & Technology*. Retrieved from http://pewinternet.org/Reports/2010/Generations-2010.aspx.

_____

Enhancing Asset Security by Integrating Internet of Things on Non-Powered Assets

Joshua Mueke Mwema, John M. Kandiri & Stephen Titus Waithaka

Kenyatta University, Nairobi, KENYA
Department of Computing and Information Technology

Abstract

Rapid proliferation of the Internet of things (IoT) has helped solve a myriad of problems across different sectors. Whilst powered assets rely on their own power source to enhance asset monitoring, a need exists to develop an asset security system for non-powered assets. Since an IoT device has network layers, it can be used with the physical layer to solve this problem. As a result, a method of enhancing asset security with IoT was implemented. A GSM chip for SMS and server connection was used to communicate the battery status and location information obtained from a GPS chip. A method of detecting tamper on the device was implemented through infrared sensors. A microcontroller was the heart of the system as it interfaced with all other devices to form the IoT system. The result of the study was an efficient prototype of an IoT asset tracking device which communicates through SMS and it logs location data to a remote server through GPRS connection. There is a room for improvement in terms of optimizing power consumption to lengthen the duration before a recharge is required.

Keywords: internet, asset security, Internet of things, non-powered asset.

1. Introduction

The internet of things (IoT) technology comprises networked devices which utilize mainly the physical layer to collect data and transmit it to a destination through the network layer (Minhaj & Khaled, 2018). As such, there is a lot of potential of automating systems through IoT as it provides actionable information. One of the vibrant applications of IoT is in asset security and monitoring where many studies have been done on development of smart security systems targeting an asset's location and condition (Valente & Neto, 2017). Utilization of IoT in data analytics has yielded predictive maintenance. For instance, OracleIoT (2020) described a system of preventing unforeseen machine breakdowns by processing data received from IoT systems with advanced analytical algorithms. In the area of asset security, IoT systems require communication mechanisms to alert the user about the asset. According to Indira et al. (2019) and Andreas et al. (2018), global system for mobile communication (GSM), Wi-Fi, and remote servers are some of the main methods used alongside IoT to achieve seamless data transmission.

Asset security in vehicles and laptops may be achieved through powering an IoT system with onboard batteries (Nirit et al., 2019). On the other hand, not all assets have an onboard power supply. At the same time, IoT systems must be powered electrically to perform embedded tasks (Andrey & Raffaele, 2019). Thus, non-powered assets such a sofa sets and TV sets

Correspondence: Joshua Mueke Mwema, Kenyatta University, Department of Computing and Information Technology. E-mail: joshuamke@gmail.com.

may be lost without noticing as Smartrak (2020) notes that there is still a challenge of securing these kinds of assets. The present study aimed at developing an asset security device that would not only report location change and upload location data to a remote server but also notify the user when it required battery recharging. Further, the system would report if it was tampered by opening its casing.

2. Methods and materials

The functional features of the IoT asset security system were achieved through design and implementation of different subsystems which were brought together reaching the proposed system. The main subsystems as shown in the figure below are; GMS module, GPS module, antitamper circuit, voltage regulation module, a logic level converter and the microcontroller. The code that controls the system was developed to function as per the flow chart. The program was written in C++ language using the arduino integrated development environment (IDE). Moreover, the circuit diagram was drawn using Proteus design suite. On the other hand, ThingsSpeak open-source server was used to create a sever account where data about the location was logged remotely using API keys method. Prototyping tools such as soldering gun, solder wire, prototyping circuit board and jumper wires were utilized in assembling the circuit.

The security system works with help of sensors, modules and a code embedded in a microcontroller. The microcontroller first checks the battery voltage followed by the status of the device casing then location change. If the battery voltage is lower than a threshold, it sends a battery low SMS through a GSM module. The GSM module communicates with the microcontroller through AT commands and hence the serial protocol is used. At the same time, the MCU checks whether the device casing is open by reading the output of the infrared proximity sensor where it sends an SMS to the owner if the case is open. Lastly, the MCU reads the GPS module to retrieve location coordinates where it sends them to a remote server through a GPRS connection. If there is a significant change in the pre-set location, an SMS is sent to the owner about the current location. If none of the conditions mentioned here are met, there no SMS triggering. Communication between GSM, GPS and the MCU is through universal asynchronous and receiver and transmitter (UART) protocol. During serial communication, high logic signals emanate from the transmitting pin of the device, a reason why a logic level converter is used to protect non-5V ports from damage by the high voltage signals. As GSM chips draw more current during transmission, i.e., transmission bursts, they can create short circuits hence in the circuit, it is connected via a dc-dc converter that can withstand current sourcing of up to 3A without need of a heat sink. Moreover, the voltage circuit is built around a voltage divider circuit to achieve a voltage less than 5v that can be fed into the analog pin of the MCU for measurement.
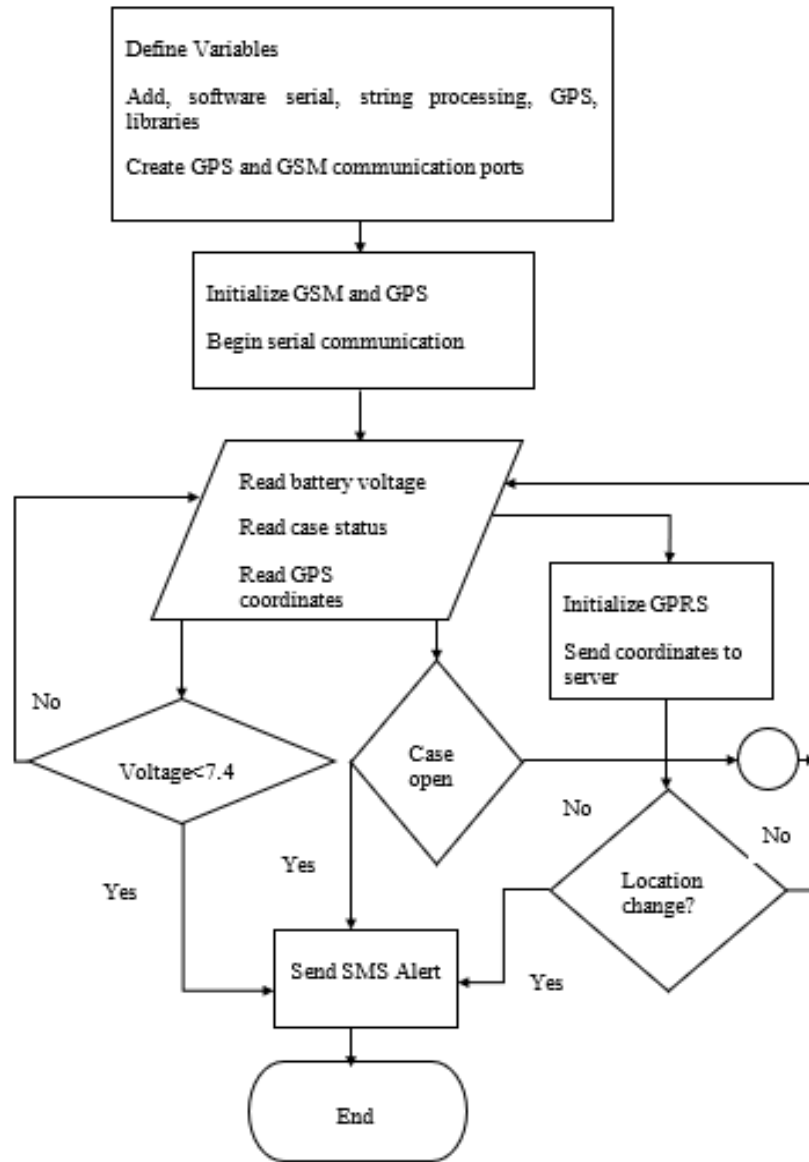
_____

Figure 2. Flow chart diagram

3. Results

3.1 *Voltage measurement*

Using two resistors R1 and R2 of 3.3k and 1k respectively, a maximum output voltage of the sensor was achieved as 1.95V. This was a safe voltage as it cannot damage the ADC channel of the arduino. A conversion factor of 4.3 which is the total resistance was used to compute the actual voltage from the reduced voltage. As the device was in use, the battery voltages were recorded during test instances as; 7.07V, 6.97V and 6.89V.
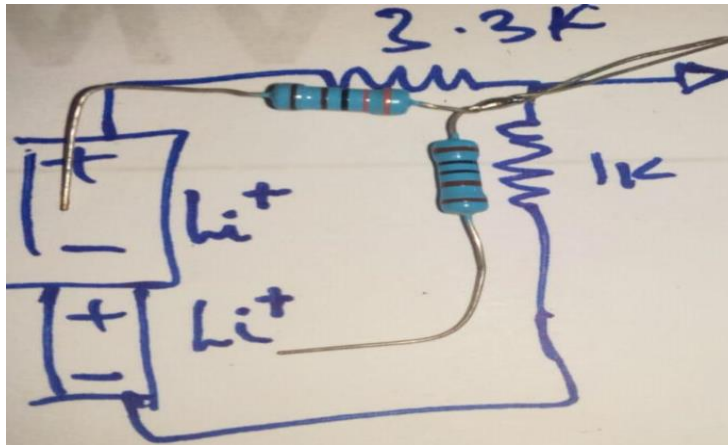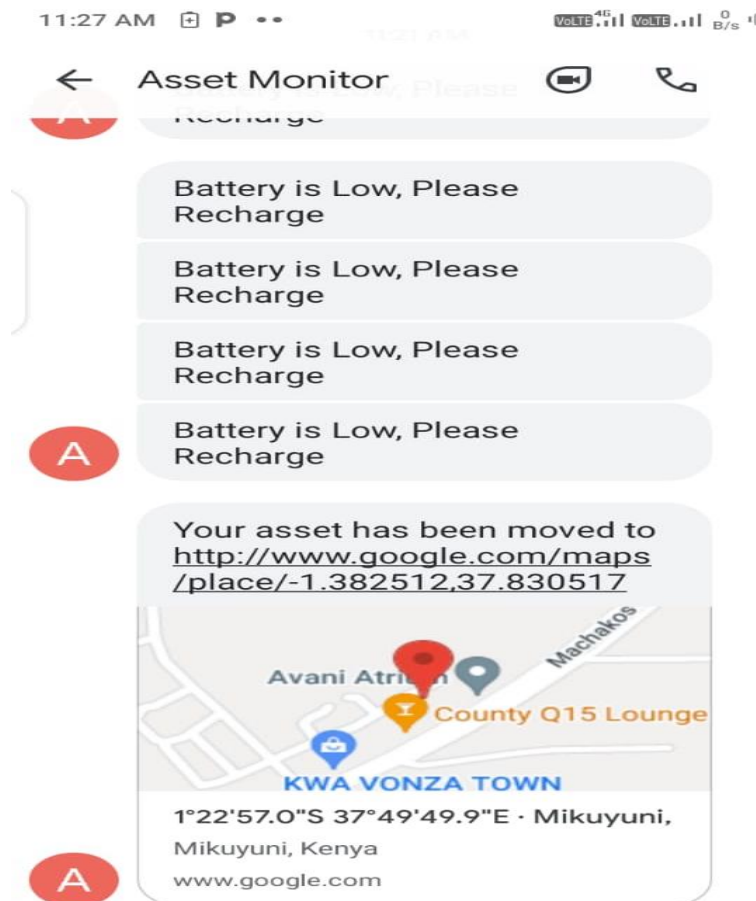
Figure 3. Voltage measurement circuit

3.2 *Location and low battery alerts*

As shown below, a battery low alert message was received when the battery was lower than the set threshold. Moreover, an SMS showing the location link was sent where the user could just click the link and view the asset location.

_____

3.3 *Server data*

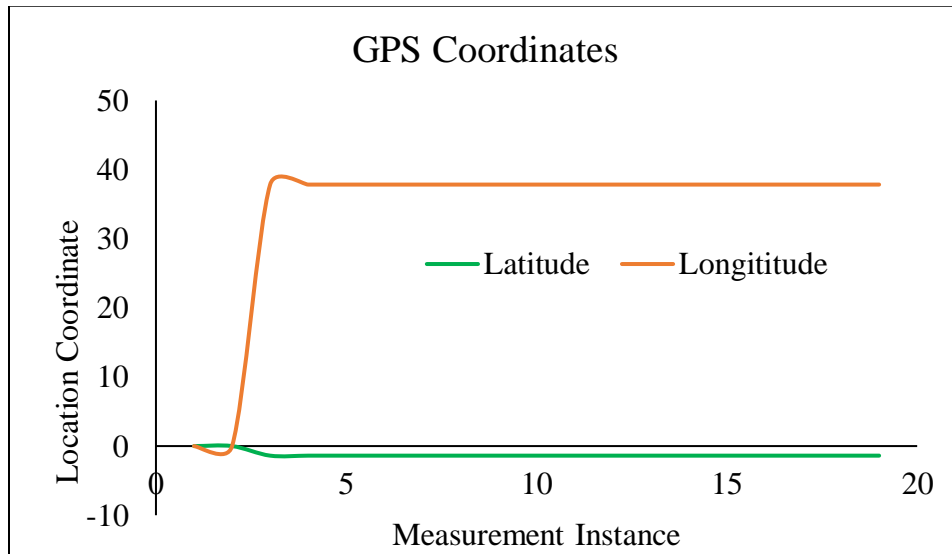The location coordinates retrieved from the server are shown in the plot below.

Figure 4. Location data retrieved from server

Figure 4 above was plotted using raw data imported from the server. However, the server has visualization features which were utilized to plot latitude and longitude data on separate graphs in real-time. The two visuals are shown below.
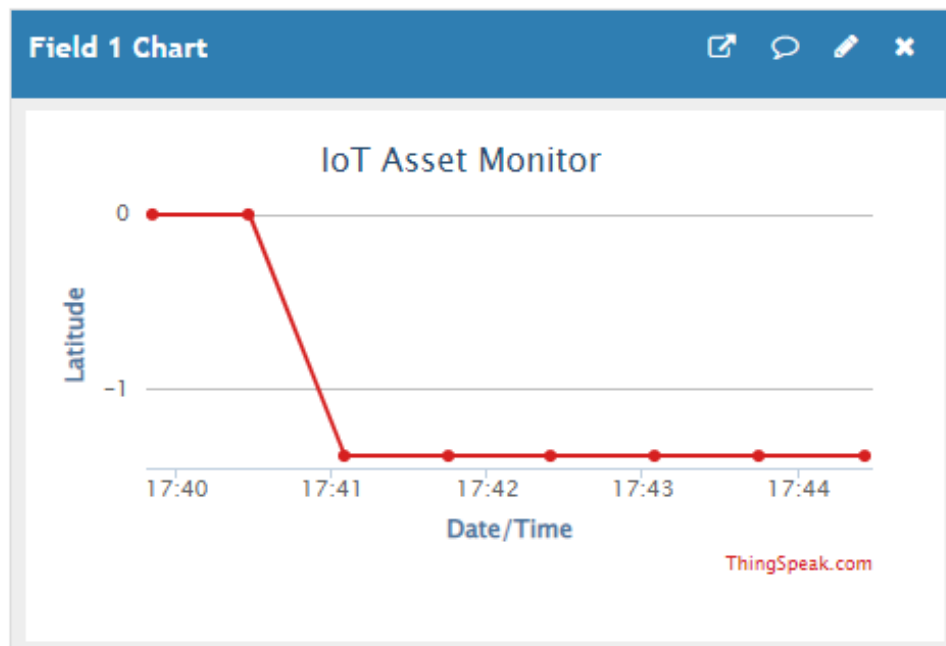

Figure 1. Latitude data

Initially, the latitude data was zero but it progressed to -1.3xxxxxx after a short time as shown in figure 4 above.
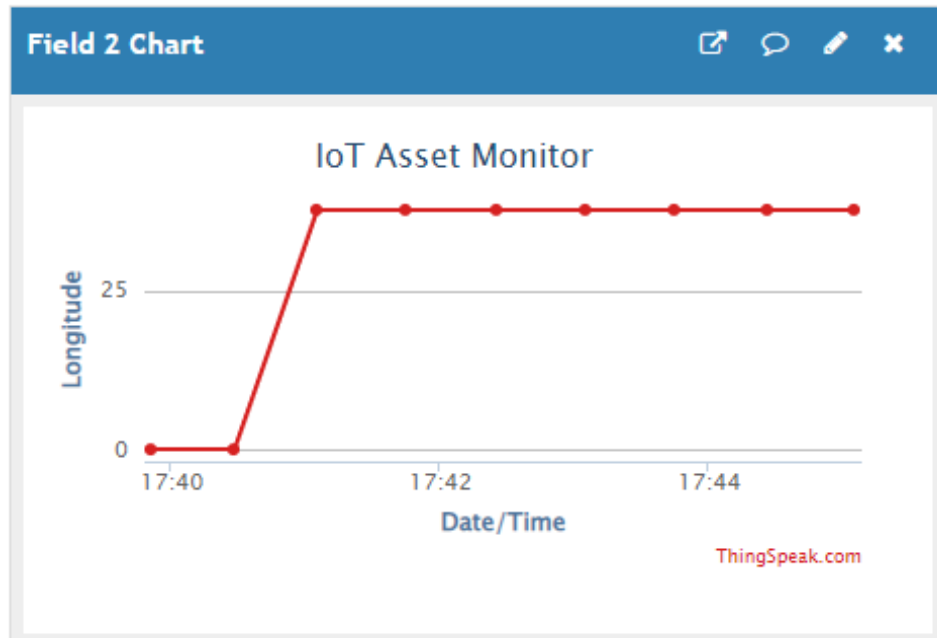
Figure 6. Longitude data

As noted in figures 4 and 5 above, there are instances where the location coordinates were both zero. After sometime, it was noticed that the value changed to several digits of different decimal values. A snip shot of serial output is shown below. The voltage is seen to decrease as GPRS transmission occurs. Data for each coordinate was logged to a separate channel using API keys shown in the serial monitor output.

The figure below shows the complete prototype with all parts labelled.

Switch

Battery

GSM

dc-dc
converter

gfgdfgc

NEO6 GPS

Voltage divider

MCU

IR sensor

Logic level converter
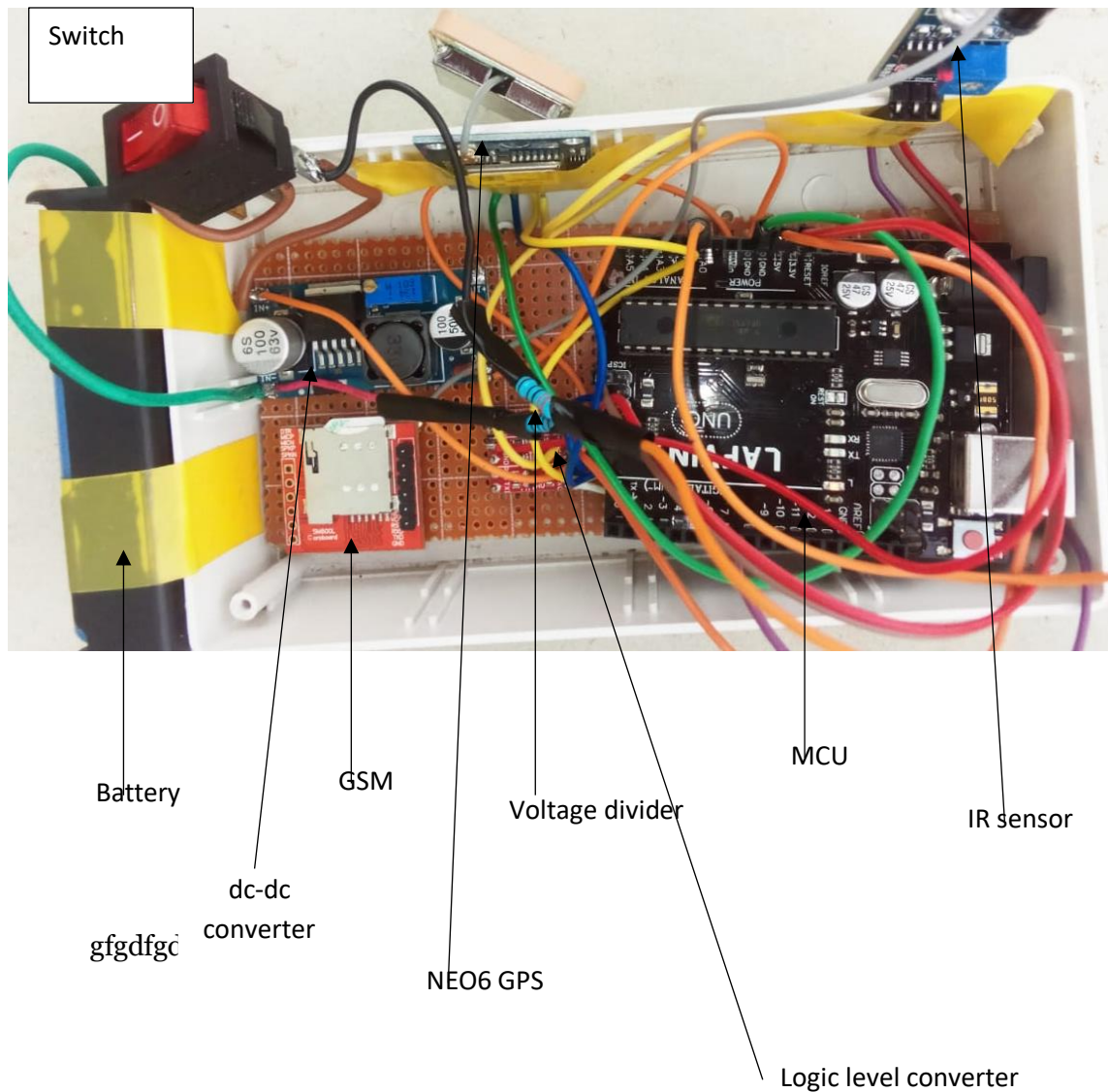
Figure 7. Completed prototype

The plastic casing was used to hold the assembled prototype together. Furthermore, some components were soldered on a PCB to reduce the number of wires required to run in the circuit. To ensure the GPS and GSM received the signals, the antennas were kept outside the plastic container.

4. Discussion

The battery voltage sensor recorded values that were less than 7.4 V as noted earlier. Since two 4.2V lithium batteries were connected in series, the maximum actual voltage would be 8.4V while the input to the ADC pin was 1.95 at full charge given the values of the resistances were in the ratio of 1:3.3. Thus, during testing the batteries had discharged hence they had lower voltages. To prevent deep discharge, an alert was sent when the battery voltage fell below 7.4 volts. This implies that each cell may have reached 3.7V which is the nominal voltage of Lithium-ion cells. Such voltage drop can be attributed by power consumption by the IoT system. More

specifically, the GSM draws a lot of current during data transmission which reduced the voltage significantly. This is in agreement with arguments of Yassin et al. (2020) that IoT systems require optimization both in hardware and software to reduce power consumption for asset tracking systems. Regarding the graphs of location coordinates, the value of zero meant that the GPS module had not connected to any satellites. This mostly occurred when the system was placed indoors where the GPS signal was obstructed by the walls of the building. Slight variations in the location coordinates even when the device was stationary implied that location determination using GPS has a certain degree of uncertainty. Findings of Meynecke and Liebsch (2021) agree with this by the fact that localization using GPS requires clear space between the GPS module and satellite signals. Although indoor localization using GPS and GSM may not be effective due to poor network coverage, the devices work well in outdoor conditions. As an asset is exposed to outdoor conditions during relocation, the GPS module will connect with satellites where the MCU will trigger an alert about location change. Therefore, the current device has advantage because of long range data transmission and remote access through server as GPS has a comparable location accuracy. The accuracy of 3m obtained in outdoor location is comparable to the 2.5 m accuracy achieved by Devlin and Kevin (2009). Overall, the device exhibited timely data transmission both to server and through SMS and is more ideal for monitoring non-powered assets that are located in outdoor spaces where signal reception is strong.

5. Conclusions

An IoT security system relying on GPS for asset location and GSM for SMS and server communication has been designed, implemented and tested. The system boasts an onboard method of measuring battery voltage, detecting when the casing has been opened and reporting the status of the system through GSM. The GSM module consumes more energy compared to the rest of the devices hence its usage can be leveraged by optimizing the embedded code. The device is found more ideal for monitoring non-powered assets that are kept in outdoor spaces given the conventional reliance of GPS on clear space for signal reception. This study concludes that IoT has the potential to change the asset security landscape by developing stand-alone devices that can monitor and report non-powered assets' location and status.

Acknowledgements

References

Andreas, P., Kostas, P., Christos, S., Haoxiang, W., & Gupta (2018). Efficient IoT-based sensor BIG data collection–processing and analysis in smart buildings. *Future Generation Computer Systems*, *82*, 349-357.

Andrey, S., & Raffaele, G. (2019, November 9). *Powering IoT devices: Technologies and opportunities*. IoT IEEE.

Devlin, G. J., & McDonnell, K. (2009). Performance accuracy of real-time GPS asset tracking systems for timber haulage trucks travelling on both internal forest road and public road networks.

_____

International Journal of Forest Engineering, *20*(1), 45-49. https://doi.org/10.1080/14942119.2009.10702575

Indira, R., Bhavya, G., Dheva, D. S., & Devaraj, R. (2019). IOT asset tracking system. *International Journal of Computer Science and Engineering*, ICFTESH, 45-50.

Meynecke, J.-O., Liebsch, N. (2021). Asset tracking whales — First deployment of a custom-made GPS/GSM suction cup tag on migrating humpback whales. *J. Mar. Sci. Eng.*, *9*, 597. https://doi.org/10.3390/jmse9060597

Minhaj, A. K., & Khaled, S. (2018). IoT security: Review, blockchain solutions, and open challenges. *Future Generation Computer Systems*, *82*, 395-411.

Nirit, D., Ashutosh, M., Mukund, A., & Anirudh, J. (2019). Real time tracking and alert system for laptop through implementation of GPS, GSM, motion sensor and cloud services for antitheft purposes. *IEEE Xplore*, *4*(19), 978-1-7281-1253.

Smartrak (2020). Smartrak. (Smartrak) Retrieved March 21, 2021, from https://smartrak.com/non-powered-asset-tracking/

Valente, F., & Neto, A. C. (2017). Intelligent steel inventory tracking with IoT / RFID. *2017 IEEE International Conference on RFID Technology & Application* (RFID-TA). Warsaw, Poland.

Yassin-Kassab, H., Rosu, M-C. (2020). An overview of own tracking wireless sensors with GSM-GPS features. *Advances in Technology Innovation*, *6*(1), 47-66.

J. M. Mwema, J. M. Kandiri & S. T. Waithaka – Enhancing Asset Security by Integrating Internet of ...

64

_____

A Hybrid Model for Text Summarization
Using Natural Language Processing

James Mugi Karanja & Abraham Matheka

Kenyatta University, School of Engineering, Nairobi, KENYA

Abstract

Text summarization plays an important role in the area of natural language processing. The need for information all over the world to solve specific problems keeps on increasing daily. This poses a greater challenge as data stored on the internet has gradually increased exponentially over time. Finding out the relevant data and manually summarizing it in a short time is a challenging and tedious task for a human being. Text Summarization aims to compress the source text into a more concise form while preserving its overall meaning. Two major categories of text summarization methods exist namely: extractive and abstractive. The extractive technique concentrates on determining key themes using frequency analysis of sentences in the corpus of the text. Abstractive methods write a new summary with newly generated texts which do not appear in the corpus itself. This paper presents a hybrid model for text summarization using both extractive and abstractive techniques. Term Frequency (TF) – Inverse Document Frequency (IDF) was used for extractive and T5 Transformers for abstractive summarization. Iterative Incremental Methodology was adopted in the study. The hybrid model emerged as not the best choice compared to the extractive and abstractive as it had been hypothesized in the study when the accuracy and execution time of the summary generated was considered.

keywords: extractive model, abstractive model, hybrid model, natural language processing.

1. Introduction

The volume of data online continues to grow exponentially thus pausing a great challenge on how to extract relevant information from the massive amount of data (Nguyen, Santos & Russell, 2019). The technique of text summarization has been very effective in information summarization and retrieval thus helping in time-saving while searching for relevant and critical information. This in turn helps in quick decision-making (Hassel, 2007). Text summarization is the process of using software to reduce the length of a text document to make a summary having important considerations from the original document. This is done by highlighting the most important parts of the text (Goyal, Behera & McGinnity, 2018). The input type determines the type of summarizer to be used. You can have a single document summarizer where the input is a small amount of text content or a multi-document summarizer where the input can be derived from various sources and long documents (Goyal et al., 2018). The complexity of the model to be created increases as the amount of text to be summarized rises.

Correspondence: James Mugi Karanja, Kenyatta University, School of Engineering, Nairobi, KENYA. E-mail: karanja.mugi@ku.ac.ke.

Generic summarizers treat the input without bias or prior knowledge; Domain-specific summarizers employ domain information to generate a more accurate summary based on known facts, and Query-based summarizers only contain known responses to natural language questions regarding the input text. In Query-Based Text Summarization (QBTS), a query is taken as input in this method, and depending on the query, the model can create the text's summary by choosing phrases and sentences which are closely related to the query posed as the input (Boorugu & Ramesh, 2020). The length of the input in Single Document Text Summarization (SDTS) is usually short while the length of the input in Multi-Document Text Summarization (MDTS) is usually longer and numerous documents are offered as input for summary creation. MDTS is usually more difficult than SDTS because you have to integrate multiple documents' summaries into a single document (Boorugu & Ramesh, 2020). Extractive methods create a summary from the source material by extracting important phrases, keywords, and sentences. The abstractive approaches provide a summary that resembles a human being's written abstract. The extractive method assures that the created summaries are grammatically and semantically correct, while the abstractive method generates more diverse and new content (Yao et al., 2018).

2. Problem statement

The need for information all over the world to solve specific problems keeps on increasing daily. This poses a greater challenge as data stored on the internet has gradually increased exponentially over time (Nguyen et al., 2019). The structure of how processes used to be executed in various industries has also changed due to the onset of the COVID-19 epidemic. In the education sector, it has led to the adoption of E-learning in most Institutions of Higher Learning as students are encouraged to get access to research and learning materials online instead of visiting physical libraries. With the vast amount of data online, extracting meaningful information might not be easier (Goyal et al., 2018). This might lead to a lot of time utilized during research and learning as students go through various learning materials online. This calls for a solution that can help transform the huge amount of data into summarized information that is easily consumable. Research previously conducted concentrated on the summarization of only plain text. This calls for the development of a hybrid model for text summarization that will be able to extract and summarize text from various sources, i.e., pdf, plain text, and website.

3. Literature review

This chapter analyzes related works, techniques for text summarization, and various models in use.

4. Text summarization

The process of creating a summary necessitates first reading and comprehending the original content. The main components of the paper are stated based on the known events, facts, or situations to satisfy the summary's objective. The summary would not include all of the material in the original text, but only the parts that were judged relevant. This is self-evident, given that the summary's objective is to decrease the quantity of information in the source papers. The summary is subsequently prepared in a suitable output format after defining the main aspects of a document. In general, there are three components to summarization: input, analysis, and output. Humans often require a comprehension of the native language in which the text material is written. The goal of the summary and the target audience would both need to be determined throughout the analysis. The final step before the summary is delivered to the user would be to create a proper output format for the summary (Radev & Erkan, 2015).

_____

There are several things to consider regardless of whether the summarizer is a person or a machine. In terms of input: The summarizer would have to select how to approach reading the material based on how it is structured. For example, information relevant to the analysis stage may be found in the headers of chapters or the labels of Figures and Tables. Some document metadata, such as keywords in HTML websites, may also be useful. If the document was classed and the summarizer has access to the class or domain to which it belongs, it may be feasible to use domain-specific information to help in the analysis and output phases. The summarizer may be influenced by the language used in the text documents. Human summarizers normally need to know the language in which the material was written (Yao et al., 2018).

5. Related Works

Every day, massive volume of information is published on the internet (Boorugu & Ramesh, 2020). This necessitates the use of a solution that can assist in the transformation of large amounts of data into summarized information. This necessitates the application of a text summarizing technique, which aids in the reduction of massive material into key points. The method works by keeping important information while producing a condensed version of the text.

Automatic summarization (Chu, Song & Jaimes, 2015) or indexing (Garg, Hassan & Chaudhury, 2015) has been investigated in a variety of study disciplines, including video, audio, and document (text) processing. Text summarizing techniques were employed to construct concise summaries of documents. A frequency thresholding method was developed in an early text summarizing study. The study used term frequency, inverse document frequency (TF-IDF) method (Shimada, Okubo, Yin & Ogata, 2018), which has been found to attain a reasonable degree of performance, has been introduced to improve frequency-driven approaches. TF-IDF concentrates only on extractive text summarization where it uses word frequency to generate a summary. Even though some words might have a higher frequency, they might not be the main points thus leading to a misleading summary.

Eberts et al. (2015) suggested a system for automatically condensing educational video content. Their method pinpoints the exact moment and location in video footage where presentation slides appear. In their method, they combine image processing and machine learning approaches. They also created electronic lectures and screencasts with the technology. The findings suggested that the summaries created gave viewers more information. The research concentrates on video content and fails to address summarizing text content which is usually used in most institutions.

A method for summarizing oral lectures was proposed by Chen et al. (2011). On a graph created, a random walk is performed making use of automatically extracted key phrases and latent semantic analysis with probabilities in their method. They used their method to obtain each document summary from lecture documents using the extractive summary generation method (Chen et al., 2011). The researcher failed to elaborate more on how to extract the content from various sources to generate the summary. Li et al. (2014) suggested a completely automatic approach for capturing the full presentation utilizing camera techniques such as panning, tilting, and zooming to extract the semantic structure of a typical academic lecture video (*Ibid.*, 2014). General video summarizing approaches were shown to obtain more precise display structures than their system. This concentrates on video content summary but doesn`t talk of text content summary generation which is easily accessible to a large population.

It is sometimes suggested that studying ahead of time for a class is critical for students to familiarize themselves with keywords from the study, and learn new concepts and terminology. Shimada et al. (2018) stressed the necessity of giving students a glimpse of what they will study ahead of time. Furthermore, effective preparation before lectures start has been linked to

improved understanding and grasping of the concepts taught during the lecture. Students are frequently requested to study a textbook or preview content to prepare for their next class at university (Shimada et al., 2018). There is need of having a way of creating summarized content that can be used in study previews instead of going through a whole topic in a textbook which will lead to time-saving among the learners.

6. Hybrid summarization

Combining more than one feature of text summarization leads to the development of a hybrid model. Rani et al. (2017) proposed a hybrid model which makes use of the word frequency and the location of the paragraphs. Words contained in paragraphs at the beginning of the document are given more weight than words in the preceding paragraphs. Even though the model developed is hybrid, it only makes use of extractive technique of text content summarization.

MuraliKrishna et al. (2013) offer a hybrid summarizing system in which sentences are extracted from documents based on the sentence scoring method. The average of the values evaluated using statistical and linguistic methodologies is used to calculate the sentence-scoring method. The duplicate information in these retrieved sentences is handled using an iterative clustering process. The final result, known as the document summary, provides the most significant sentences in the document related to the query without redundancy. The generated sentences can be sorted by their sentence score or the order in which they appeared in the source. The researcher makes use of extractive method of summary generation and doesn't mention abstractive technique.

A hierarchical hybrid multi-document model using extractive technique was invented by (Celikyilmaz & Hakkani-Tur, 2010). The model is designed with the capability of splitting a document into major subtopics which in turn are divided into more small topics. The content in the subtopics is summarized independently via extractive technique and the results are combined to form a final summary. The outcome of this system would be more fine-tuned if they would have been channeled to a model using abstractive technique. Even though the research developed a hybrid model, it only made use of one major technique of content summarization.

A multiple-text document system that uses hybrid summarization techniques was developed by (Dave & Jaswal, 2016). The system uses extractive techniques in the initial stage and the output is channeled to an abstractive model which uses Word Graph generation which locates important nodes from the extractive summary using heuristic rules. Heuristics rules aid in easier problem-solving and provide a shortcut to solving difficult problems but don't necessarily give an optimum solution. This might affect the output of the summary. The system only tackles the summarization of text content pasted on it and doesn't provide the allowance of summarization of online text content, i.e., from Wikipedia. Instead of using Word Graph, the research proposes the use of T5 transformers for the abstractive summary generation to increase the accuracy of the output summary and also provide the allowance of getting the text content to be summarized from various sources, i.e., extracted from pdf documents and website URL (Uniform Resource Locators) like Wikipedia.

7. Text summarization models

7.1 *Term Frequency (TF) – Inverse Document Frequency (IDF)*

The TF-IDF algorithm was utilized in the extractive summary generation. Term Frequency (TF) helps to count the frequency of words. The frequency discovered is used to assess the word's significance. The more frequently a word appears in the source document, the more important it is. The straightforward explanation for TF is that it counts the frequency with which

a word emerges or is seen in a document (Meena et al., 2020). IDF provides unique words with a higher value and repeated words with a lower value. TF occasionally overestimates the significance of stop words depending on how often they appear. Inverse Document Frequency determines the rarest of words that appear in the document to address TF's issue. IDF is the inverse of TF; when the two are combined, the result is TF-IDF refers to the product of TF and IDF. TF−IDF is a powerful algorithm for summary generation though it is only applicable for extractive summary generation and can't rewrite the summary using different words other than the one in the main document (Meena et al., 2020).

8. Text rank algorithm

The text rank is an unsupervised method that uses weights as a value to rate sentences. The foundation of the text rank algorithm may be traced back to page rank on Google system, which performs the ranking of websites regarding their links and their significance (Mihalcea, 2004). As the name implies, a directed graph is built using phrases. This is referred to as the graph-based ranking method. The phrases are referred to as nodes or vertices, and edges are used to connect nodes that are related (Mallick et al., 2019). The text rank algorithm is a recommender-based system in which the vertices joined by the edges recommend the relevance of the phrases in the graph. The weights assigned to the sentences are used to rank the sentences and the summary is generated from the sentences having more weight. The text Rank Algorithm employs only the extractive method of summary creation. A system employing both extractive and abstractive summary generation models can be of greater help in gauging the validity of the summary created.

9. Sequence to sequence model

Time-Frequency Representation Summary (TFRS) method employs an abstract summary model known as a sequence-to-sequence model to improvise new words while maintaining the meaning of the original text. It uses an encoder-decoder model which translates sequences of varying lengths as input and output (Song, Huang & Ruan, 2019). The encoder-decoder component's Long Short-Term Memory (LSTM) is useful for capturing long-term dependencies. The training and inference phases of the encoder-decoder model are separated. In the training and inference phases, both the encoder and the decoder are utilized. Sequence-to-sequence models have provided feasible solutions for abstractive summarization but are still hard to tackle long text dependency in the summarization task (Liao, Zhang, Chen & Zhou, 2020).

10. Transformers for text summarization

Transformer is an encoder-decoder model and converts all NLP problems into a text format. They make use of Transfer machine learning where pre-trained models are used to perform different tasks thus providing higher performance (Gupta, Chugh, Anjum & Katarya, 2021).

Due to the large amount of data used to train the transformers; they usually create well optimized models thus making them to be most preferred.

A research conducted by Gupta et al. (2021) on various types of transformers for text summarizations portrayed that T5 transformer outperformed the other types of transformers as shown in the table below.

Models	Evaluation Metrics		
	ROUGE-1	*ROUGE-2*	*ROUGE-L*
Pipeline - BART	0.38	0.28	0.38
BART modified	0.40	0.28	0.40
T5	**0.47**	**0.33**	**0.42**
PEGASUS	0.42	0.29	0.40

Figure 1.

11. Methodology

An Iterative Incremental Methodology was adopted in the system development. Several testing and reviews were conducted at each step thus enabling system specifications to be error-free and reliable. Testing and debugging were conducted on smaller iteration hence making the task easy and effective compared to testing complete system requirements at once. In the beginning, simpler implementations of a hybrid model for text summarization were designed and developed. Additional functionalities were incorporated into the system in various iterations. The design modification was made and new functional capabilities were added in every cycle. The iterations continued until the development process was completed. This methodology enabled the user to evaluate the system functionality periodically until the final product was delivered. This created room for capturing new requirements and implementing them.
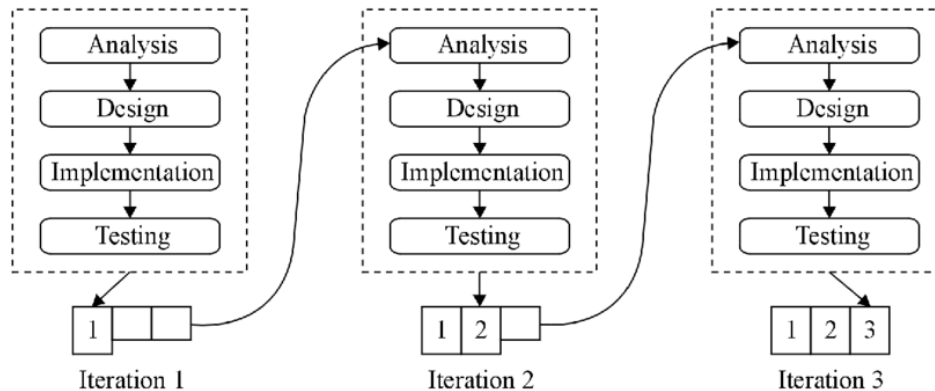

Figure 2. Iterative incremental methodology

12. Planning phase

The hybrid model for text summarization should be able to summarize text content from a pdf document, plain text pasted on the system, and text content scrapped from the website URL (Uniform Resource Locators). Extractive text summarization was conducted first and the output channeled to the abstractive model to generate a hybrid summary.

13. Analysis

In this phase, the specifications of the hybrid model for text summarization were studied based on the problem that had been identified in the planning phase. Analysis was conducted to choose the best logic, database models and to identify any other requirements. The

_____

system architecture was stipulated in this stage. Term Frequency (TF) – Inverse Document Frequency (IDF) and T5 Transformers were to be used in the development of a hybrid model for text summarization.

14. Design and development

The design of a hybrid model for text summarization prototype was produced in this phase. The requirements captured in the previous phases were used to develop the system. An automatic text summarization model was developed in this phase. The model had the capability of summarizing contents extracted from pdf, Wikipedia, and raw text. Python programming language was used in the model backend development. CSS and HTML were used for front-end development. Streamlit framework was used in the front-end development of the model. Streamlit framework helped turn python codes into web apps in a very short time for free and no front-end experience was required. Building an app was conducted with a few lines of code and simple API calls. Widgets were added easily like declaring variables and no backend code was required to describe routes, handle HTTP web requests, either connect a frontend, draft HTML, CSS, and JavaScript. The model adopted both abstractive and extractive summarization techniques.

15. Summary creation process

Data Preprocessing was conducted to convert the data into a machine-readable form of the vector. The process started with the Tokenization of Sentences. The text was divided into sentences. This was implemented via the use of a sentence tokenizer from the NLTK toolkit in Python. Once the paragraphs were divided into sentences, all special characters and stop words were removed. It's conceivable that the text contains some characters that aren't needed. All of the characters that are not needed were eliminated. Word Tokenization was conducted using word space where each of the article's phrases was broken down into words. After the word tokenization, each word's weighted occurrence frequency was determined and then used to generate an extractive summary. The output from the extractive summary was keyed into the T5 transformer model to generate the hybrid summary.

16. Testing phase

Testing began after the current build iteration had been developed and implemented to find and track any potential defects or problems that may have existed in the model. System testing was carried out for each iteration to determine if all the user requirements were captured and implemented. The hybrid model for text summarization was tested to check if it met the research objective. The plain text was pasted on the system and submitted to create a summary for both the abstractive and extractive options. This step was repeated for pdf content and plain text extracted from a website URL (Uniform Resource Locators) to make sure that the research objective was achieved.

17. Evaluation phase

The Iterative life cycle ended at this stage. If there were bugs and requirements not met in the testing stage, the development was subjected to iterations. If no bugs were found, the hybrid model for text summarization was deployed for use. Once the hybrid model for text summarization passed the testing stage, it was ready for deployment in production.

18. Implementation

Python programming language was used in the model backend development while CSS and HTML were used for frontend development.

19. Results

To get started, a dashboard panel was created to act as the user interface for the project. Streamlit Platform is used in the creation of the graphical user interface. This is conducted by importing all the required packages and then creating the GUI (Graphical User Interface).
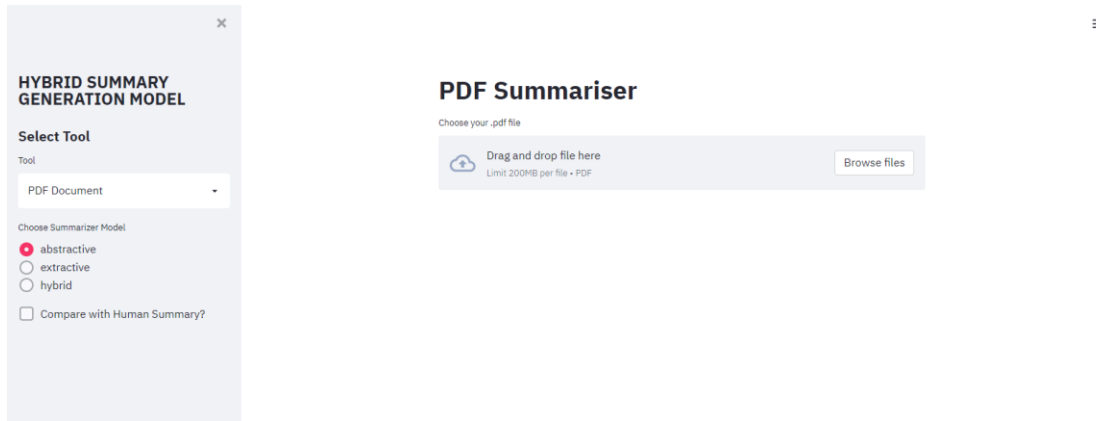

Figure 3. Dashboard

20. Extractive summary model

This summarizer model constructed summary by selecting relevant sentences derived from the original text. The TF-IDF algorithm was utilized in the extractive summarization. Term Frequency (TF) helped to count the frequency of words. The frequency discovered was used to assess the word's significance. The more frequently a word appeared in the source document, the more important it was. The straightforward explanation for TF is that it counted the frequency with which a word emerged or seen in a document. IDF provided unique words with a higher value and repeated words with a lower value. TF occasionally overestimated the significance of stop words depending on how often they appeared.

21. Abstractive summarization model

The abstractive model created its phrases and sentences to provide a more comprehensive summary, similar to the one human being developed. We made use of a transformer network that relied primarily on many levels of attention. For memorizing the sequence of words in the input sequence, it didn't employ RNN and instead relied on attention layers and positional encoding. The global dependencies formed by using several attention layers aided in the parallelization of input processing. Encoder and decoder layers were coupled to a multi-head attention layer and feed-forward network levels in the transformer model. The model used cosine and sine functions to recall the location and sequence of words, resulting in positional encoding. The encoder and decoder layers used a multi-head attention layer and applied a mechanism called self-attention. The input was keyed into 3 linked layers to generate query (Q), key (K), and value (V) vectors. The vectors were subdivided into n vectors (Gupta et al., 2021).

_____

22. Hybrid summarization model

This model was developed by combining the extractive model with the abstractive model. The output of the extractive model was used as the input for the abstractive model thus generating a hybrid summary.

22.1 *Hybrid model rouge score results*

Results from online content (Wikipedia)

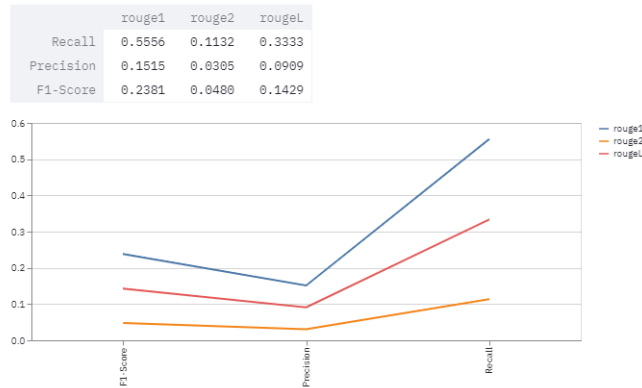|	rouge1	rouge2	rougeL
Recall	0.5556	0.1132	0.3333
Precision	0.1515	0.0305	0.0909
F1-Score	0.2381	0.0480	0.1429

Figure 4. Hybrid results from online content (Wikipedia)

This graph shows results for the summary generated using a hybrid model with an online content source of data. The values for recall, precision, and f1-score for rouge1, rouge2, and rougeL are displayed in the table and a graph is drawn to show their relationship. Rouge 1 has the highest values followed by rouge2 and rougeL has the least values. A very big difference is noted in the values for recall, precision, and f1-score across the 3 rouges (rouge1, rouge2, rougeL). This is portrayed by the line graphs for the rouge score being away from each other.

22.2 *Extractive rouge score results*

Results from online content (Wikipedia)

	rouge1	rouge2	rougeL
Recall	0.4938	0.2264	0.2375
Precision	0.3990	0.1827	0.1919
F1-Score	0.4413	0.2022	0.2123

Figure 5. Extractive results from online content (Wikipedia)

This graph shows results for the summary generated using an extractive model with online content source of data. The values for recall, precision, and f1-score for rouge1, rouge2, and rougeL are displayed in the table and a graph is drawn to show their relationship. Rouge1 has the highest values followed by rouge2 and rougeL has the least values. A very big difference is noted

in the values for recall, precision, and f1-score for rouge1 while the scores for rouge2 and rougeL have a slight difference. This is portrayed by the line graphs for the rouge1 score being away from the rest while the line graph for rouge2 and rougeL being closer to each other.

22.3 *Abstractive rouge score results*

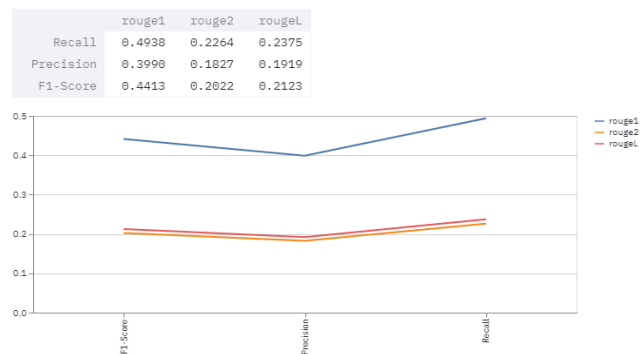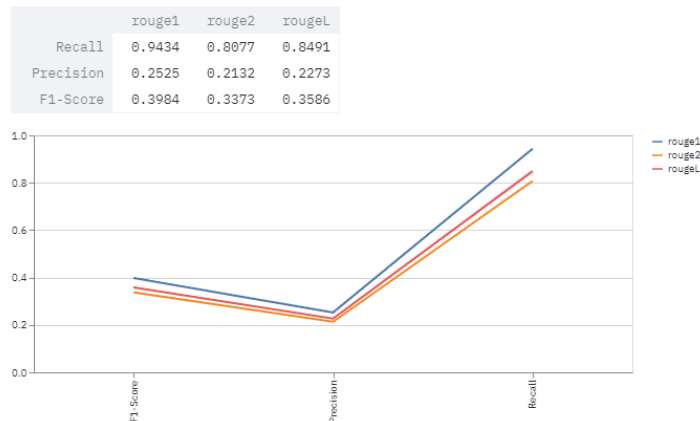Results from online content (Wikipedia)

Figure 6. Abstractive results from online content (Wikipedia)

This graph shows results for the summary generated using an abstractive model with online content source of data. The values for recall, precision, and f1-score for rouge1, rouge2, and rougeL are displayed in table and a graph is drawn to show their relationship. Rouge1 has the highest values followed by rouge2 and rougeL has the least values. A small difference is noted in the values for recall, precision, and f1-score across the 3 rouges (rouge1, rouge2, rouge L). This is portrayed by the line graphs for the rouge score being closer to each other.

23. Discussion

Table 1. Recall average scores

		HYBRID	EXTRACTIVE	ABSTRACTIVE
	Online	0.5556	0.4938	0.9434
	Pdf	0.5333	0.3822	0.9434
ROUGE 1	Text	0.6	0.3618	0.8103
	TOTALS	**1.6889**	**1.2378**	**2.6971**
	AVERAGE	**0.562966667**	**0.4126**	**0.899033333**
	Online	0.1132	0.2264	0.8077
	Pdf	0.1364	0.0705	0.8077
ROUGE 2	Text	0.1111	0.0596	0.6667
	TOTALS	**0.3607**	**0.3565**	**2.2821**
	AVERAGE	**0.120233333**	**0.118833333**	**0.7607**
	Online	0.3333	0.2375	0.8491
	Pdf	0.3556	0.2102	0.8491
ROUGE L	Text	0.3455	0.1908	0.7586
	TOTALS	**1.0344**	**0.6385**	**2.4568**
	AVERAGE	**0.3448**	**0.212833333**	**0.818933333**

_____

From the results in table 1, the Abstractive model has the highest average score for Rouge 1, Rouge 2, and Rouge L followed by the Hybrid model and finally the Extractive model. If we consider the recall ability of the models, it is clear that the Abstractive model emerges as the best model followed by the Hybrid model and the last one is the Extractive model. The summary generated by the abstractive model has more similarity with the one generated by a human being as compared to the summary generated by the other models.

Table 2. Precision average score

		HYBRID	**EXTRACTIVE**	**ABSTRACTIVE**
ROUGE 1	Online	0.1515	0.399	0.2525
	Pdf	0.1212	0.303	0.4967
	Text	0.1667	0.2778	0.2374
	TOTALS	**0.4394**	**0.9798**	**0.9866**
	AVERAGE	**0.146466667**	**0.3266**	**0.32886667**
ROUGE 2	Online	0.0305	0.1827	0.2132
	Pdf	0.0305	0.0558	0.2132
	Text	0.0305	0.0457	0.1929
	TOTALS	**0.0915**	**0.2842**	**0.6193**
	AVERAGE	**0.0305**	**0.094733333**	**0.206433333**
ROUGE L	Online	0.0909	0.1919	0.2273
	Pdf	0.0808	0.1667	0.2273
	Text	0.096	0.1465	0.2222
	TOTALS	**0.2677**	**0.5051**	**0.6768**
	AVERAGE	**0.089233333**	**0.168366667**	**0.2256**

Table 2 shows the precision score for all the models. From this table, the Abstractive model was found to have the highest precision score for Rouge 1, Rouge 2, and Rouge L, while the Extractive model was the second and the Hybrid model had the lowest precision score. If we choose a model in terms of precision score, the Abstractive model was the one recommended.

Table 3. F1-Score average

		HYBRID	**EXTRACTIVE**	**ABSTRACTIVE**
ROUGE 1	Online	0.2381	0.4413	0.3984
	Pdf	0.1975	0.338	0.3984
	Text	0.2609	0.3143	0.3672
	TOTALS	**0.6965**	**1.0936**	**1.164**
	AVERAGE	**0.232166667**	**0.364533333**	**0.388**
ROUGE 2	Online	0.048	0.2022	0.3373
	Pdf	0.0498	0.0623	0.3373
	Text	0.0478	0.0517	0.2992
	TOTALS	**0.1456**	**0.3162**	**0.9738**
	AVERAGE	**0.048533333**	**0.1054**	**0.3246**
ROUGE L	Online	0.1429	0.2123	0.3586
	Pdf	0.1317	0.1859	0.3586
	Text	0.1502	0.1657	0.3438
	TOTALS	**0.4248**	**0.5639**	**1.061**
	AVERAGE	**0.1416**	**0.187966667**	**0.353666667**

Table 3 shows F1-score calculated from all the models. From this table, the Abstractive model had the highest F1 score, followed by the Extractive model and finally, the Hybrid model

J. Mugi Karanja & A. Matheka – A Hybrid Model for Text Summarization Using ....

_____

had the least score. If we were to consider the F1-score, the Abstractive model would be the best choice.

The results for rouge score were used to help in choosing the best-performing model. From the result, there was a very slight or no difference in the rouge score when we consider various sources of inputs (pdf, plain text, and online). This implied that the source of data/input doesn`t affect the quality of the summary generated and also the precision and recall ability of the model used.

The abstractive model was found to have higher recall ability as compared to the other models. It was followed by the Hybrid model and the last one was the Extractive model.

The Abstractive model had the best average precision as compared to all the other models. It was followed by the Extractive model and the last one was the Hybrid model.

On F1-score, the Abstractive model had the best average F1-score, followed by the Extractive model and the Hybrid model had the lowest average F1-score.

24. Execution time

Table 4. Execution Time results

EXECUTION TIME IN SECONDS					
	Online	**Pdf**	**Plain Text**	**Total**	**Average**
Abstractive	80.8308	94.428	67.0551	**242.3139**	**80.7713**
Hybrid	86.7931	95.0698	68.4162	**250.2791**	**83.42636667**
Extractive	2.7926	21.5619	0.8644	**22.4263**	**7.475433333**

From Table 4, the Extractive model took the least time to generate the summary then followed by the abstractive model, and finally, the hybrid model took more time as compared to the extractive and abstractive models. If we consider the time taken for execution, the extractive model becomes the best choice followed by the abstractive model.

25. Models Ranking

Table 5. Models ranking results

ROUGE SCORE & EXECUTION TIME MODELS POSITION RANKING							
	Recall	**Precision**	**F1-Score**	**Execution Time**	**Total**	**Average**	**Position**
ABSTRACTIVE	1	1	1	2	5	1.25	1
HYBRID	2	3	3	3	11	2.75	3
EXTRACTIVE	3	2	2	1	8	2	2

Table 5 displays the rank position of all the models in regards to recall, precision, f1-score, and execution time. The models were ranked depending on the position they hold. The positions were summed up and then averaged to get their overall position. From the research outcome displayed in Table 5, the abstractive model held position 1, the extractive model position 2, and the hybrid model position 3.

_____

26. Conclusion and future works

26.1 *Conclusion*

The study hypothesis was that the hybrid model would have the best results as compared to the separate extractive and hybrid models. From the research findings, each of the models was found to be having its strengths and drawbacks. When considering the recall ability of the models, the abstractive model was found to be the best followed by the hybrid model, and lastly the extractive. The results in terms of precision of the models portrayed the abstractive model to be the best, followed by the extractive model, and finally, the hybrid model was the least. From the f1-score, the abstractive model emerged first, the extractive the second, and the hybrid model the last.

The models' execution time was noted down and used to make an analysis. From the results obtained, the extractive model took the least time to generate the summary followed by the abstractive model and the hybrid model.

The performance of the models generated from the rouge score helped in ranking the models. The abstractive model emerged to have the best results as compared to the extractive model and hybrid model. If the model choice is made depending on the rouge score, the abstractive model was the best choice. The models' performance based on time execution portrayed the extractive model as the best choice followed by the abstractive model. When we find the average score positions from the rouge score and timely execution, the Abstractive model became the leading one followed by the extractive model and the hybrid model was the last. Therefore, the hybrid model isn't the best choice compared to extractive and abstractive as it had been hypothesized in the study when the accuracy and execution time of the summary generated is considered.

26.2 *Future works*

The research was based on text summarization thus more research should be conducted on the automatic summarization of graphical content and video content. It has also not been easy to summarize research papers effectively especially summarizing papers with scientific formulas, equations, graphs, and tables. A tool needs to be established which will be able to understand contents from equations, formulas, graphs, and tables for effective summarizing of scientific papers.

Acknowledgements

References

Boorugu, R., & Ramesh, G. (2020). A survey on NLP based text summarization for summarizing product reviews. *Proceedings of the 2nd International Conference on Inventive Research in Computing Applications, ICIRCA 2020*, 352-356. https://doi.org/10.1109/ICIRCA48905.2020.9183355

Celikyilmaz, A., & Hakkani-Tur, D. (2010). A hybrid hierarchical model for multi-document summarization. *ACL 2010 - 48th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, (July), 815-824.

Chen, Y.-N., Huang, Y., Yeh, C.-F., & Lee, L.-S. (2011). Spoken lecture summarization by random walk over a graph constructed with automatically extracted key terms. *Twelfth Annual Conference of the International Speech Communication Association*.

Chu, W.-S., Song, Y., & Jaimes, A. (2015). Video co-summarization: Video summarization by visual co-occurrence. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3584-3592.

Dave, H., & Jaswal, S. (2016). Multiple Text Document Summarization System using hybrid Summarization technique. *Proceedings on 2015 1st International Conference on Next Generation Computing Technologies, NGCT 2015*, (September), 804-808. https://doi.org/10.1109/NGCT.2015.7375231

Eberts, M., Ulges, A., & Schwanecke, U. (2015). Amigo-automatic indexing of lecture footage. *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, 1206-1210.

Garg, R., Hassan, E., & Chaudhury, S. (2015). Document indexing framework for retrieval of degraded document images. *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, 1261-1265.

Goyal, P., Behera, L., & McGinnity, T. M. (2018). A context-based word indexing model for document summarization. *IEEE Transactions on Knowledge and Data Engineering*, *25*(8), 1693-1705. https://doi.org/10.1109/TKDE.2012.114

Gupta, A., Chugh, D., Anjum, & Katarya, R. (2021). *Automated news summarization using transformers*. Retrieved from http://arxiv.org/abs/2108.01064.

Hassel, M. (2007). *Resource lean and portable automatic text summarization*. https://www.csc.kth.se/utbildning/forskar/avhandlingar/doktor/2007/HasselMartin.pdf.

Li, K., Wang, J., Wang, H., & Dai, Q. (2014). Structuring lecture videos by automatic projection screen localization and analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *37*(6), 1233-1246.

Liao, P., Zhang, C., Chen, X., & Zhou, X. (2020). Improving abstractive text summarization with history aggregation. *Proceedings of the International Joint Conference on Neural Networks*. https://doi.org/10.1109/IJCNN48605.2020.9207502

Mallick, C., Das, A. K., Dutta, M., Das, A. K., & Sarkar, A. (2019). Graph-based text summarization using modified textrank. In J. Nayak, A. Abraham, B. M. Krishna, G. T. Chandra Sekhar & A. K. Das (Eds.), *Soft computing in data analytics* (pp. 137-146). Singapore: Springer Singapore.

Meena, S. M., Ramkumar, M. P., Asmitha, R. E., & Emil Selvan, G. S. (2020). Text summarization using text frequency ranking sentence prediction. *4th International Conference on Computer, Communication and Signal Processing, ICCCSP 2020*, 0-4. https://doi.org/10.1109/ICCCSP49186.2020.9315203

Mihalcea, R. (2004). *Graph-based ranking algorithms for sentence extraction, applied to text summarization*. (4), 20-es. https://doi.org/10.3115/1219044.1219064

MuraliKrishna, V. R., Pavan, Kumar, Y. S., & Satyananda, R. C. (2013). A hybrid method for query based automatic summarization system. *International Journal of Computer Applications*, *68*(6), 39-43. https://doi.org/10.5120/11587-6925

Nguyen, H., Santos, E., & Russell, J. (2019). Evaluation of the impact of user-cognitive styles on the assessment of text summarization. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*, *41*(6), 1038-1051. https://doi.org/10.1109/TSMCA.2011.2116001

Radev, D. ., & Erkan, G. (2015). Proceedings of Document Understanding Conference Workshop. *The University of Michigan at Duc*, 120–127.

Rani, S. S., Sreejith, K., & Sanker, A. (2017). A hybrid approach for automatic document summarization. *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 663-669. https://doi.org/10.1109/ICACCI.2017.8125917

_____

Shimada, A., Okubo, F., Yin, C., & Ogata, H. (2018). Automatic summarization of lecture slides for enhanced student preview-technical report and user study. *IEEE Transactions on Learning Technologies*, *11*(2), 165-178. https://doi.org/10.1109/TLT.2017.2682086

Song, S., Huang, H., & Ruan, T. (2019). Abstractive text summarization using LSTM-CNN based deep learning. *Multimedia Tools and Applications*, *78*(1), 857-875.

Yao, K., Zhang, L., Du, D., Luo, T., Tao, L., & Wu, Y. (2018). Dual encoding for abstractive text summarization. *IEEE Transactions on Cybernetics*, *50*(3), 985-996. https://doi.org/10.1109/TCYB.2018.2876317

_____

COAS

AIMS AND SCOPE

The OJIT, as an international multi-disciplinary peer-reviewed *online open access academic journal*, publishes academic articles deal with different problems and topics in various areas of information technology and close scientific disciplines (information society, information communication technology - ICT, information architecture, knowledge organisation and management, information seeking, information management, electronic data processing – hardware and software, philosophy of information, communication theory and studies, mass communication, information ethics, library and information science, archival science, intellectual property, history of computer technology, development of digital competencies, ICT in education and learning, ICT education, etc.).

The OJIT provides a platform for the manuscripts from different areas of research, which may rest on the full spectrum of established methodologies, including theoretical discussions and empirical investigations. The manuscripts may represent a variety of theoretical perspectives and different methodological approaches.

The OJIT is already indexed in Crossref (DOI), BASE (Bielefeld Academic Search Engine), Google Scholar, J-Gate, ResearchBib and WorldCat - OCLC, and is applied for indexing in the other bases (Clarivate Analytics – SCIE, ESCI, and SCI, Scopus, Ulrich's Periodicals Directory, Cabell's Directory, SHERPA/RoMEO, EZB - Electronic Journals Library, etc.).

The authors of articles accepted for publishing in the OJIT should get the ORCID number (www.orcid.org).

The journal is now publishing 2 times a year.

PEER REVIEW POLICY

All manuscripts submitted for publishing in the OJIT are expected to be free from language errors and must be written and formatted strictly according to the latest edition of the APA style. Manuscripts that are not entirely written according to APA style and/or do not reflect an expert use of the English language will *not* be considered for publication and will *not* be sent to the journal reviewers for evaluation. It is completely the author's responsibility to comply with the rules. We highly recommend that non-native speakers of English have manuscripts proofread by a copy editor before submission. However, proof of copy editing does *not* guarantee acceptance of a manuscript for publication in the OJIT.

The OJIT operates a double-blind peer reviewing process. The manuscript should not include authors' names, institutional affiliations, contact information. Also, authors' own works need to be blinded in the references (see the APA style). All submitted manuscripts are reviewed by the editors, and only those meeting the aims and scope of the journal will be sent for outside review. Each manuscript is reviewed by at least two reviewers.

The editors are doing their best to reduce the time that elapses between a paper's submission and publication in a regular issue. It is expected that the review and publication processes will be completed in about 2-3 months after submission depending on reviewers' feedback and the editors' final decision. If revisions are requested some changing and corrections then publication time becomes longer. At the end of the review process, accepted papers will be published on the journal's website.

_____

OPEN ACCESS POLICY