On the Security of Governmental Hashtags Against Hijacking. A Case Study of #توكلنا Hashtag.

Ohoud Alshabib [1], Amal Alruwaili [2], Waad Almadhy [3], Dr. Osama Ouda [4]

[1] Jouf University, College Of Computer And Information Sciences , 241204013@ju.edu.sa.

[2] Jouf University, College Of Computer And Information Sciences , 421204016@ju.edu.sa.

[3] Jouf University, College Of Computer And Information Sciences , 421205952@ju.edu.sa.

[4] Jouf University, College Of Computer And Information Sciences , omalsayed@ju.edu.sa.

Abstract

Social media plays an important role in spreading data effectively. However, although social media programs have numerous benefits, they also have a share of dangers. Government organizations utilize social media to produce modern thoughts within the intrigued of the government, improve decisions and unravel issues. Twitter is one of the well-known social media programs utilized by numerous government offices as a way to discover the conclusions of the community because it gives a quick and adaptable instrument for disseminating data through the hashtag. The Saudi Ministry of Health launched the "Tawakkalna" application to limit the spread of Coronavirus in the kingdom of Saudi Arabia. Meanwhile, the Saudi Ministry of Health launched the hashtag #توكلنا on Twitter to know people's opinions, receive their suggestions about this application, and provide quick solutions to them if they encounter problems in using this application during the period of social distancing. Malicious attackers might try to steal popular government hashtags with the goal of distorting the reputation of the government in front of its people. The goal of this Paper is to investigate the security of hashtag #توكلنا. We utilize the hashjacker tool to protect this hashtag during the COVID-19 pandemic to ensure that this hashtag can serve as a useful communication tool between the Saudi people and their government during the pandemic. In our research paper, we used the hashjacker tool to protect the hashtag #توكلنا, for the Saudi Ministry Of Health. Through this tool we were able to identify and distinguish tweets on twitter and whether they were hijacked or not.

Keywords: twitter , saudi arabia , hashtag , hijacking   , pandemic .

Introduction

Social media (i.e. twitter, facebook, instagram ….) centers on a particular set of devices that have been particularly planned for social communication. these devices vary in terms of their reason and diverse strategies utilized , but all of them share one objective, which is to encourage communication between clients and share the same substance in a social organizational environment. (Bertot..et.al , (2012))

Social media has the capacity to progress and reestablish connections between individuals in this virtual world without looking at any distinction between them. Online-centric social media such as youtube, facebook, and twitter have an extraordinary notoriety and these strategies are creating since

they work in a two-way (web 2.0) was so that numerous of clients share the same substance to reach certain objectives, and this incorporates the comfort given by these social media As well as encouraging viable communication between the individuals and the government on a theme or issue (kim, et al). A parcel of social media programs have been created and so distant, one of those created programs is the twitter program, which is around sharing news through the web and collaboration between clients in so-called "tweets". clients spread data more broadly among themselves through "retweet". In 2011 there were 65 million tweets with a rate of around 650 seconds per a day. Tweets were initially confined to 140 characters, but on november 7-2017, this constrains was multiplied for all dialects but japanese, korean and chinese. Twitter is still utilized by numerous government offices to discover out what the suppositions of individuals and society in specific approximately an issue, such as the government of central jawa province. (Jati, (2020))

Recent Studies on Information Transparency Using The Hashtag in Twitter

In 2020 (Central Jawa Provincial Government Study) through the participation with twitter indonesia, the government of central jawa province propelled an instrument called kemala (kelola melantas layanan) for their program this device looks for to move forward social media to actualize the commitments of the open specialist based on law no. 14 of 2008 in arrange to reveal open data and this increments the quality of administrations to get it and know the interface of the community and this was done by using the hashtag #kemalajateng within the twitter program (see Figure1).

In this hashtag, the community and the citizen can yield a proposal or complaint and will get a reaction straightforwardly from the central jawa common government. the government Famous that the twitter program made a difference in them increment citizens' certainty in it. the government of central jawa territory, in their twitter account (@provjateng) and other organizations, is attempting to ended up educate that react to citizens and don't waver to yield their complaints and the matter was sporadic, so the government kept on to get complaints through hashtags and without utilizing the hashtag . (Jati, (2020))

Figure1. Hashtag #KemalaJateng . (Jati, (2020))

In expansion, the government eluded complaints to the pertinent offices (see Figure 2). These complaints were not submitted straightforwardly to the offices. On the opposite, the government looks for to address the issue and the agency's knowledge . (Jati, (2020))

Figure2. Tweet response from government. (Jati, (2020))

Not only the central jawa provincial government hashtag, but even the saudi arabia government hashtag related to covid19, which is the hashtag #توكلنا, through which the ministry of health publishes important news on its twitter account @saudimoh, there is also An official spokesperson for the ministry of health in the kingdom of saudi arabia who receives the messages and responds to them in @spokesman_moh, so that we know Everything new about this application and it also receive the opinions of the people it seeks to solve problems, if any, through the hashtag #توكلنا (see Figure 3 and 4).

Figure3. Hashtag #توكلنا

Figure 4. A spokesman for the ministry of health responds to some of the rumors and denies them. (Hassounah…et.al ,(2020))

Indeed, the ministry also speaks through this hashtag to inform every muslim that he is now able to perform the rituals of (umrah) in (makkah) and give permits for that according to specific rules and conditions. All these important and wonderful things are from the ministry of health in the kingdom of saudi arabia and it is very important for us as a people to strive to protect government

and important hashtags of manipulators who seek to harm or discredit the ministry of health in the kingdom of saudi arabia. (see Figure 5).

Figure 5. Services provided in tawakkalna application.

The objective of all social media is to encourage the utilize and sharing of substance with other clients and the interaction among them. social media programs are planned on this premise to effortlessly trade data indeed when clients are each one of them in a far off and diverse geological region. Social media stages are planned for exchange and interaction rather than the already far-reaching conventional implies such as radio and television, which work within the ancient way as a broadcast stage, and are recognized by cutting edge innovation and communications such as computers, cell phones, social media such as twitter, facebook and numerous applications have something unmistakable, which is that radio and TV work in a unidirectional way whereas cutting edge innovation works in a two-way course and may reach work in a multi-directional) .(Kumar.Jha ,(2020))

It ought to be known that this interaction with social media issues critical challenges, counting in connection to protection, security, information administration, openness, administration and other related issues to data arrangement. We endeavor to utilize the hashjacker apparatus to extend the straightforwardness of data on social media .Particularly on twitter utilizing the hashtag, this hashtag is exposed to negative advertisements which will not permit the government from hoodlums to transmit the straightforwardness of data.It is anticipated that the method of expanding the data security of data to the protection of the ministry of health in the kingdom of saudi arabia .

In 2019 (Olga Rodak) the western world saw a move from government to administration after the war, they are related issues always endure, agri-food such as: natural effect, the quality and wellbeing of creatures and the welfare of cultivate creatures. This investigates was done utilizing the hashtag #februdairy on twitter to publicize the british dairy industry, but Shockingly this hashtag has been captured by creature security activists. In this case: two vital designs of ensuring social media, particularly twitter, were highlighted by restriction bunches, to begin with: hashtag seizing. besides: the straightforwardness of the collective commitment. As a result of this hashtag being captured, social relations must be reconfigured through advanced innovation utilizing social media through strategies that really upgrade control over the issue of securing cultivate creatures with showcase mechanisms. (Rodak,(2019))

In 2020 (Wasim Ahmed…et al) it is recognized that the security of the persistent and his privileged insights may be an exceptionally imperative matter between the specialist and the Patient, and not to

exasperate people's certainty within the restorative profession. this investigates was carried out by utilizing the hashtag #shareastoreinonetweet within the twitter program. This hashtag included (1206) tweets by specialists, specialists and other wellbeing experts out of (43374) tweets that were shared in may 2018 and the substance of those tweets was assessed in january 2019 to determine the cases in which names or data approximately their cases were shared. And after that conduct an examination of the substance of the tweets and discover in them the revelation of other patients 'information, and this thing abuses the patients' privacy. (Ahmed...et.al , (2020))

Studies similar to our study

Researchers (Nikita Jain ... et al) conducted their ponder on the subject of hashjacker location and investigation of hashtag capturing on twitter, and which we depended on in our inquire about, utilizing the hashtag on twitter to conversation approximately McDonald's. and their hashtag was #McDStories, utilizing the hashtag #McDStories was impelled trusting that the hashtag would propel heart-warming stories roughly cheerful suppers. instep, it pulled in snarky tweets and mcdonalds depreciators who turned it into a #bashtag to share their #McDHorrorStories and he utilized the hashjacker instrument and clarified how to use the tool as we are going say it afterward in our inquire about, but he did not apply the assurance to the McDonald's hashtag, it was as it were a recommendation, but he classified other categories that have nothing to do with the subject of McDonald's, such as the category of android and apple ... etc. the matter reached them at the end of their research with a Description of the tool and its method of work, even if they not clarified the protection of a specific hashtag. (Jain...et.al , (2015))

At that point came a few ponders that alluded to this investigate (hashjacker — discovery and investigation of hashtag seizing on twitter)and (A, Chokora…et al) they said in (analyzing arizona osha damage reports utilizing unsupervised machine learning) they said in their investigate that consequent to preprocessing the disaster's depiction, a term frequency-inverse archive recurrence (TFIDF) examination is actualized to weight terms for information recuperation. The TFIDF regard builds moderately to the events a word appears up within the record, however is offset by the recurrence of the word within the corpus, which assists with changing for the way that some words appear up more oftentimes as a rule. (Chokor..et.al , (2016))

In a consider conducted by (Emilio Ferrara)and the title of his investigate (measuring social spam and the effect of bots on data dispersal in social media) utilize the hashjacker-detection and examination of hashtag seizing on twitter as a reference in his consider discovering spam campaigns in his data and divided bona fide talks related to tobacco from social spam, he executed a present day procedure. His post utilization comes around: the catchphrases appearing up inside the spam dataset talk to 3.06 million curiously tweets posted more than one tweet. He will center his thought, for the rest of his examination, on these 74,000 energetic spammers. (Ferrara , (2018))

New Challenges Related to Privacy

Why are social media users anxious? is there any justification for their fears? for the most part these concerns come from the nearness of social media in people's lives to a huge degree. 45% of the world's populace employments social systems, and usually prove that there are 3.48 billion individuals who communicate in any way in social media programs, concurring to the information

that i have collected (Hootsuite). This communication can put clients at hazard in numerous ways, particularly when individual information falls into noxious hands. as expressed in (seat believe) that 13% of americans have had their accounts seized in social media programs by unauthorized people, and these breaches may lead to hurt to clients and in common social media stages that have expansive sums of individual data work beneath the Supervision of government is constrained, and on the other hand hoodlums misuse this restricted government supervision as alluring targets for committing extortion and burglary. threats to privacy on social media cyber criminals are interested in deceiving social media Users into gaining access to their sensitive information, stealing their personal data and accessing their accounts, so here are typical threats to social media .(Tulane University , (2021)) :

Data Mining : It is that each individual who employments the web will take off a path of data and information, and indeed when he makes an account in one of the unused social media, he will take off data of his possess which will incorporate title, age, city in which he lives, and other interface. On the other hand, companies collect information almost this user's behavior with their stage and take advantage of this information to superior target advertisements to clients, and in a few cases a few companies share client information with other companies or Third parties without the user's information or assent. And phishing attempts : In phishing, hoodlums attempt to get to individual and delicate data, which is one of the foremost common strategies, frequently through a content message on a portable phone, a phone call, or an electronic e-mail, as in case it were an e-mail from a true blue and dependable institution, and in these messages a trap for their data such as: secret word for a bank account or bank card points of interest . Also malware sharing : These malevolent programs have been planned in arrange to be able to get to users' gadgets and the data they contain. Once the hacking prepare is total for this gadget, it can be utilized to take private data (such as spyware) or through the method of blackmail of stores (such as ransomware), and in social media stages a perfect framework that makes a difference client of these noxious programs. Regularly, in the event that the user's watchword is gotten, the cybercriminal can, through phishing, seize the user's individual account and disseminate malevolent programs to hurt the user's companions or individual contacts. Moreover Botnet Attacks : They are extraordinary robots in social media, which is a robotized account through which you make substance or take after particular individuals at whatever point a particular term is specified. These bots can shape a so-called botnet, where they spread through this organize noxious and hurtful programs on social media, take client information, and dispatch conveyed refusal of benefit assaults DDoS (Denial Distributed of Services) that can offer assistance programmers get to and enter users' devices. (Tulane University , (2021))

Social Media Privacy Issues In 2020

Attacks related to privacy will continue and increase, especially in 2020 and with the approaching elections in america, these attacks are expected to increase, and earlier now Politico reported that the campaigns of attacks will increase and will target the democratic candidates and may have started. And the attackers utilize the same strategies that the trolls utilize to form a data war that targets americans in specific, and advertisements are regularly spread through bot accounts to target gatherings of people, and it is troublesome to completely know what will happen since of these assaults on the 2020 state in presidential elections. (Tulane University , (2021))

Prıvacy Polıcy ın Twıtter

The Twitter program could be a common program that anybody can utilize in all parts of this world, and there are non-public strategies inside it, such as utilizing the private messaging highlight inside the program. You'll be able to utilize the program beneath a nom de plume for you in case you lean toward. (Twitter Inc , (2021))

After you utilize the twitter program as it were to see tweets, twitter will get a few data almost to you such as the sort of gadget utilized as well as the IP address. You'll moreover with twitter select to share other data around you such as email and phone number. twitter employments this data in arrange to extend the security of your individual account and to show tweets. More significant to your substance, comparable occasions and advertisements. (Twitter Inc , (2021))

witter moreover gives you an include to control the data that it'll collect from your individual account, such as controlling things such as get to your account and your address book. You'll be able to download all your data from twitter. In addition to the data that's shared on twitter, it'll to utilize your tweets or the substance simply enjoyed and retweet to be after that to decide what data suits you or intrigued in it and the languages you talk to appear you more pertinent substance concurring to your data. Twitter gives clients straightforwardness in this data, and you'll be able to revision to it at any time. Hashtag is Planned with this tag #, and hashtag could be a key word to encourage the method of looking for a particular data . (Twitter Inc , (2021))

Twıtter Moments Guıdelınes and Prıncıples

Among the thousands and millions of day by day tweets, what is called tweeting or composing ordinary minutes is the most excellent thing that happens on twitter. and through those tweets, numerous individuals can have fun whereas utilizing the twitter application, whether it is at the oscars or at sports clubs at the world glass or indeed in external space. What we see on twitter is most likely what is really happening around us. The taking after are the guidelines or rules that the twitter curation group will take after in making and selecting moments (Twitter Inc , (2021)):

Moments curation policy : these moments are outlined to discover incredible substance within the twitter program. Group of the twitter program attempt to show substance in an easy-to-use and fulfill way to clients and give compelling substance on twitter. these global Guidelines on twitter are designed to illustrate the best possible experience for both content consumers and those whose content is viewed in the moments. And creating great moments : moments are outlined to highlight the alluring substance on twitter, and it must be borne in intellect that the substance must be fitting for all components of the community on twitter, and in some cases these minutes are displayed on the media organize and shown on the twitter stage, the moments will be chosen based on what is happening within the media, for illustration when determination of disputable moments, the shown pictures will be shown on the news circulating within the press, be that as it may, the avoids displaying news that damages the protection , additionally does not rehash the collection in one site for third parties. Also selecting great tweets : the moments on twitter spin concurring to what individuals show from the substance, and it permits the show of person voices and will be faithful to them and will not essentially alter the first setting of the tweet and at last it'll dodge showing the substance of twitter workers. Moreover bias, accuracy, standards : these moments are implied to

grandstand premium, compelling, and non-tampered substance. When deciding substance to post, twitter

Takes under consideration the substance of this tweet and any Other media, as well as the title of the client who is posting this tweet. And in more details (Twitter Inc , (2021)) :

a. Bias : Twitter will make the choice agreeing to what record the information is, as is the case when a disputable tweet is distributed, and it'll appear the tweets that get open interaction and an expansive sum of interest on twitter. the individuals in charge of the twitter program don't need to display their possess sees, but, or maybe they must highlight the sees as they are on the stage.

b. Accuracy: When twitter bargains with critical or essential news, it'll show tweets that contain accurate information approximately the subject, but within the occasion that it appears tweets that contain wrong data, it'll update those tweets with a visual rectification or appear an upgraded tweet and in a couple of cases, twitter erases the tweets or fix it.

c. Standards: whoever needs to distribute substance must dodge terrible words, savagery or any other sort of delicate substance but in cases where a particular story must be told and distributed. Twitter will not advance illicit behavior or take a picture of that substance, but, or maybe it'll allow an alarm to the client that this tweet contains on touchy substance.

And the last one style guidelines: the moment could be a bunch of tweets that are told in creative and persuading ways for the reader employing a wide run of media and incorporate the leading tweets that reflect the picture of the scene or sound to require advantage of these moments given by the twitter program.

Some considerations for creating moments

•Each moment should provide a piece of information or a specific aspect of twitter.

•Every moment should be touching a specific aspect or inspiring the viewer to share it.

•Each moment should be at a time commensurate with the scenes and not precede events.

•Each moment should be in the spotlight on the side of a specific story, event, or people.

Some twitter moments by the twitter curation team include a text below the tweet to add to the content of the previous tweet or be a translation of the tweet. And avoiding conflicts of interest : the twitter organizing group is never dependable for expanding incomes or the Development of twitter clients within the future or for any relationship with directors of twitter companies, and tweets will be shown based on what serves the interface of the open and never in favor of commercial promotions, not indeed twitter's intrigued in terms of exchange, in any case in case it gets to be a program twitter or one of its competitors is the center of a discussion or an issue, and it has gotten to be common, so the twitter group will watch out of the issue, counting the reaction by the twitter company when the fitting reaction is accessible. Also moments created by Individuals and twitter partners twitter permits people as well as accomplices to support tweets, and those tweets are shown on the twitter.com tab (for all clients of the android framework as well as IOS).Therefore the twitter program audits each tweet some time recently showing it and it must meet twitter's possess

guidelines, and a few tweets may be supported by a third party. Other than twitter, in any case, twitter does not take this care into thought when selecting accomplice minutes for consideration on the minutes and investigate tabs. It ought to be borne in intellect that all minutes, whether supported by a third party or not, follow to the complete common administrative benchmarks of twitter. (Twitter Inc , (2021))

What Is The Hashtag?

Hashtag is planned with this tag #, and hashtag could be a key word to encourage the method of looking for a particular data utilizing twitter, and with roughly 50 million tweets per day, it employments this hashtag to organize the data, and within the utilize of the hashtag is an organizational prepare to talk about a particular subject or something among twitter clients and the hashtags showed up political within the iranian races for the year #iranelection and it Was the primary broad news on twitter in 2009, and thinks about show that twitter could be a majority rule implies utilized within the media, and numerous web clients look for to get data around canadian legislative issues through the hashtag #cdnpoli, and in this comes the esteem of the hashtag the political from the characteristic and genuine time truth to the shared information .(Small , (2011))

What Is The Transparency

Transparency means: the method of sharing data in a non-closed way. The concept of transparency within the administration prepare is centered around the compelling interest of All those who have a particular intrigued in decision-making and to guarantee that this administration looks for after the interface of the citizen and is totally free from corruption .(Malhotra , (2015))

Effects of information transparency

The proprietors of companies and everybody who contains an authority position always inquires himself, how is it conceivable to extend the generation handle for everybody who works for me? there are numerous answers to this address, but with the improvement of social media, there has been a strategy of data straightforwardness, particularly on the off Chance that there's a particular choice that must be taken, and it is known that straightforwardness within the working environment is an open communication between all workers within the company, for illustration, to realize the objectives of the collaborator in arrange to realize them and in that an increment within the level of believe among all representatives of the company .(Weitzel , (2014-2021))

The benefits of information transparency

At the point when administrators are more open with their laborers approximately the destinations of the organization, agents increment a hoisted degree of believe towards their executives. More correspondence fabricates superior associations between the two. Agents feel more inclined to work more sincerely for someone that they have a not too bad association with. What's more , compelling correspondence permits the representatives to feel more great voicing their assessments and posing inquiries, which could yield significant Contribution from workers with regards to critical thinking just as lucidity about work measures. (Weitzel , (2014-2021))

Sound associations among the officials and their agents creates believe, however furthermore steadfastness. In this way, a more tight association among trade and agent results In an elevated

degree of productivity from agents since they feel more great within the work environment , another advantage of straightforwardness is more specialist commitment. Straightforwardness progresses agent commitment by including their centrality and causing them to feel more included. Agents are bound to work more eagerly when they realize that Their commitment is critical and when their director reaffirms their significance . (Weitzel , (2014-2021))

Not with standing empowering believe and more prominent commitment, straightforwardness gives specialists a predominant comprehension of what their work is interior the organization and how their commitment has an effect in finishing long and transitory targets. Usually a significant factor since it licenses the agents to achieve their work in light of a specific objective, provoking more effective work. Laborers feel more included when they comprehend why certain things are going on within the organization and can center their work more towards accomplishing a passed on objective. This gives more setting to the laborers, giving them a more significant comprehension of the organization's vision and Qualities. At the point when specialists comprehend the importance of their portion within the 10,000-foot see, their degree of effectiveness increases . (Weitzel , (2014-2021))

How we can introduce more transparency to our organization

Transparency is fair finished through steady and compelling correspondence among businesses and their agents. At the point when regards, goals, and goals are clear to everyone within the affiliation, the way of life of the organization gets to be more grounded and more consistent (Weitzel , (2014-2021)) . One compelling approach to execute straightforwardness into your work environment is by utilizing a program, for case, twitter. twitter is an american scaled down blogging and long run casual communication organization on which clients post and relate with messages known as "tweets". Enlisted clients can post, as and "retweet tweets", in any case unregistered clients can fair examine them .(Twitter Wikipedia ,(2021))

Twitter transparency

The basis of work on twitter is transparency, which has driven to unlimited opportunity of expression, and it is discernible that numerous governments are utilizing twitter to discover out the suppositions of citizens on an issue. accepting within the control of open expression and motivation in it made the twitter group distribute one of those reports demonstrating transparency, and in 2012 the objective was to know the public's sees on the issue of government weight. Everything has changed from 2012 to 2020, and it is exceptionally Critical to shed light on the usage of twitter's rules and the persistent work to disturb any of the operations welcomed from anyplace or indeed inquire for data around account holders .(Twitter Inc , 2020))

What's New in Twıtter (Twitter Inc , 2020))

•A modern website that contains all the information that can be seen.

•Data visualizations to facilitate the comparison process over time.

•Unit of comparison between all countries

•Certain techniques to help with how to use the twitter program.

•Transparency and all updates to the application are registered with a specific record

•Modern ways to enforce twitter rules from july 2018 to december 2019.

•Modern political laws to better merge with the twitter rules.

Work On Informatıon Operatıons

The document of the data prepare ought to happen on a standard premise and intermittently after its expulsion from the twitter program, and as of late the biggest divergence was among the 32,242 accounts. (Twitter Inc , 2020))

Platform manipulation in twitter

This term came to confront manipulation through the twitter platform to spread misconceptions in what are called robots on the twitter program, and the policy focuses on this concept, especially the behavior because here the content is not important and this is done through mail from people seeking to manipulate the name of the platform on twitter and Twitter revealed this behavior and confirmed studies have decreased by 10% because when twitter requests a phone number or an electronic email for the account holder, verification is done by filling in a code sent to the user's email or phone in order to make sure that there is a person behind that account and not a robot account. (Twitter Inc , 2020))

Terrorism and violent extremism

The twitter laws never allow the so-called terrorism and the necessary decisions have been taken regarding 86,799 twitter accounts, and 74% of the accounts have been frozen using twitter's internal tools by the team, and the twitter company does not object to cooperating with it by users to find out Whoever uses terrorism and twitter will continue to eliminate this category as much as possible through its own services in accordance with certain laws. (Twitter Inc , 2020))

Twıtter Rules Enforcement

The segment was extended to be in line with the unused twitter laws, which was for the primary time and a parcel of information was shared around what related to the damaging arrangements, and this came in understanding with santa clara's laws with respect to straightforwardness and responsibility for everything related to the substance. Due to the Profundity of center on the insulting substance, the arrangements that take after precise frameworks and detailing devices in a particular way, and more information was included for about 12 diverse ranges of its segments within the approach and accounts that were solidified or closed due to damaging the twitter laws expanded by 47%, and usually certain the rate will not stay stagnant, because it will increment or diminish with the modern developments. (Twitter Inc , 2020))

a. Abuse : Due to the expanding rates of social media utilize and innovation in common, the rate expanded until it come to 95% of the accounts on which the fundamental choices were taken, counting security breach or manhandle. Amid the period specified within the report, there was an incremezt within the accounts that twitter chosen on around these accounts . (Twitter Inc , 2020))

b. Hateful conduct : The hateful behavior expanded to reach 54% of the number of accounts that twitter decided on, and that hateful behavior expanded to include dehumanization on july 9, 2019 , twitter took appropriate action on those accounts . (Malhotra , (2015))

c. Sensitive media, including graphic violence : There was a 39% increase in the number of accounts actioned for violations because of its infringement media policy. (Malhotra , (2015))

d. Promoting suicide & self-harm : The client ought to not permit others to drop beneath those calling for self-harm, and they must be detailed since there was a 29% increment within the accounts with which the essential activity was taken. (Malhotra , (2015))

e. Illegal or certain regulated goods or services : In addition to that, there are new disclosures about anyone who violates twitter's policy on goods or any illegal service, and that represented 60,807 account . (Malhotra , (2015))

f. Private information : To share the individual data of certain individuals without their information or assent could be an infringement of the twitter laws, and the inside devices on twitter have moved forward, which driven to the total execution of this approach and there was an increment within the rate until it come to 41% of an account that abused these laws and the fundamental activity was taken . Included within The chronicle. This file is utilized by researchers and analysts from all the posts of the world, it contains more than Terabytes of Media and has more than 83,000 accounts and much more than 200 million tweets, and the method bunches that have been made in more than 15 nations have been discharged, and this thing gave analysts a particular see at how it created forms related to benefit data, and this too broadens how we work with accomplices to make strides understanding forms around data operations. Prior two educate specialized in investigate, the australian key approach founded (ASPI) and the stanford web observatory, collaborated to analyze and survey information regarding revelation, and after that it facilitated the primary #infoops2020 conference to realize Organization with carnegie to combat affect operations. in this event, it was concurred to extend openings for participation and inquire about on the so-called ion supporting thoughts traded between twitter clients and the look community . (Twitter Inc , 2020))

g. Violent threats : Twitter saw a decrease in the number of accounts seeking violations of violent policies, including threats, by 5% . (Twitter Inc , 2020))

Legal Requests

In addition to the twitter rules imposed on users, it is also advisable to take measures to respond to legal purposes and requirements . (Twitter Inc , 2020))

Information requests (legal requests for account information)

It ought to be famous that governments and organizations mindful for upholding the law gotten an ask for data around the world by 21%, and it is detailed that the overall number of accounts has expanded to 63%, and the volume of data ask is the biggest from 2012 until presently, and from 2012 an ask for government data was gotten from 91 diverse nations. Identity of borrowed or mysterious accounts. With the proceeded ask for data gotten from the joined together states of america to put the

most noteworthy rate of demands within the legitimate handle for getting data for a particular account, 26% of worldwide demands for an account showed up from inside the united states of america, and after that the moment the biggest ask comes from a japan accounts for 22% of worldwide orders. (Twitter Inc , 2020))

The utilize of a pseudonym or anonymous discourse is one of the essentials on twitter in some cases in arrange to guard the voice of the open and guarantee the entry of their voices, And frequently twitter gets non-governmental data demands to uncover private account information on twitter namelessly or beneath pseudonyms. Denying to do so, twitter rejected 23 demands from america to uncover the personality of borrowed or mysterious accounts. (Twitter Inc , 2020))

Removal requests (legal requests for content removal)

Twitter has obtained approximately 27,538 legal requests to remove 98,595 content from twitter accounts, which is a very large number since the first transparency report was issued in 2012 . The global volume of legal requests around the world Originated from three countries: japan, russia and turkey, and was 86% . Increasing from the legal requests from Japan significantly to about 143%, which represents 45% of the requests from around the world received by twitter, and it is worth noting that the 12,496 requests coming from japan focus on drugs and other psychotropic substances . (Twitter Inc , 2020))

Copyright and Trademark Actions

a. Copyright violations: This infringement expanded by 13% in arrange to expel the subject concurring to the thousand years copyright act, and this rate influenced roughly 163% of the accounts on twitter . (Twitter Inc , 2020))

b. Trademark takes note: Expanded by 7% of the overall number of past commercial takes note within the final report . (Twitter Inc , 2020))

Method and Result Discussion

The hashtag is utilized on twitter to know the diverse sorts of tweets and their classifications and to encourage more precise recovery of tweets. Capturing happens when a bunch of clients employments a common hashtag to advance a theme that's totally diverse from the setting of the discussion within the original hashtag . (Chokor..et.al , (2016))

This hashtag works to supply a more viable tweet in arrange to draw in the consideration of clients since of this typical tag, and with the increase of users' visit to this hashtag, it'll be placed within the list of "trending subjects" and in a surprising and quick way, it'll turn into An imperative point that numerous clients utilize and talk on twitter. And numerous offenders tend to capture this common hashtag for posting negative advertisements, posting unseemly substance, or slandering a company or individual, and in this sort of hashtag, it is counterproductive and harming to twitter's reputation. in this manner, in this inquire about, we look for to secure the central java common government's hashtag from seizing and programmers by considering the hashjacker tool. Hashjacker instrument this apparatus identifies and analyzes the seizing of tweets from programmers and offenders. It may

be through joining an awful interface or connecting a connect that's not relevant to the subject additionally incorporates a talk of unessential content. Hashtag tweets are focused on to draw in the consideration of clients and to assault the trade of commercial companies, particularly those that have incredible notoriety, where the aggressors express negatively pointed at Ruining, as well as those hoodlums look for to spread awful substance to hurt those governments or companies on social media. the thought in our investigate is that the hashjacker device first collects the foremost well-known hashtags day by day, around a Month, that's , inside 30 days, and after that the hashtags we got are sorted into one of these categories , see Table 1.

Table 1:

Category for Different Hashtags

Category	#Hashtag name
The Ministry of Health in the Kingdom of Saudi Arabia	الجواز_الصحي# توكلنا#
Epidemic	#Coronavirus #COVID19 #COVID19uk
University	#Bachelor,#Master #Graduate_Student
Sports	#Football #Tennis #Basketball

We are presently attending to portray how this modern apparatus can really distinguish the seizing of the initial hashtag of a specific category. The starting of this apparatus will make an investigation based on a lexicon in which all the tweets that contain a specific hashtag from a particular classification are found, so that the tweet that contains the foremost number Of sentences or words indistinguishable to the word reference, so we conclude that it may be a tweet that's more important than others. And since this instrument takes after the common drift, the hashtag is distributed and talked almost on social media, and after that within the event that the tweet taken after a particular heading or coincided with other tweets for a particular hashtag, at that point the tweet will be non-hijacked. as for the tweets that conversation around disconnected themes, typically a sign and there's a tall likelihood that they are pernicious tweets or have other targets, whether they are taunt or spam mail pointed at purposefulness hurt. Why ?

Because it is clear that it does not follow the focus of the conversation as in the other tweets , it is not interested in talking about the content of the hashtag, as it seeks to harm and does not intend to talk about the classification of this hashtag the goal of these malicious tweets is to misuse a specific hashtag as well as to attract the attention of users speaking inside the hashtag to follow this tweet.

And such tweets that follow a common hashtag system is to Spread all unwanted messages and the purpose is to distribute them to the users to reach the largest possible number.

Here is a description of the flow chart that the tool uses to detect hashtag hijacking .See Figure 6 . (Jain...et.al , (2015))
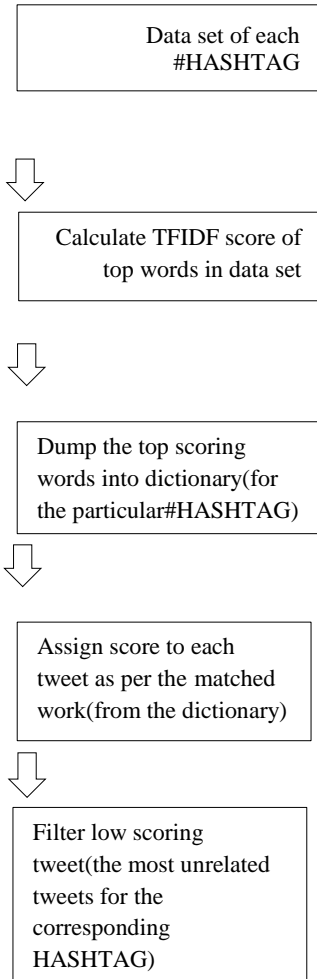
```
┌─────────────────────┐
│   Data set of each  │
│      #HASHTAG       │
└─────────────────────┘
           ⇩
┌─────────────────────┐
│ Calculate TFIDF score of │
│  top words in data set   │
└─────────────────────┘
           ⇩
┌─────────────────────┐
│ Dump the top scoring │
│ words into dictionary(for │
│ the particular#HASHTAG)   │
└─────────────────────┘
           ⇩
┌─────────────────────┐
│ Assign score to each │
│ tweet as per the matched │
│ work(from the dictionary) │
└─────────────────────┘
           ⇩
┌─────────────────────┐
│ Filter low scoring   │
│ tweet(the most unrelated │
│ tweets for the       │
│ corresponding        │
│ HASHTAG)             │
└─────────────────────┘
```

Figure6: hashjacker flow chart . (Jain...et.al , (2015))

Each hashtag has data, and HashJacker displays the largest words used in tweets by means of what is called TFIDF for all words in about 1000 tweets, after which the TFIDF value will increase in an orderly and consistently with the number of times the word used appeared in the document. It should be borne in mind that a word may be repeated in a group and this helps to understand the fact that there are words that will appear frequently in general, and by using TFIDF we will obtain the most frequent words in a particular set of data. After we got the most important words through TFIDF. The hashjacker tool will create a file for this file called Dictionary in order to put in it the most important and most important words that you got from the above in each hashtag.

It is worth noting that these multiple dictionaries will take the first 20,000 tweets from different hashtags and in different categories. Because it is possible that there are hashtags belonging to more than one category, it is certain that the HashJacker tool will also create a common dictionary for each type considered, which combines the most important words for all hashtags in a category. The words that will get the highest score are already the most relevant and most used words in relation to a

Ohoud Alshabib, Amal Alruwaili, Waad Almadhy , Dr. Osama Ouda .

specific hashtag from a category. And since the dictionary will be based on a specific category, it is possible to re-use it in the future for common hashtags belonging to the same category and re-include more common words in it such as in Table 2.

Table 2:

Top Scoring Words for Different Hashtags

HASHTAG	Top scoring words
#توكلنا	معالي وزير الصحة , وزارة الصحة , لقاح كورونا
#COVID19	Symptoms , Washing hands , Keeping distance , Set in home
#Master	Degree , Quiz , Research paper , Complete study
#Tennis	Sport marketing , Team , Golf , Workout

In addition, each result will be marked and each tweet will be marked based on the matches we get from each tweet with the words in the dictionary, and since the twitter program determines the number of characters for each tweet published to be 140 characters, which gives an average of 30-35 words (paying attention to the letters per word) and each tweet can get the highest score if it's at the same level or category. A tweet can score as low as 0 or 1. these low-rated tweets are viewed as most harmful first, irrelevant or potentially irrelevant messages. Finally, hashjacker tool will separate each low-score tweet and place it in a completely separate file from the tweets with the highest score in each hashtag. given that the study was conducted roughly, we concluded that the majority of hacked tweets are from a group of users who create accounts to distort a specific hashtag on social media, or it may be from a group of accounts that seek to spread offensive or satirical content on social media by Exploiting these hashtags. The hashjacker tool is a very good tool in identifying and classifying hashtags distinct from the hashtags that cyber criminals seek for harm of any kind, and this tool will help the ministry of health in the kingdom of saudi arabia to protect the hashtag From hackers and for the government to keep receiving the opinions of citizens and seeking to solve their problems without a criminal has tampered with the hashtag #توكلنا. Finally, this tool will calculate the percentage of tweets that have low scores, as follows:

The percentage of hijacked tweets for a specific hashtag = (the number of low-grade tweets / the total number of tweets) * 100% such as in table 3, and it is possible to perform many tests for this tool and it also does not need to create a new dictionary for each new hashtag, on the contrary, you will put this hashtag in any a category that belongs to it and then reuse the dictionary again, and it is also

possible to add a new word that does not exist to the dictionary for that category . See Figure from 7 to 10.

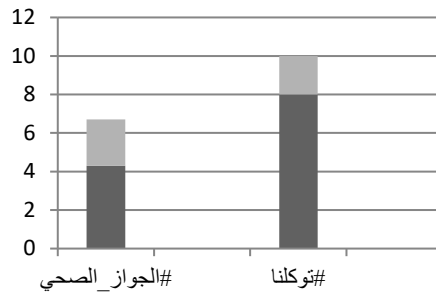*Category of the ministry of health in the kingdom of saudi arabia:*

Figure7: Percentage of hijacked tweet for category- the ministry of health in the kingdom of saudi Arabia.
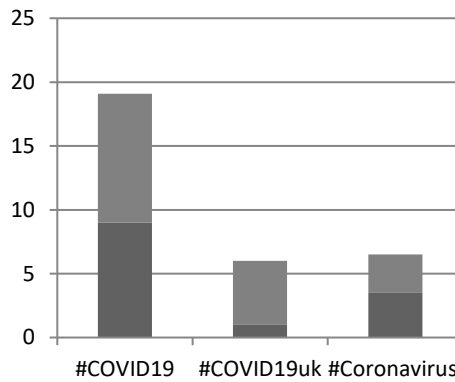
Category of epidemic :

Figure8 : Percentage of hijacked tweet for category- epidemic
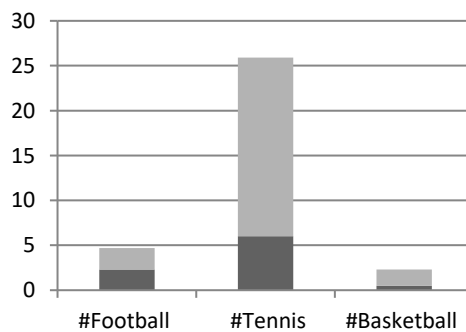
Category of sport :

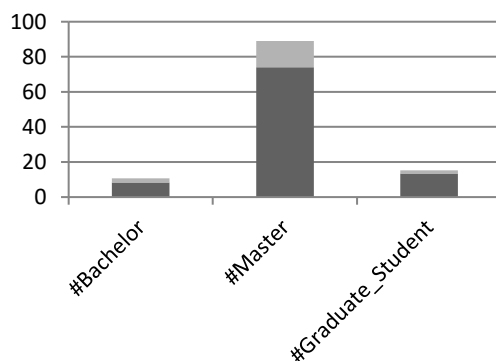*Figure9 :* Percentage of hijacked tweet for category- sports

Category of university:

Ohoud Alshabib, Amal Alruwaili, Waad Almadhy , Dr. Osama Ouda .

Figure10 : Percentage of hijacked tweet for category- university

Table 3 :

Scores for Different Category.

Category	Tweet	Score
The Ministry of Health in the Kingdom of Saudi Arabia	للراغبين في حجز تصاريح عـمره تواصـل خـاص ..	0
Epidemic	169 more deaths and 9893 new covid-19 positive cases reported	10
University	Stop talking against Ameer muaawyia #kohat	0
Sports	#Kashmiri Women display their skills in Sport by participating in the #Uplona Girls Open Volleyball Tournament.	20

*(The number of low-grade tweets / the total number of tweets) * 100% , approximate and default number :*

Suppose #توكلنا= %100*(65687/52)= 0

Because the experimental results show that only 52 tweets out of 65687 checked tweets are low-grade tweets that are not relevant to the studied hashtag #توكلنا.

In #COVID19 = (9740/90,000)*100%= 10 √, because the experimental results show that only 9740 out of 65687 checked tweets are non-malicious tweet that talks about the hashtag #COVID19 content.

And in #Tennis = (1980/9800)*100% = 20 √ , because the experimental results show that only 1980 out of 9800 checked tweets are non-malicious tweet that talks about the hashtag #Tennis content.

Also #Master = (40/50.000)*100%= 0 , because the experimental results show that only 40 tweets out of 50.000 checked tweets are low-grade tweets that are not relevant to the studied hashtag #Master.

And when we get the number zero, that is an indication that it is a hijacked tweet, in contrast to the tweets that we got from them with a high grade.

Conclusıon

A hashtag may be an image utilized on social media channels like Twitter, which empowers clients to discover a particular subject or type of content. Using hashtags in tweets is a means of aggregating information on a topic , Are not case-sensitive, but adding capital letters does make them easier to read: #BelieveInDreams vs. #believeindreams , many government are using it as a marketing strategy while others use it to vilify their competitors. in our research paper, we used the hashjacker tool to protect #توكلنا hashtag for the saudi ministry of health , and this tool can be easily used to detect spam through trending hastags, and we found in this tool the best way to reduce hacker manipulation of exploiting hashtags. Once detected various and hackers means as well as intelligent Awareness can be used to curb this to an Extent possible. Remember, not only can people be defrauded in actually only but also the fraudsters use information online to send bogus materials, solicit or phish to harm all users.

In the recommendations of our research, we can say that we propose to the developers of the twitter program to reformulate privacy and policy by giving more safety to the hashtags that relate to governments , and in this policy to guarantee every government sectors the safety of their desires, and when the policy is strong, therefore we do not need to use a tool to protect governmental and important hashtags. Because with the development of technologies and tools from hackers, this tool may not work well in the future, while when the twitter policy becomes strong and provides more security for government hashtags, then the results will vary greatly.

Acknoweldgemet

References

1. J. C. Bertot, P. T. Jaeger, and D. Hansen.(2012) .The impact of polices on government social media usage: issues, challenges, and recommendations , Government information quarterly, Vol. 29, 30-40.

2. H.Jati . (2020) . Application of social media twitter as a strategy for government's transparency: study on #kemala jateng program , forum ılmu sosıal Facility of sosial science Vol.47, 1-7.

3. M.Hassounah , H.Raheel , M.Alhefzi. (2020) . Digital response during the covid-19 pandemic in saudi arabia , JMIR Publications advancing digital health & open science, Article ID 32790642 , Retrıeved from-https://www.jmir.org/2020/9/e19338/.

4. C.Kumar.Jha .(2020) . Information control, transparency, and social media: implications for corruption. In the book " Media controversy : IGI global publisher of timely knowledge" Chapter 23, 1-19.

5. O.Rodak .(2019). Hashtag hijacking and crowdsourcing transparency: social media affordances and the governance of farm animal protection , Springerlink, Vol.37, 281-294 .

6. W.Ahmed , R.Jagsi , T.G.Gutheil, M.S.Katz .(2020). Public disclosure on social media of identifiable patient information by health professionalsals: content analysis of twitter data , JMIR Publications advancing digital health & open science , Article ID 32870160 , Retrıeved from- http://preprints.jmir.org/preprint/19746.

7. N.Jain , P.Agarwal, and J.Pruthi. (2015) .Hashjacker- detection and analysis of hashtag hijacking on twitter , International journal of computer applications Vol.114, 17-20.

8. A.Chokor , H.Naganathan , W.K.Chong ,and M.El.Asmar.(2016). Analyzing arizona osha injury reports using unsupervised machine learning , Sciencedirect Vol.145, 1588-1593.

9. E.Ferrara. (2018) . Measuring social spam and the effect of bots on information diffusion in social media , Springerlink Vol.24, 229-255.

10. Tulane University.(2020). Key social media privacy issues for 2020 , School of professional advancements, Retrıeved from- https://sopa.tulane.edu/blog/key-social-media-privacy-issues-2020.

11. Twitter Inc .(2021). Twitter privacy policy , Privacy policy , Retrıeved from- https://twitter.com/en/privacy.

12. Twitter Inc.(2021). Twitter moments guidelines and principles, General guidelines and policies , Retrıeved from-https://help.twitter.com/en/rules-and-policies/twitter-moments-guidelines-and-principles.

13. T.A. Small.(2011). WHAT THE HASHTAG? , Taylor&Francis online, Vol.14, 872-895 .

14. C.Malhotra .(2015). Role of social media in promoting transparency in an open government era in saarc countries with special reference to india , IM OD EV Improving public policies in a digital world Vol.1.

15. A.Weitzel .(2014-2021). Effects of information transparency , Teambay, retrıeved from-https://teambay.com/information-transparency/.

16. Twitter .(2021) .Twitter From Wikipedia the free encyclopedia , Wikipedia the free encyclopedia, Retrıeved from-https://en.wikipedia.org/wiki/Twitter.

17. Twitter Inc .(2020). Introducing the new twitter transparency center , Twitter company ,Retrıeved from-https://blog.twitter.com/en_us/topics/company/2020/new-transparency-center.html.