JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF INFORMATICS AND INNOVATIVE SYSTEMS

PROJECT TITLE: MITIGATING UNCONTROLLED SOCIAL MEDIA UTILITY

PRESENTED BY:

COLLINS O. BUNDE I132/0858/2013

EDWARD ONYANGO I132/3025/2013

A PROJECT PROPOSAL SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF BACHELOR OF SCIENCE IN COMPUTER SECURITY AND FORENSICS

2016©

SUPERVISOR

DR. SAMUEL LIYALA

ABSTRACT

Having been started as a technology focusing on the consumer, social media offers critical business advantages to enterprises, it is being leveraged as a powerful tool driving business objectives such as improved customer relations, better brand recognition and enhanced employee recruitment. While it affords companies many potential benefits it presents inherent security risks such as cyber loafing, cybersquatting, phishing and data leakage. Enterprises willing to incorporate social media into their business process must adopt an all-inclusive social media strategy that encompasses effective social media policies. Furthermore enterprises should consider establishing local enterprise intelligent systems that focuses on social media monitoring, for real-time social media monitoring and analysis of employee and customer sentiments. The study attempts to identify the existing risks in social media, and suggest a mitigation strategy for effective social media use that defines social media policy formulation guideline for use within the Kenyan industrial realm, Furthermore we propose a contextualized methodology for brand reputation tracking for the enterprises that can be replicated locally among the Kenyan companies. Finally, we suggest the implementation of such system for enhancing privacy and security in social media by the Kenyan enterprises. The methodology used for this project is the agile methodology, where scrum is used in this case for the system development and exploratory research design is used for the research, with regards to risks and privacy issues in social media. This project is significant since it helps in addressing the social media risks and provides well defined mitigation strategies against privacy and security concerns through enforcement of policies, in addition to real time monitoring through sentiment analysis which will provide rich insightful intelligence that will help in profiling users both insider and outsiders by arresting bad posts before they create an uproar in the social media.

Table of Contents

Keywords/Abbreviations

n-Gram - In linguistics, an n-gram is a sequence of n words used in text or speech. When doing natural-language analysis, it can often be handy to break up a piece of text by looking for commonly used n-grams, or recurring sets of words that are often used together.

C.I.A confidentiality Integrity Accountability

E.S.N.Ts Enterprise Social Media Networking Tools.

A.P.I Application Programming Interface

REST Representation State Transfer

Chapter 1

1.0 Introduction

Web 2.0 and social media has widened opportunities for communication and networking, the information flows on the internet can be accessed through the use of open source intelligence tools, by crawling them and analyzing them for various reasons such as behavior prediction and opinion mining The introductory part of this chapter talks about social media use in within the Kenyan enterprises. Kenya like many other countries, has embraced the ideals of social businesses. Where social media technologies are being exploited successfully by the corporates to sell their brand identity, increase sales of product and services as well as augment customer satisfaction. Further on it goes to give an overview of the problems that arise from the use of these enterprise social media networking tools (E.S.N.Ts). Risks of these social media tools to the enterprises is the problem statement. Consequently it underlines the study objectives and its significance stating that the project will be significant in emancipating the enterprises on the need for establishing effective social media strategy which includes policies, risk assessment guidelines before the uptake of any of these social media tools and implementing local enterprise intelligent systems that can monitor in real time social media use within these enterprises. It is a study of how enterprises can effectively mitigate against the risk of social media use within the enterprise environment. The scope is limited to Kenya with a focus on the formulation of a social media strategy and the implementation of local enterprise intelligent systems as measures that will eventually help in mitigating this risks. Furthermore the assumptions of the study are indicated and the research questions formulated.

1.1 Background information

Social media offers business advantages to both private enterprises and government agencies. Organizations use this social media to reach out to the digital audience efficiently and cost effectively. Social media tools are thus being embraced in the business world as important game-changing customer communication platforms. They necessitate networking with current and potential customer besides offering aid in promoting brand awareness in many different markets.

Furthermore, it has revolutionized the way people engage with the internet, opening a world of constant ‘anywhere, anytime’ access thanks to the plenty and popularity of smartphone devices such as Android, Blackberry and iPhones (A Cyveillance White Paper, 2015).

In a 2009 survey of companies that participate in online social media communities, 70 percent of respondents reported using social media of some kind in their businesses (Gordon, n.d.). Over 40 percent of such companies had employees whose job function included spending time on social media sites in order to maintain an organizational presence. More than a quarter of these companies maintained social media sites for employees’ personal announcements and social events. Fewer than ten percent blocked access to social media for any employees.

In Kenya, the more affordable tablet computers and the rise of smartphones in the country, not to mention, the competitively priced mobile data, converge to bring us to the age of the smart

audience (Kaigwa, Mark; Madung, Odanga; Samer, Costello;, 2015). Kenya, like in many other countries, enterprises are embracing the ideals of social media businesses. Social media technologies have been exploited successfully by the corporates to sell their brand identity, increase sales of product and services as well as augment customer satisfaction. This has already resulted into a dichotomy between the enterprises that have embraced this technology and those that are yet to, or mostly avoid it.

The bottom line is with all these roles and responsibilities; every enterprise needs an elaborate risk management plan that addresses the inherent risks. The starting point is always supposed to be a social media risk assessment, which should help institutions ascertain the key intrinsic risks that, exists and identify the appropriate mitigation strategies. Dr. Michael Ong once said, "Good Risk Management fosters vigilance in times of calm and instills discipline in times of crisis." (Macharia Kihuro, 2015).

Practical case studies in the Kenyan context of how social media has affected enterprises or employees. Chase Bank being put under receivership after the online uproar that tore the company to its core, arguably Chase Bank was brought down by social media. The case of a famous Disc Jockey, Delacreme being a victim of cyberbullying when explicit videos were posted on twitter and attracted nationwide tweets and retweets ultimately almost jeopardizing his career this and many untold stories are the impact of unmonitored social media.

1.1.0 Kenyan vision 2030 ICT policy objectives

Our project considers the Kenyan ICT objectives as the driving force, that is by harmonizing our goals with the I.C.T policy objectives we can as well be playing a part in the in the realization of the vision 2030 goals. The I.C.T policy objectives:

1. Points to establishing a cyber security strategy as a key objective of national security. In this regard, effective social media policy should be inculcated into the cyber security strategy for the achievement the vision 2030 goals.
2. Encourages the development of cyber security strategy with the aim of driving economic and social prosperity while at the same time protecting cyberspace reliant societies against cyber threats. Enterprises will ultimately promote economic growth by implementing appropriate mitigation strategies against social media risks.
3. An analysis of the security impacts of social media, for instance, cyberbullying would provide a regulatory framework and insight into technical solutions and law enforcement strategies and this would be appropriate for the detection and prevention of cyber threats across the country according to the Kenyan ICT policy strategy.

1.2 Statement of the problem

Social media has tremendous benefits but it also has serious security risks for organizations. Without a proper social media strategy, enterprises will fail to realize these benefits. Corporates are faced with a big challenge due the inherent risks they face with the rapid uptake of social

media. Corporate policies have largely not kept up with the rapid adoption of social media. Hence there's need for a social media policy to augment the existing information security policies within these enterprises. Besides people (employees) being the weakest link in the security chain, they do not make it any better for corporates, with the latest security technologies that cannot protect employees, giving away information on social networks, or using their own insecure mobile devices for business purposes. Furthermore, the corporate perimeter is not defined due to the ubiquitous nature of smartphones and other mobile devices that have facilitated the anywhere anytime internet reality. A security model that includes a social media strategy, policy formulation and risk assessment coupled up with local enterprise intelligent systems can help them mitigate this risks.

1.3 Objectives

1.3.1 Main Objective

- To identify the existing risks of social media utility across Kenyan enterprises.

1.3.2 Specific Objectives

- To suggest mitigation strategy for effective social media use that defines, policy formulation guideline for the enterprise use within the Kenyan industrial realm.
- Propose a contextualized methodology for brand reputation tracking within the enterprise that can be replicated locally among the Kenyan companies.
- Suggest the implementation of such intelligent systems for enhancing privacy and security in social media

This project monitored twitter content generated by Kenyans, online content that was monitored included tweets, comments, and mentions. This was exclusively done by monitoring English words since our classifier was only trained on the English words. We did not focus on spoken vernacular languages such as Swahili and Sheng'.

From the twitter platform, we were able to classify user comments as either positive or negative and append a polarity or confidence score which determined how negative a particular sentiment was. The findings, as well as the methodology used, are discussed in the later chapters.

1.4 Research questions

What risks are most Kenyan enterprises exposed to as result of uncontrolled social media use?

How will local enterprise intelligent system such as sentiment analysis systems, help these enterprises in mitigating the risks arising from the social media use.

What are the mitigation strategies implemented for social media surveillance to help in dealing with suspicious user activities on a timely basis, like flagging off negative comments/statements, or issues of negative brand images to these enterprises? Before they get out of hand?

1.5 Scope of the study

The project will be executed as a study of the enterprises within Kenya, taking a casse study of Safaricom, the study will focus on the existing social media risks to these enterprises and the resultant impact to the enterprise environment. In addition, we look at social media strategy as being of importance to companies that have embraced social media and the need for risk assessment and formulation of effective social media policies. Finally, we implement a sentiment analysis system to monitor, employee and consumer comments towards corporates and their brands. The methodology used will be the scrum methodology for the implementation phase where we use sentiment analysis techniques by leveraging the use of the Twitter platform as the social media case study for the project. The research design will be exploratory research since it relies on secondary research such as reviewing available literature in addition to the use of the internet.

1.6 Assumptions of the project

Most Kenyan enterprises are using twitter social media platform since we are going to do sentiment analysis leveraging the twitter API. This methodology is applied on only one social media site with the assumption that it might highlight its applicability and the success rate of the same technique to other social sites.

Most of the Kenyan enterprises are using social media as part of the social business objective to promote their brands and to increase their customer base.

Most of the Kenyan enterprises have not implemented any local cyber intelligence techniques of monitoring the social media, for instance, sentiment analysis and opinion mining.

1.7 Significance of the study

This project is necessary because it will help enterprises in reducing risks that arise from data leakages, cyberbullying, and social media squatting and social engineering effects of social media and maintain their privacy against the malicious online users. In addition, these institutions will be able to account for the existing employee productivity hours by putting up measures to control cyberloafing during work hours.

The project will also emancipate the enterprises on the need for establishing effective social media strategy which includes policies, risk assessment guidelines before the uptake of any of these social media tools and implementing local enterprise intelligent systems that can monitor in real time social media use within these enterprises.

1.8 Limitations of the study

The system is limited to the twitter platform currently thus doesn't cover other social media sites.

Most employees may not be willing to give up their privacy to participate as they have a freedom of expression and privacy on the web.

Chapter 2

2.0 Literature Review

2.1 Research literature

2.1.1 Social Media

Isaca, (2016) describes social media as a technology that involves the creation and dissemination of content through social networks using the Internet. Social media is (Chi, 2016), the internet and mobile technology based channels of communication in which people share content with each other. A Cyveillance White Paper, (2015) terms it as an unprecedented phenomenon that has opened new worlds of opportunity for organizations around the globe and that while the potential and rewards are seemingly limitless, so are the challenges and risks. Organizations are faced with a hard time establishing and enforcing effective social media strategies. Some organizations afraid of the dangers of exposure, prohibit social media use. On the other hand, companies not wishing to be left behind, have leveraged E.S.N.Ts without developing effective social media use policies or conducting a risk assessment. Nonetheless, both alternatives are unsatisfactory. Enterprises that do not embrace social media fail to realize its significant benefits and are at a disadvantage to their opponents that do.

In 2009, the U.S. military considered a near-total ban on social media sites throughout the Department of Defense. Military officials cited inherent technical security weaknesses and lack of security safeguards on social media sites (Schachtman, 2009).

2.1.2 Social Media Monitoring

Social media monitoring tools are designed to gather facts about consumer actions, purchase, and attitudes without having to survey customers. Companies can gather information about customer preferences by observing how they talk about companies on platforms such as Facebook and Twitter. Social media monitoring give's companies much more insight into how their products fare in the marketplace and how to be successful with customers and prospects. In the same respect, we can leverage the use of these monitoring tools to monitor employees and customer behavior of narcissism and malevolence and be able to arrest comments that may result into reputation damage to the companies. Most industries and companies are more focused on the market share rather than also taking proactive measures to mitigate security risk. The concept of thought is that having the right security processes will help a company realize benefits in the long term and as a result can save the company from financial losses and reputational damage.

A report by Price Water Coopers in 2013 on how organizations detected their most significant breach (pwc, 2013) consisted of the following. Routine internal security monitoring detected 42% of the worst breaches, while 30% were obvious from the business impact (e.g. system outage, assets lost). 9% of organizations worst security incidents were discovered by accident up to 6% in 2012. This really justifies the need for continuous monitoring as a factor key in mitigating the risks.

Users often appear to be unaware of the fact that their data are being processed for various reasons, such as consumer behavior analysis, personalized advertisement, opinion mining or profiling that's according to Gritzalis, Kandias Stavrou & Mitrou. In order to raise awareness, researchers have conducted attacks on realistic environments, consisting of Social Media communities or groups. (Gritzalis; Kandias; Stavrou; Mitrou; 2012). For this reasons social media monitoring will play a key role in enhancing social media security. They also suggest that, the rise of social media usage has influenced researchers towards opinion mining and sentiment analysis, which constitute computational techniques in social computing. As presented by King etal, social computing is a computing paradigm that involves multi-disciplinary approach in analysing and modelling social behaviour on different media and platforms to produce intelligence and interactive platform results

2.1.3 Social Media Strategy

Social media strategy is fundamental for every company to realize the benefits of social media. Cyveillance White Paper, (2015) indicates that as social media continues to grow and become a part of daily life for millions of people, organizations must take responsibility for developing a proactive social media strategy that protects all facets of the business while taking advantage of the opportunities it offers to engage with customers. An effective social media presence boiled down to its very core, is made up of four parts: A strategy, Monitoring and refining, Active social

presence and passive social presence (Gray, n.d.). It is important to have a plan, so that an enterprise can know how they are going to measure the success of these tools, besides putting into consideration what happens when something goes wrong. Hence there is every reason for every Kenyan enterprise to have an objective in mind when taking up these social media tools (Gray, n.d.).

According to Michael Cross the author of the book Social media security. The primary purpose of a social media security strategy is to give people the ability to do what's needed without compromising security. In creating one, you need to identify what areas need to be secure, how security will be achieved, and who will be responsible. The strategy should encompass any areas related to using social media, inclusive to the corporate workstations people may use, mobile devices issued to employees, network security, and firewall restrictions (Cross, n.d.).

2.1.4 Risk, Privacy Concerns, and Security Issues

Not wishing to be left behind, many enterprises are seeking to leverage social media tools. Since the tools are new to many enterprises and do not require new infrastructure, social media technologies may be introduced to the enterprise by business and marketing teams without IT involvement, a project plan or risk assessment (Isaca, 2015). It is, therefore, important that the enterprise creates a proper plan to address the risks that accompany the technology.

According to Cross, (n.d.), the trade-off of security applies to almost anything you can think of in technology, accounts, network access, equipment, and content. He continues to say that there are many threats on the Internet and many tactics, settings, and tools to protect you and your systems. According to the Cisco 2013 Annual Security Report, the highest concentration of online security threats are on mass audience sites, including social media In addition to giving anyone the power to disseminate commercially sensitive information, social media also gives the same power to spread false information, which can be just as damaging.

2.1.4.1 Insufficient Authentication Controls

In lots of social media applications, sensitive information is spread among many different locations. This makes it more likely that an inexperienced user will introduce a weakness that will adversely affect the entire system. For example, there might be some administrative accounts for which the correct security controls are not in place, such as sufficiently strong passwords. An attacker could use a brute-force attack to determine the password of one account; if other accounts are connected to it through a single-sign-on arrangement, the attacker would then have administrative access to a number of systems.

2.1.4.2 Cyber loafing

According to pwc. (2013) most staff-related incidents normally involve staff misuse of the Internet or email. This happens in more than three-quarters of large organizations and around two-fifths of small businesses. Cyber loafing is the wastage of time in unnecessary web browsing and social media sites.

2.1.4.3 Data leakage

Kenya cybersecurity report 2015 by SERIANU, indicates that with many companies in Kenya allowing their employees to take their devices to work, hackers are taking advantage of poor security on the gadgets to access sensitive company information. “With the continued adoption of enterprise mobility, a growing percentage of workers are using their personal devices to access corporate resources,” states the report. “When these devices are not secured this introduces a wide range of security threats.”

2.1.4.4 Social media squatting

Cyber-squatting is becoming more common each and every day in Kenya. Both celebrities and Government officials alike are being targeted by cyber squatters. Enterprises are also not left behind in this, we got people who masquerade as genuine company accounts on Facebook, Twitter, and LinkedIn. Business owners should therefore register their domain name immediately they start operations to avoid cyber squatters from registering the name under their details as they wait for the business owner to approach them so they can demand payments (Cybersquatters-hit-e-commerce, 2016).

2.1.4.5 Phishing

Though phishing is not exclusive to social media, there has been a current spike in phishing attacks linked with social media sites (Fisher, 2011). Many people view social media sites on cell phones or other mobile devices. Which makes it harder to distinguish actual and bogus web sites. Additionally, social media enables attackers to send phishing messages that appear to come from someone that the victim knows. Having obtained login information for a few accounts, scammers will then send out messages to everyone connected to the compromised accounts, often with an enticing subject line that suggests familiarity with the victims (Baker, 2009)

2.1.4.6 Information Leakage

With the advent of “always-on” connectivity, the traditional lines between work and personal life are becoming blurred. Particularly, younger workers use the same technologies in the office as at home. Additionally, social media sites like Facebook and Twitter create the illusion of familiarity and intimacy on the Internet. The result is that people may be inclined to share information on the Internet that their employer would have preferred to keep private. Individuals may not be divulging trade secrets, but the cumulative effect of small, seemingly innocuous details can enable a business's competitors to gain valuable intelligence about that company’s business situation and future plans.

2.2 System Overview

At the beginning of the project, it was suitable to provide a methodology that fitted well to the enterprise. A review of the available literature on online social media monitoring and opinion mining was therefore used to better understand the methodologies used in preceding projects as regards the sentiment analysis for social media monitoring.

Umati project incubated from a popular innovation hub in Kenya called I-hub, developed a system for election monitoring, for the identification and detection of hate speech. The methodology that was thought ideal for Umati was one that considered the dynamic and unique characteristics of the Kenyan online space, e.g. the multiple languages spoken online, the need for local monitors who understood not only the vernacular languages but the ethnically divided politics in Kenya, the lack of a workable definition of hate speech that suited the Kenyan context and budgetary limitations. The Umati project having provided a way to collect and study hate speech incidents from the Kenyan cyberspace, is however only limited to the election periods which renders it inactive for a longer period before another election. On the contrary our project focuses on the local enterprises all year round whereas the Umati project was a nationwide project for the entire country at election periods, the other difference is that we are analyzing brand hate messages while the Umati project focused on hate speech.

The Panopticon project analyzed how a multifaceted information shared on social media can be used in order to achieve a dual purpose of: (a) Dealing with the insider threat prediction and prevention, as malevolent insiders and predisposition towards computer crime being closely related to the personality trait of narcissism. They proposed a method of outlier detection in social media via influence, usage intensity, and Klout scores evaluation in order to detect users with narcissistic behaviors. They also proposed a method for group analysis under the prism of group homogeneity, being an important characteristic to prevent the manifestation of insider threats. The Panopticon project was the most suitable to replicate to the Kenyan enterprise by analyzing insider and outsider threat of conceit. Though they focused only on the insider threats our project also focusses on the outsider threat as well and how the self-righteous activities of both employees and customers might affect the enterprise leading to such factors as flash mobs, boycotts, and reputation damage.

Ijarie project done in Germany used both sentiment analysis and image analysis to help determine and detect cyberbullying in the cyberspace. They defined cyberbullying as an attack that depends on frightening, intentionally insulting, awkward or annoying people via mobile phones or on the internet over social networking websites, instant messaging and emails application. The image analysis involved the extraction of meaningful data from images by means of digital image processing techniques, this entailed procedures such as reading bar coded tags by implementing the skin color detection algorithm to detect images. Text analysis, on the other hand, was characterized by using preprocessing techniques such as stop word removal and stemming. Bag of words model being the primary stage in Natural Language Processing was mainly used in sentiment analysis.

According to The Register, researchers from the University of Cardiff have been awarded more than $800,000 by the US Department of Justice to develop a pre-crime detection system that uses social media. The system relies on drawing on big data from social media to identify on potential crimes before they happen. The project builds on existing work conducted by

Professor Matthew Williams, Director of the Social Data Science Lab at the Data Innovation Research Institute and Professor of Criminology at Cardiff University and Dr. Peter Burnap, senior lecturer in Computer Science & Informatics, to look for signatures of crime and disorder in open source communications, not crimes themselves.

The new project is meant to collect Twitter posts containing terms that already have been labeled as hate speech by human annotators over a period of 12 months. These two measures will then be entered into statistical models to identify if there is a correlation, that is, whether an increase in hate speech in a given area is also statistically linked to an increase in recorded hate crimes on the streets. The project is meant to use the same similar machine learning techniques to build new hate speech algorithms based on the US data. If the project succeeds, then social media data may be used in conjunction with conventional data sources to improve predictions of hate crimes offline. These new forms of data are also attractive as they can provide new information on changing risks in near real time, unlike conventional data that is often weeks or months out of dates.

Istats a company in Kenya that does social media monitoring solution that tracks activities of Kenyan brands on twitter. It gives a report about a measure of brand conversation with the aim of outlining the most talked about brands on twitter and by extension the most visible brands. They gave a report on data based on 1.13M tweets spanning across 95 brands. With categories including Media, Energy, FMCG's, Motoring, Insurance, Retail, Service, and Telcos. The data was compiled to give a comparative report on how brands are performing on twitter in terms of conversation – driven by the popularity of twitter locally. The metrics used in the project were: Number of posts mentioning a brand, number of unique people mentioning a brand and the average following of the sum of users mentioning the brand to estimate reach.

They also analyzed the brands with the most positive sentiments. In their report Durex achieved the highest score in the positive sentiments, the sentiment score checked for key words and phrases that depict the tone of a post. The post is then scored between 0.0 and 1.0 where 1 is fully positive and 0 is negative. However, the tone was contextualized since they sometimes are not directed on the brand.

The limitations of these is that for a business to thrive the social business strategy is key to promoting their brands and reaching out to potential customers, but with the rising cyber-crime cases, where malicious persons are now leveraging these social media tools to cause exposure to enterprises, we can still use social media monitoring towards analyzing and modelling social behavior that may be able to provide cyber threat intelligence enabling organizations to protect their information infrastructure and employees from physical and online threats found within and outside the network perimeter.
After the review of the literature in the fore mentioned areas, we did propose a system design framework below for the implementation of the project:

The system can be accessed from either a laptop or a desktop computer, then leveraging the use of the twitter API, with python code we can access tweets in real-time. Once we access the tweets we do sentiment analysis on the stream of tweets by tracking a given brand name In this case we considered a Safaricom as the case study of the project. Safaricom being a brand with large following, due to its position the market and the large customer base, it attracts a number of mentions on a daily basis. Once we analyze the tweets, each tweet is thus classified as being either positive or negative and assigned a polarity score. In making our system more intelligent, we have a watchdog application that monitors for tweets with a higher polarity score more than 0.8. These tweets are then sent via email to the individual responsible for social media monitory in a company or any relevant personnel with that duty, these tweets can then be analyzed to determine which ones are more sensitive to the company, and which ones would result into damage of company or brand reputation. After which a decision can be made whether to arrest or flag the post before it creates an uproar in the social media. Alternatively, a decision can be made to send cease and desist letter to the individuals. This data is also useful in profiling this users through opinion mining.

The system implementation thus has the goal of providing the company with information they need to address quickly and protect their businesses and reputation. This is because most industries and companies are more focused on the market share rather than, also taking proactive measures to mitigate security risks. The concept of thought is that having the right security processes will help a company realize benefits in the long term and as a result can save them from financial losses and reputational damage.

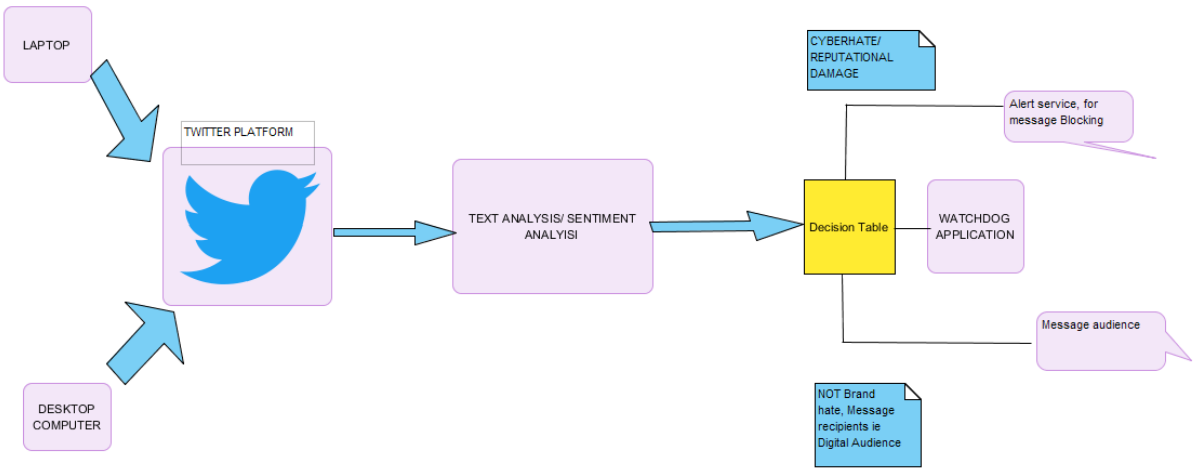


Fig 1. Adapted from Ijariie-Issn (O)-2395-4396

Chapter 3

3.0 Methodology

3.1Research Methodology

Saunders et al. (2007) describes research *to be something that people carry out in order to find out find out things in a logical way, thereby increasing their understanding* (Sunders et al.2007,5). That "something" is the purpose why this study was carried out to increase our understanding on the area.

As a matter of fact, it is imperative that a sequential process be followed for a research to be rigorous enough. Research procedures are a sequence of stages should be followed for a research. Even though these steps have been ordained to follow each other sequentially, Saunders et al.(2007) disputes that saying that it might not be the case sometimes(Saunders et al.2007,8). Alternatively researchers might decide to take an approach that is completely dissimilar for the research.

This section focuses on the research methodology undertaken, with justification aas to why we based our research on the particular methodology and not the others. The topic "Mitigating Uncensored Social Media Utility" was arrived at after a brainstorm of what to research on the area of social media.

3.1.1 Purpose for Research

To carry out a research an objective must be defined, these explains the reason researchers are three dimensional; explanatory, descriptive and exploratory. The difference only exists in the way the research questions are formulated (Saunders, et al.2007:133) The three are different in the following ways.

Descriptive Study

This are more structured research that defines a problem phenomenon with a specified hypothesis. According to Blumberg et al (2008), descriptive research are always more complicated requiring astute skills on research to be considered and implemented successfully.

Explanatory Study

This research tries to vilify and establish the reason for the existence of relationships of a phenomena or situation that appears more than one aspect. (Kumar 2011: 11 and Saunders et al. 2007, 134). This study attempts to reason out, why things happen the way they do.

Exploratory Study

For the research bit of this section, exploratory study was used, Exploratory research is implemented to establish what is happening and to gather new insightful conclusions, and

knowledge and ask questions that can help in a establishing a problem in a different perspective (Robson 2000 as cited by Saunders et al. 2007, 133). Churchill and Lacobucci (2009) says that it is normally done not necessarily to test an explanation but rather to gain insight. Saunders et.al (2007), describes the study to be conducted in a three different ways that is:

- a review of literature
- interviews for experts on the subject
- Focus group interviews

Research Design and Analysis

Blumberg et al. (2008), argues that even though there are many definitions for research design not a single one shows the full range of its fundamental aspects. It is the blueprint for achieving project goals and finding answers to questions in addition, choosing the appropriate research design is very important as result of the many existing options available. For the purpose of the study, a series of data collection approaches used included gathering relevant material form sources such as: Books, Journals, and Reports, websites, blogs and trends.

3.2 System Methodology

Agile

Agile resulted as a "solution" to the shortcomings of the waterfall methodology. it follows an incremental approach, Instead of the sequential design process. While choosing the methodology we considered the following factors why agile was suitable.

- In Agile rapid production is more important than the quality of the product.
- Clients are able to change the scope of the project.
- It is suitable when there is no clear picture of what the final product should look like.
- It is best for skilled developers who are adaptable and able to think independently.
- Best suited for products intended for an industry with rapidly changing standards.

The twitter sentiment Analysis system uses an agile methodology for the development process. According to ken Schwaber (Schwaber K, 2003) one of the initiators of the agile scrum method, agile is a process for managing complex projects. He puts emphasis on the fact that the methodology is not just limited to software development, but given the tendency for software development processes to be very complex, agile is well suited for the managing them (Brooks, 1978).

3.1 Overview

The scrum method is incremental, with each increment called a sprint, each sprint is recommended to last for 4 weeks. Before the sprint there is a planning meeting for each sprint,

where a customer decides which features should be implemented in the upcoming sprint. During the sprints the teams meets on a daily basis on short meetings called scrum. A sprint review meeting is held at the end of each sprint, and the customer is able to see the existing accomplishments for the preceding sprint. Teams can also hold sprint retrospective meetings to, where they can look at the process and then try to find out what went right and what can be improved. The figure bellow shows the flow of the methods.

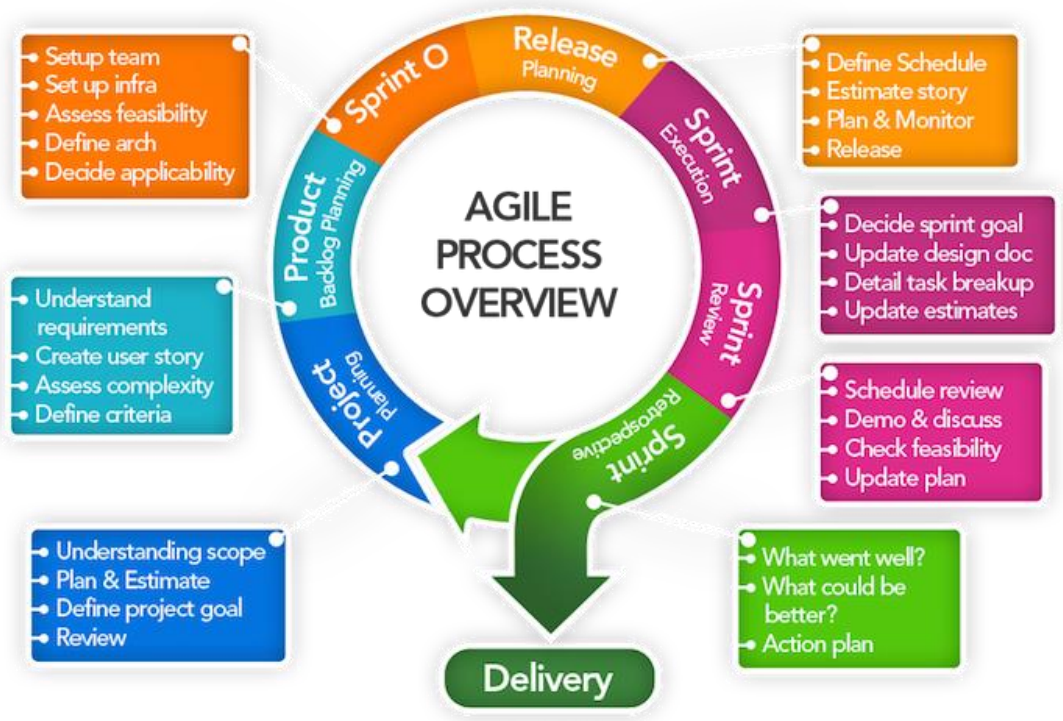


Fig 1. Overview of the scrum method

3.2 Scrum Artifacts

Scrum being an agile method, follows that the formality of the project is as low as possible, it thus facilitated necessary changes to be made to the project. The customer is also able to see the project progress since this improves their motivation and involvement. Artifacts of scrum includes:

1. The product backlog
2. The sprint backlog

3. The sprint burndown chart.
4. The impediment lists

3.2.1 The Product Backlog.

It can be considered equivalent to the requirement specifications, but it has one big difference in that it does not entail long description of each requirement, it has only single sentence description for each requirement. Being a list of such single sentence requirements, it is thus the customer's priority to keep it prioritized and updated. The customer can add requirements to the list and the team is then responsible for providing estimates of how long the implementation might take.

3.2.2 The Sprint Backlog

It is a list of tasks maintained and compiled by the team based on he the items in the product backlog, initially selected to be part of the sprint. The list is similar to the product backlog, but with a big difference. The items on the product backlog are features requested by the user, the sprint backlog is a list of tasks the developers must do to implement the items that the customer chose from the product backlog. The customer doesn't need to know about the items on the sprint backlog.

A general rule of thumb is that the tasks on the sprint backlog should always be relatively short, that is between one hour and two days.

3.2.3 The Sprint Burn down Chart

This measures the progress of the sprint instead of the project. It is important because it helps the team discover tasks they did not consider but that must be added to the sprint backlog. Since the chart displays the amount of work remaining and not the amount of work completed, the graph can in fact increase from one day to the next.

3.2.4 Impediment list.

An impediment is anything holding back development in some way or another. It is the scrum master's responsibility to deal with any such impediments. The list is simply a set of tasks that the scrum master uses to track the impediments that needs to be solved.

3.4 The Sprint Planning Meeting.

In the first session, the Customer chooses high priority items from the product backlog that should be completed in the upcoming Sprint. The customer explains the items to the team and they give an estimate on how long it will take to complete it. The sprint backlog is filled so that

the sum of the item estimates is about the same as the available work time of the team during the upcoming sprint.

3.5 The Daily Activities.

During the sprint, the developers work on the items in the sprint backlog. Every day the developers synchronize their progress in a daily Scrum meeting that should last no longer than 15 minutes. During the meeting, all the developers will tell the others what they did since the last Scrum, if there are any impediments obstructing their work and what they are planning on doing until the next Scrum. Another important day to day activity is updating the sprint backlog and burndown chart.

3.6 Sprint Review Meeting

At the end of the sprint, the team meets with the customer and presents the result of the sprint. The users demonstrate the functionality they have completed and gets feedback from the customer. If the demonstrated functionality is what the customer wanted, then this gives the team a feeling of accomplishment as well as the customer a proof that the project is moving in the right direction. If the demonstrated functionality isn't quite what the customer was looking for it is now easy to explain how it is different and what should be done next. In some cases, it is enough to make a few changes while in other cases the implemented functionality must be discarded.

3.6. Sprint Retrospect Meeting

The intention of this meeting is to help the team improve their development process. The meeting is attended by the team, the scrum master and the customer (optional). During the meeting the team members take turns saying what went well during the last sprint, and what could be improved. After all team members have had their say, they prioritize the possible improvements and discuss them in order. The meeting should not last more than 3 hours.

3.7 Project Startup

Ken Schwaber has had much success with his kick-starting of Scrum projects as described in the book Agile Project Management with Scrum (Schwaber K. M., 2002). This process goes as follows.

The Scrum Master works with the customer and prepares a backlog. Then the Scrum Master, the Customer and the Team uses one day to go over this backlog. During this first day the customer explains the items in the backlog to the team, and the team estimates how much work it would take to implement this. The customer then prioritizes the items in the backlog and divides the backlog items into sprints. The following day is the first day of the first sprint. This first sprint isn't very different from the following sprints, except that the first part of the sprint planning meeting has already been completed.

Chapter 4.

4.1 Research Findings

4.1.1 The existing risks of social media utility across Kenyan enterprises.

4.1.1.1 Cyber loafing

According to pwc. (2013) most staff-related incidents normally involve staff misuse of the Internet or email. This happens in more than three-quarters of large organizations and around two-fifths of small businesses. The average affected company had about one breach a month, though some reported many more. Cyber loafing is the wastage of time in unnecessary web browsing and social media sites. A number of firms such as G4S and Child Welfare Society of Kenya (CWSK) say they have been forced to come up with policies on how employees access social media during office hours. They have also invested in filtering and antivirus tools. Often employees do unproductive activities like chatting, surfing the Internet, downloading videos and music during office hours, activities that are bandwidth-hungry and slow the speed of internet for those using it for business.

4.1.1.2 Data leakage

Kenya cybersecurity report 2015 by SERIANU, indicates that with many companies in Kenya allowing their employees to take their devices to work, hackers are taking advantage of poor security on the gadgets to access sensitive company information. “With the continued adoption of enterprise mobility, a growing percentage of workers are using their personal devices to access corporate resources,” states the report. “When these devices are not secured this introduces a wide range of security threats. “People may be inclined to share information on the Internet that their employer would have preferred to keep private. Individuals may not be divulging trade secrets, but the cumulative effect of small, seemingly innocuous details can enable a business's competitors to gain valuable intelligence about that company’s business situation and future plans.

4.1.1.3 Social media squatting

Cyber-squatting is becoming more common each and every day in Kenya. Both celebrities and Government officials alike are being targeted by cyber squatters. Enterprises are also not left behind in this, we got people who masquerade as genuine company accounts on Facebook, Twitter, and LinkedIn. Business owners should register a domain name immediately they start operations to avoid cyber squatters from registering the name under their details as they wait for the business owner to approach them so they can demand payments (Cybersquatters-hit-e-commerce, 2016).

They are mainly known to target large organizations. Insight Kenya limited, delayed to renew the domain name of one of its clients and in a couple of days, it had been grabbed by another

company. The company is currently in a bidding war to get the name or lose it to other bidders. Independent ICT analyst and registrar Alex Gakuru says there are cyber squatters who have already registered popular names in an Africa, especially Swahili names waiting for locals to use them only to later propose to sell it to them (Cybersquatters-hit-e-commerce, 2016). “The squatters usually follow big brands whose owners are either not keen to renew their addresses or have never bothered to take them online,” Said Mr. Wambugu adding that those cases are rampant in Kenya and South Africa. Declining to give examples of the actors. Examples of Kenyan domains that have been taken down includes Uhuru.co.ke a website, which was launched in 2011, was widely used in the 2013 elections and was used to showcase the President’s bio as well as the Jubilee party manifesto(Eric Wainaina, 2016)

4.1.1.3 Phishing

Though phishing is not exclusive to social media, there has been a current spike in phishing attacks linked with social media sites (Fisher, 2011). Many people view social media sites on cell phones or other mobile devices. Which makes it harder to distinguish actual and bogus web sites. Additionally, social media enables attackers to send phishing messages that appear to come from someone that the victim knows. Having obtained login information for a few accounts, scammers will then send out messages to everyone connected to the compromised accounts, often with an enticing subject line that suggests familiarity with the victims (Baker, 2009).

More than 5,000 Kenyan Facebook users have lost millions of shillings in a social hacking scam that lasted for a year, as revealed by a local cyber security firm, Serianu. The firm said that it unmasked a Kenyan who broke into personal accounts of Facebook users in Nairobi, Mombasa, and Eldoret and used the access to solicit funds from thousands of people linked to the breached accounts.

4.1.1.4 Reputation risk

The truth is, social media is undoubtedly one of the most powerful forms of online marketing but it isn't simply a matter of waving a wand and success "auto-magically" happens. It takes time, focus and, most importantly, a particular skill set developed over time. As a matter of fact, if you do it the wrong way, social media could end up working against you. Take, for example, a case where RMA motors Tweeted to a client to have a cold beer in lieu of service. This was is a response to a complaint from the high profile client to the effect that his brand new Range Rover was always breaking down. The actual Tweet read, "will be in touch to discuss one-to-one your outburst here & what can be done. For today; have a cold tusker & enjoy the day." Within hours, this Tweet had gone viral with hundreds of Kenyans on Twitter expressing their outrage. This was a classic example of how social media can go wrong if you do it wrong or put unskilled or inexperienced people in charge of this channel. ("5 Reasons Kenyan Businesses Should Embrace Social Media", 2016)

4.1.1.5 Cyber bullying

Cyber-bullying is defined as publishing materials about a victim severely defaming and humiliating them. It isn't something that anyone or large corporate would want to be associated with. Ironically in Kenya, many companies and personalities actively take part in it by quickly aligning their digital content to a widely known unrelated trending topic. A case in point is during a wave of the #PoleKwaMwirigi tweets, retweets, Facebook jokes and funny memes on twitter which went viral globally through Kenyans on twitter (KOT).Enterprises that apparently took part in the hashtag included: Airtel Kenya, Cold Stone Kenya, Fast Food Restaurant Kentucky Fried Chicken (KFC) and lastly NTV one of kenya's leading broadcaster also took part in the conversation according, Chacha,(2016).

4.1.2 Mitigations strategy

To effectively control social media use by both employees and enterprises, a well-documented strategy need to be developed, with the input of all the relevant stakeholders. This includes: the business management, the human resource, officials entitled for risk management, and the legal representation. An approach of this perspective by holistically integrating emerging technologies into the business will help to ensure risks are considered, with the view of the broader business objectives. A strategy to address the social media risks, should focus primarily on user behavior, with the development of policies and offering support for training and awareness programs which covers.

- Individual use in the workplace:
    - Is it allowed or not
    - Is it a nondisclosure of business oriented content
    - Is it a discussion of work related topics
    - Inappropriate content and conversations
- Individual use out of the workplace
    - Nondisclosure of business oriented content
    - Ordinary disclaimers for employee identification
    - The hazards of posting a lot of personal information
- Business/Enterprise use
    - Is it allowed
    - Is there a process to gain approval for use
    - What is the scope of information allowed to flow

- What are the disallowed activities
- Consider the escalation process for consumer related issues.

Proper training and education are imperative, vulnerabilities of social media usage should be well apprised to every employee. Organizations can also consider. A standard “Social Media Safety 101” class as a good starting point. Consequently, a compact and all-inclusive social networking policies should be put in place, and enforced through continuous monitoring leveraging the intelligence tools such as sentiment analysis, for monitoring real-time posts. In addition a proactive, continuous monitoring is highly essential for success, hence all organizations should take responsibility by knowing the greatest goal beyond these social media sites.Lastly, business departments must subscribe to a solid organizational feedback loop. This is with regards to the common tendency for departments to point the finger to another department. Businesses should consider the fact that whenever a breach occurs it is more than just a public relations issue, or rather, just a normal security, legal IT, human resources or security issue. Every department has a specific role to play which can either make or break an organizations social media policy.

The figure below shows the risk mitigation techniques for Employee social media use.

Threats and Vulnerabilities	Risks	Risk Mitigation techniques
Employee posting of pictures or photos that link them to the enterprise	- Privacy Violations - Loss of competitive advantage - Reputational damage - Brand Damage	- Social media monitory. - Ensure existing policies address postings of employees. - Develop awareness training and campaigns.
Cyber loafing	-Productivity loss -Increased risk of exposure -Strains on bandwidth	- Social media monitory. - Managing social media accessibility through content filtering.
Employee access to social media through enterprise – supplied mobile devices(PDA's and smartphones)	- Data leakage - Phishing	- Routing enterprise smartphones through corporate network filtering technology to restrict social media usage. - Social media policy - Conduct a rigorous training and awareness campaigns emancipating employees of the risks posed by social media sites.
Social media Squatting	- Reputation Damage - Brand damage	- Social media monitory to analyze fake, unverified company accounts.
Using personal Accounts for work related postings/information	- Loss of competitive advantage - Reputational Damage - Privacy Violations	- Formulate effective social media policies - Update policies regularly - Training and awareness campaigns

		- A standard “Social Media Safety 101” class

4.1.3 Suggested Policy Formulation guideline for Social Networking Sites.

<Company Name>

Social Networking Policy

1. Overview

Social networking is gradually being seen as a fundamental element of work as well as personal life. Whilst industries are simultaneously gaining approval for social media tools as a means for endorsing goods and services, at the same time improving retention of workers, there is always an ever present risk of employee abuse or sensitive information pilferage.

2. Purpose

The drive is to provide a framework for contractors, workers and other personalities carrying out work for <Company Name> , on the acceptable use of the Enterprise social networking tools at work and in personal usage situations.

3. Cancellation or Expiration

This policy document has to be reviewed and updated as required in line with the dynamic nature of these social media tools. Therefore the policy does not particularly have an expiry date

4. Scope

The Social Media Policy applies to all those personalities working on behalf of <Company Name> regardless of whether they are part time or fulltime employees, on contract, casual workers, business partners, temporary agency workers and vendors .

5. Policy

5.1. Speaking on Behalf of <Company Name>

Specific individuals doing work on behalf of <Company Name> will, in lieu of their position, be familiar about particular aspects of <Company Name> and for that may be legalized to talk on the behalf of <Company Name>

- One must not express his/her views on behalf of <Company Name> unless (the

person) is influential on the matter And has been legalized, by the book, to communicate on behalf of < Company Name> by the manager or liable <Company Name> executive.

- You must not give out information confidential or proprietary. Only public available
- Information or information which you have been authorized to share may be disseminated.

- Be transparent. Clearly identify yourself, that you work for <Company Name>, and what your Role is.
- Be professional. This includes being honest , respectful and factual at all times.
- Do not refer to the Products or services of vendors, client's customers or partners without obtaining their consent.

5.2. Personal use of Social Media Activities

It is well known that particular personalities working on behalf of <Company Name>will be active on social media.

If one is discussing products or services provided by <Company Name> , then one obligated to identify themselves as an employee distinctly show that the views are theirs and do not epitomize the views of <Company Name> .

You must not express disapproving statements about <Company Name>, its employees or officers, or any Product or service provided by <Company Name>.

You may not trade or recommend any product or service which would compete with products or services sold by<Company Name>.

When on the job, social media access should be confined to limited personal use.

6. Enforcement

Any individual found to be in violation of this policy may be subject to disciplinary action, up to and including termination of employment or contract and potentially legal action.

7. Definitions

Limited personal use – A philosophy that employees are permitted limited personal use of <Company Name>computing resources when that use does not:

- Interfere with business usage of<Company Name>resources.
- Is performed on non-worktime

4.2 System/Software Development

4.2.1 Project Planning

The first step in the scrum methodology is the Planning phases. Ours was to develop a sentiment analysis system that analyzes twitter sentiments. The Planning phase comprise of the following:

1. Development of a comprehensive backlog list.
   The backlog list contains:
   - Initiate project
   - Requirement analysis and conceptual designs
   - Modelling and Implementing the classifier
   - Closure
     The sprint backlog

	Product Backlog	**Sprint backlog**	
1	Initiate project	Project concept	
		S.Q.C.T targets	
		Formulate W.B.S	
		Project Charter	
2	Execute Project	Requirement Analysis	
		Literature Review	
		Conceptual Framework	
		UML and Use cases	
3	Analyze Sentiments	Obtain a dataset of (+ve) and (–ve) tweets	
		Train classifier	
		Obtain Twitter API	
		Do sentiment analysis on tweets	
		Graph live tweets (+ve/-ve)	
		Send email notification for tweets with high polarity	
		Scrape tweets to a database for analysis	
		Visualize most tweeting accounts in a word cloud	
	Closure	Final Project Presentation	
		Submission of project	

2. Defined delivery date and functionality of immediate releases.
   The delivery date was scheduled to be on 2nd November 2016

3. Selection of the release most appropriate for immediate development.

The release most immediate for the development would be the Beta version of the sentiment analysis system.

4. Definition of project team(s) for the building of the new release.

Scrum Master	Dr. Liyala	The project supervisor -Leads the team -Facilitates and coordinates -Helps removing the obstacles -Safeguards the process
Project manager(Owner)	Collins Bunde	- Defines initial content and Timing of the release - Manages evolution of project content - Deals with Backlog, risk and release content.
Development team	Edward Onyango and Collins Bunde	Edward – Documenter Collins – Lead Developer

5. Risk assessment and risk controls.
6. Reviewing and providing possible adjustment of backlog items and packets.
7. Development tools validation and infrastructure.
8. Verifies management approval and funding

4.2.2 Architecture/High Level Design

1. Review of assigned backlog items.
2. Identify immediate changes for implementing backlog items.
3. Conduct domain analysis to a level required to build, enhance, or update the domain models to reflect the new system context and requirements.
4. Harmonize architecture of the system to support the new requirements and context.
5. Identification of issues in implementing the changes
6. Design review meeting, each team presenting approach and changes to implement each backlog item. Reassign changes as required.

Technical – We conducted a technical feasibility to know whether we have all the technical skills required for the project. What we considered are the skills in machine learning and python

programming, we had the necessary resources and infrastructure to support up to only the implementation bit.

1) Economic –the project work was facilitated despite, with financial constraints, this entailed using economically the resources available for the implementation of the system.
2) Operational – the operation feasibility of the system, gave consideration to using any enterprise in Kenya as a case study for the project.

The possible stakeholders of the project are:

1) Companies and Enterprises
2) Twitter Company -It has rate limited use to its API (must agree to terms and conditions) 3) Programmers
4) Project supervisor

This type of project is explicitly conceptualized on the basis of **a proof of concept.**

4.2.2.1. Preconditions

Preconditions form the context within which the project must be conducted. This includes the legislation, working condition regulations, and approval requirements. Such kind of requirements is not influenced from within the project. Some of the existing preconditions for this project are:

1. The twitter Streaming API is rate limited to a certain number of requests per day, hence this should be adhered to or else, twitter will shut one (the twitter collector) out.
2. The enterprises, who would be willing to take such a system will have to fit it within their social media policy guideline, measures as to what an employee should or should not post and who are responsible for the social media responsibility in the company should be distinctly defined. Finally, how these persons are held accountable should well defined.
3. There might be privacy issues that arise from monitoring what others are posting hence there is a need for harmonization of this issue in the best way possible. This is because according to the constitution every individual has freedom of expression.

4.2.2.2 Design limitations

1. The design limitations of the system are that it only analyzes text messages, hopefully, in future, it can be advanced by doing image analysis of photos and images posted on most of this social media sites.
2. In addition, the implementation is only limited to one social media site, when most enterprises are using more than one social media tool such as Facebook, Instagram, WhatsApp and Linkedin in their social business strategy.
3. The classifier was modeled with no consideration to the neutral tweets, this should with time be considered for future advancement. Neutral tweets are tweets that are neither positive nor negative.

4.1.3 Analysis and Requirements

In analysis and requirements we looked at the following important areas that characterized the functionality of the system

1. Machine Learning

Machine learning is the scheme and study of software items that make future decisions from past experiences; it is the study of programs that learn from data. The primary goal of machine learning is to prompt an unknown rule from examples of the rule's application. The undisputed model of machine learning is spam filtering. Spam filters learn to classify new messages, by observing numerous emails previously considered as either spam or ham (Hackeling, 2014).

Why machine learning?

Factor such as growing sizes and varieties of accessible data, and the cheaper computational processing that is more powerful, and affordable data storage. This facilitates the quick and automatic production of models that can analyze bigger, more multifaceted data and deliver faster and more accurate results – And by building precise models, an organization has a better chance of identifying profitable opportunities – or even avoiding unknown risks.

Categories of machine Learning

1. Supervised Learning - A supervised learning program learns from labeled examples of the outputs that should be produced for an input. They make estimations using data.

Fig 4.1 showing an example of supervised learning for predicting whether an email is a spam

2. Unsupervised Learning – In unsupervised learning, a program attempts to discover patterns in the data rather learning from labeled data.

The supervised Learning Model

1. The first step is to train a machine learning model using labeled data.

Labeled data is normally labeled with the outcome, the machine learning model then learns the relationship between the attributes of the data and its outcome.

2. The second step is to make predictions on new data for which the label is unknown

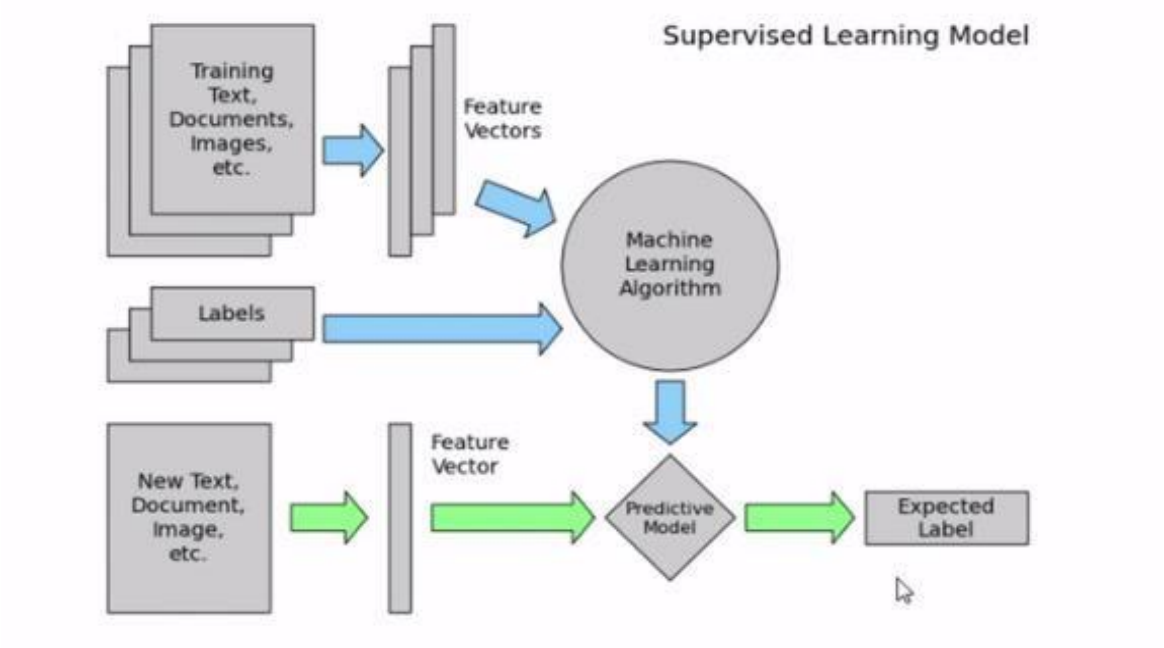


Fig 4.2 the supervised learning model

The primary goal of machine learning is to build a model that generalizes, It should accurately predict the future rather than the present.

Training data and test data

In training observations comprises the particular capability the algorithm is using to learn. Supervised machine learning complications, is where each observation consists of an observed response variable and one or more observed explanatory variables.

The test set is an analogous collection of interpretations used to evaluate the performance of the model by the use of some performance metric. Importantly no observations from the training set should be incorporated in the test set. A test set that contains samples from the training set, is difficult to evaluate whether the algorithm has learned to generalize from the training set or has simply memorized it. To effectively perform a task with new data, a program should be able to generalize well. Of great importance to note is that the predictive power of many machine learning algorithms improves as the amount of training data increases**.**

Our project uses supervised learning where we obtain a data set for both positive and negative sentiments and train the classifier to the data, after which we test the data set on a sample data before we eventually use it to classify twitter sentiments.

Performance measures, bias, and variance

To measure whether or not a program is learning to perform its duty more efficiently a number of metrics can be used. Many performance metrics measure the number of prediction errors for supervised learning problems. The two fundamental causes of error are bias and variance.

A high biased model produces related errors for an input irrespective of the training set it was trained with; the model biases its personal assumptions about the real relationship over the one demonstrated in the training data. A high variant model, conversely, produces different errors for an input depending on the training set it was trained with. While a model with high bias is inflexible, a model with high variance can be quite flexible that it models the noise in the training set.

2. Natural Language Processing

Python toolkit and a fundamental arsenal for mining the social web. The Natural Language Toolkit (NLTK) is a suite of Python libraries designed to identify and tag parts of speech found in natural English text. Its development began in 2000, and over the past 15 years, dozens of developers around the world have contributed to the project. Although the functionality it provides is tremendous (entire books are devoted to NLTK).

NLP is inherently complex and difficult to do even reasonably well, and completely nailing it for a large set of commonly spoken languages may well be the problem of the century. However, despite what many believe, it's far from being a solved problem, and it is already the case that we are starting to see a rising interest in a "deeper understanding" of the Web with initiatives such as Google's Knowledge Graph, which is being promoted as "the future of search." After all, a complete mastery of NLP is essentially a plausible strategy for acing the Turing Test, and to the most careful observer, a computer program that achieves this level of "understanding" demonstrates an uncanny amount of human-like intelligence even if it is a brain that's mathematically modeled in software as opposed to a biological one.

How the classifier works
A classifier distinguishes between good and bad words. It has the following techniques:

Uni-grams

Keep track of consecutive sequences of words, the longer sequences of words are called n-grams.

Stemming

Stemming is getting rid of prefixes and suffixes in a word, it's an algorithm that takes words and strips out its suffixes and prefixes. Example of stemming applies in the following:

1. Watching, watched -> watch
2. Liked, liking -> lik
3. Cats, catlike, catty -> cat

Stop words

They help in preprocessing data by analyzing proper text, stop words are words that are typically pulled out since they don't have much meaning. They are referred to as filler words

Fig 4.3 showing example of stop words

WordNet

The linguistic knowledge that tells one what are adjectives, and word synonyms, nltk enables one to leverage this language.

Part of Speech Tagging

A common preprocessing technique divides words into bigrams, trigrams or unigrams

It decomposes words into verbs adverbs adjectives nouns and pronouns**.** All nltk libraries have this functionality

3. Scikit Learn

One of the most popular open source machine learning libraries for Python is Scikit-learn. It provides algorithms for machine learning tasks such as classification, reduction, regression, dimensionality and clustering. Furthermore it provides modules for features extraction, data processing, and models evaluation. Built on the popular Python libraries NumPy and matplotlib, scikit-learn is popular for academic research due to its well-documented, easy-to-use, and adaptable API. Developers can use it to experiment with different algorithms by altering only a few lines of the code. It also wraps some popular implementations of machine learning algorithms, such as LIBSVM and LIBLINEAR. Other Python libraries, including NLTK, have wrappers for scikit-learn. It also includes an assortment of datasets, allowing developers to put emphasis on algorithms rather than finding and cleaning data.

4. Twitter Platform

Twitter can be defined as a real-time, vastly social microblogging facility that lets users to post precise status updates (tweets) that appear on timelines. Tweets includes one or more entities in their 140 letterings of content and reference, one or more places mapping to locations in the real world. An understanding of users, tweets, and timelines is predominantly vital to effective use of Twitter's API. Tweet entities comprises the user mentions, hashtags, URLs, and media be associated with a tweet, while places are locations in the real world. To make it all a bit more concrete, let's consider a sample tweet with the following text:

@KTNNews @Hassanjumaa @SMukangai @abullerahmed . Safaricom wanatuibia

The tweet is 83 characters long and contains two tweet entities: the user mentions @KTNNews @Hassanjumaa @SMukangai and @abullerahmed the text "Safaricom wananiibia." An API is largely abstract in that it specifies an interface and controls the behavior of the objects specified in that interface. The software that provides the functionality described by an API is said to be "an implementation of the API". An API is typically defined in terms of the programming language used to build an application (Russel, 2013).

Twitter uses the REST (Representation State Transfer Protocol), which is resource focused, and remote resources can be created, read, updated and deleted. The Twitter API requires a key and hence it is quite secure. However, it is no longer free. This API's generally provide very valuable information. The data providers, limit the number of requests per day, demand an API key or even charge for the use.

5. Software requirements

- Python programming Language

  Python programming language used because of its intuitive syntax, the amazing ecosystem of packages that trivializes API access and data manipulation, and core data structures that are practically json, make it an excellent tool that's powerful yet also very easy to get up and running.

- Ipython Notebook

  It's a powerful, interactive Python interpreter that provides a notebook-like user experience from within your web browser and combines code execution, code output, text, mathematical typesetting, plots, and more. It's difficult to imagine a better user experience for a learning environment because it trivializes the problem of delivering sample code that the reader can follow along with and execute with no hassles.

- Anaconda

A freemium open source distribution of the Python and R programming languages for large-scale data processing, predictive analytics, and scientific computing, that aims to simplify package management and deployment. Its package management system is conda.

6. Algorithms used

1. Multinomial Naive Bayes

The multinomial naive Bayes model is typically used for discrete counts. With a text classification problem, it takes the idea of Bernoulli trials one step further and instead of "word occurs in the document" we have "count how often word occurs in the document", you can think of it as "number of times outcome number x_i is observed over the n trials"

2. Naïve Bayes Algorithm

This Classification is named after Thomas Bayes (1702-1761), who proposed the Bayes

Theorem. Bayesian classification provides practical learning algorithms and prior knowledge. Some of the uses of the Naïve Bayes Classifier are:

1. Naive Bayes text classification. Naive Bayes classifiers are among the most successful known algorithms for learning to classify text documents.
2. Spam filtering is the best-known use of Naive Bayesian text classification. It makes use of a Naive Bayes classifier to identify spam e-mail. It has become a popular mechanism to distinguish illegitimate spam email from legitimate email

.

3. Gaussian Naive Bayes Algorithm

Assumes that the features follow a normal distribution. Instead of discrete counts, we have continuous features (e.g., the popular Iris dataset where the features are sepal width, petal width, sepal length, petal length).

4. Support Vector Machines algorithm

In machine learning, support vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to one or the other of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the other, making it a non-probabilistic binary linear classifier. SVMs can be used to solve various real-world problems:

• SVMs are helpful in text and hypertext categorization as their application can significantly reduce the need for labeled training instances in the standard inductive settings.

5. Logistic regression

It is the appropriate regression analysis to conduct when the dependent variable is dichotomous (binary). In this case, the output of the classifier is binary data that is used later on the actual tweets. The logistic regression is a predictive analysis. Logistic regression is used to describe data and to explain the relationship between one dependent binary variable and one or more metric (interval or ratio scale) independent variables.

We used more than one algorithm since two algorithms are not equivalent and will not necessarily produce the same accuracy given the same data. Since the results for each method/ classifiers are significantly different

4.3.1 Design and Development

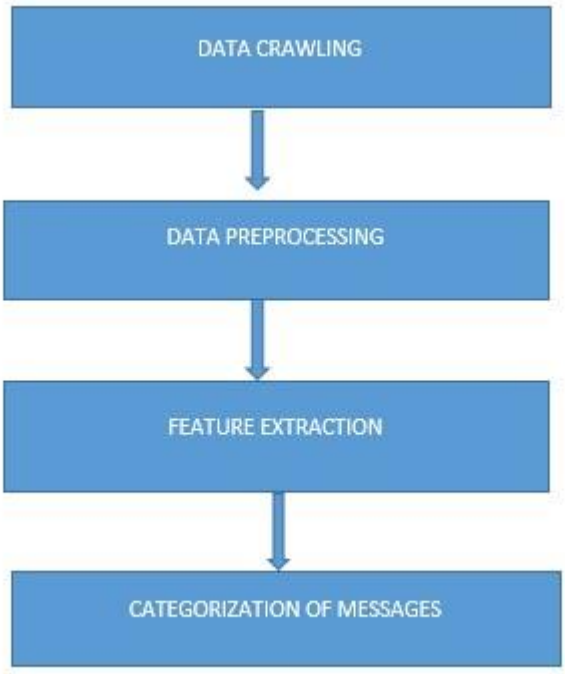
4.3.1.1 System Requirements

The following are the system requirements for the system:

REQUIREMENTS	SPECIFICATIONS
Laptop/Desktop computer	- 64 bit - 6GB RAM - Windows 8,10 - Linux
Python	Python 3.5 or 2
Anaconda (for data science)	64 bit version
database	MySQL or SQLite
Mail server	Gmail, Yahoo or any other
Twitter Account	For scraping and streaming tweets

4.3.1.2 System Use Cases

During this phase, functional, support and training requirements are translated into preliminary and detailed designs. Decisions are made to address how the system will meet functional requirements. A preliminary (general) system design, emphasizing the functional features of the system, is produced as a high-level guide as the one below.

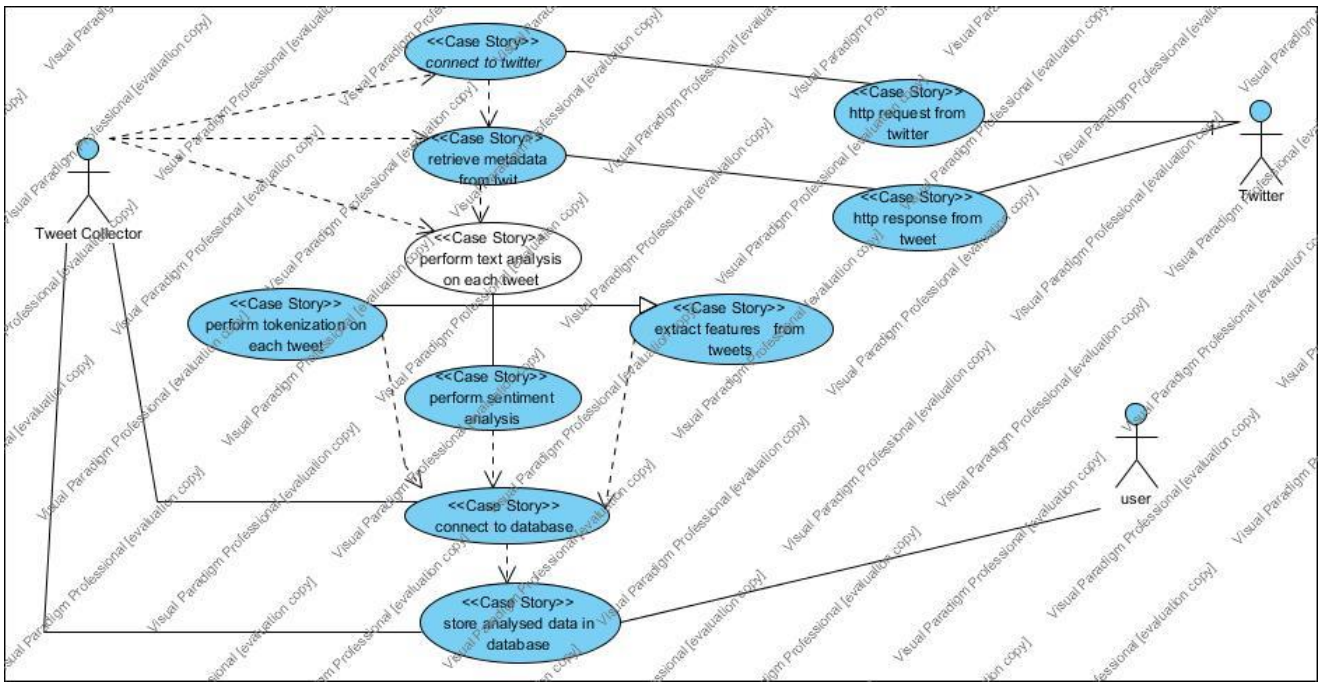


1. Data crawling- In this module, we captured data from Twitter by using our customized crawler written in Python. Data we got from twitter contain sentiments that are either positive or negative.
2. Data pre-processing - Data captured from twitter contains many missing fields, duplicate tweets. Pre-processing of the dataset involve following steps:
    1. Missing fields are replaced by NULL.
    2. Stemming - The idea of stemming is a sort of normalizing method. Many variations of words carry the same meaning, other than when tense is involved. The reason why we stem is to shorten the lookup and normalize sentences
3. Categorization of Messages

In this module, we used a number of machine learning methods, which needed pre-labelled training data for automatic learning: Naive Bayes classifier, a classifier based on Decision trees, Support Vector Machines (SVM) and Multinomial naïve Bayes classifier.

The following diagrams are used to describe the functionality of the Twitter Sentiment Analysis System.

Fig 4.1 sentiment analysis use case diagram



The following use case describes the interaction with the different entities with the twitter API to access tweets and do sentiment analysis on them.

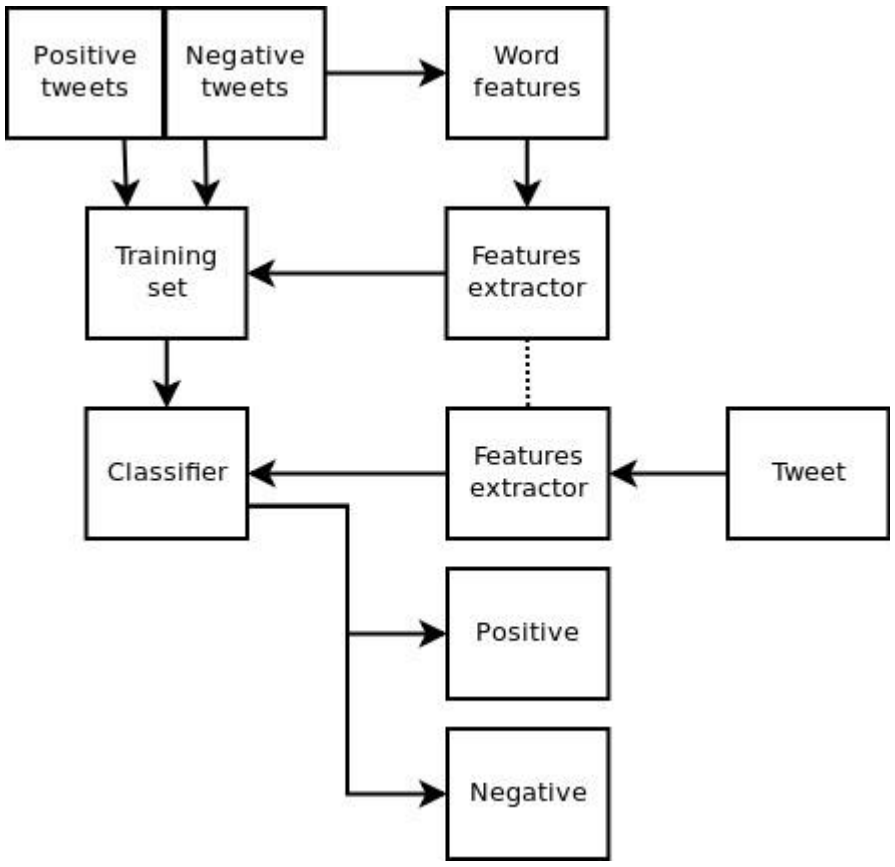


Fig 4.2 Object diagram for the sentiment analysis.

4.3.1.3 System Architecture

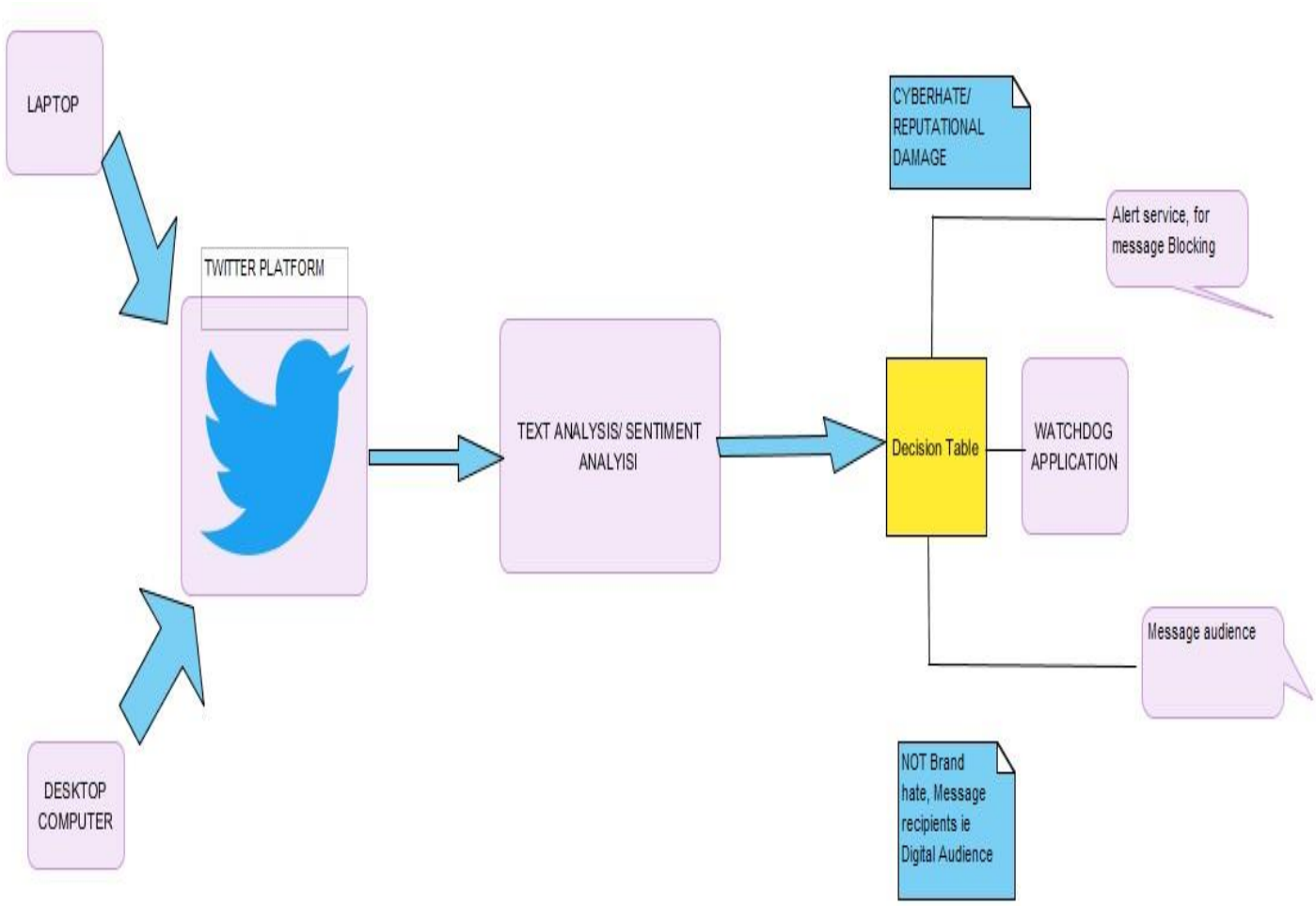
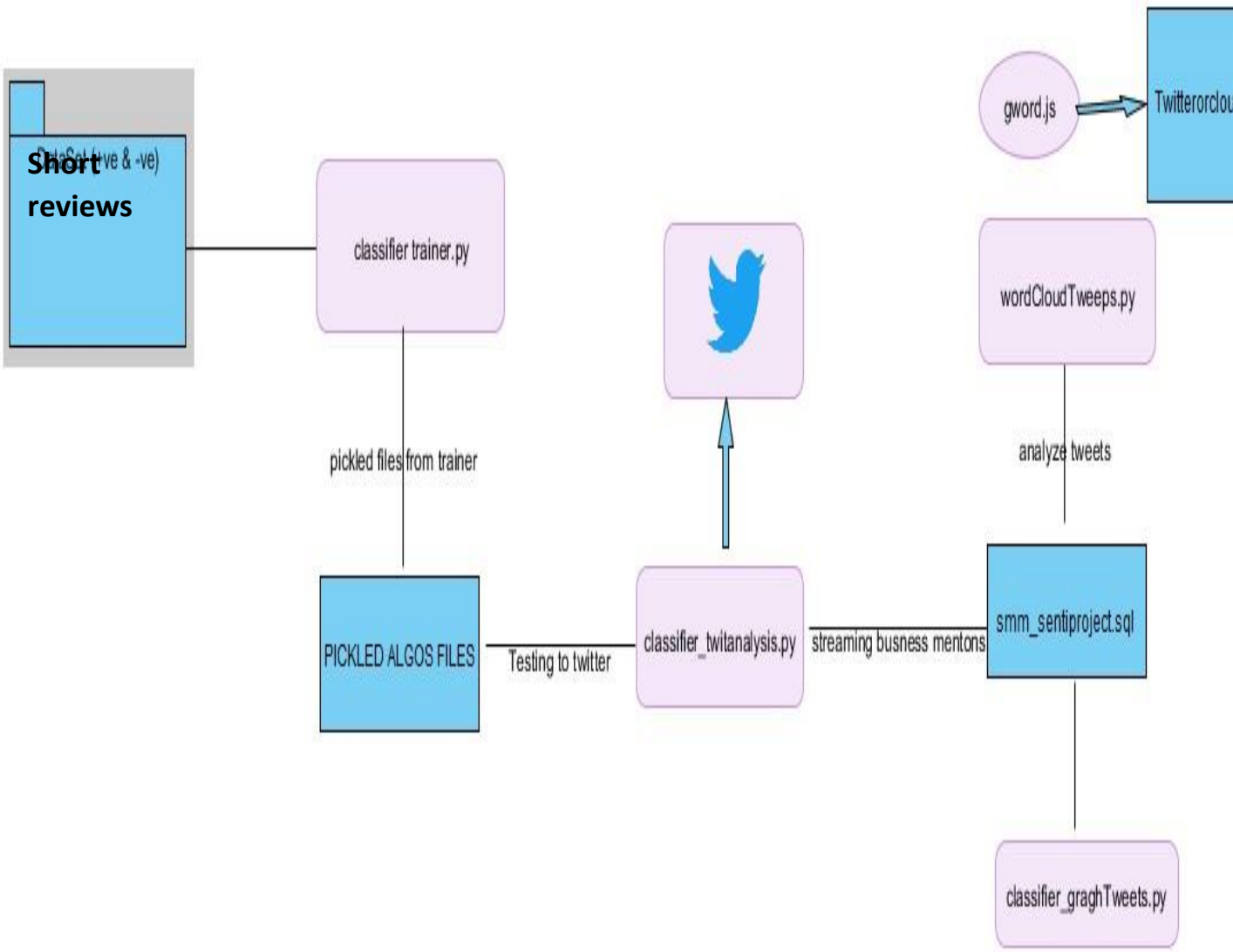


Fig 4.3 system architecture

We can access the twitter API, from any laptop or desktop computer, then we can do a stream on the tweets for the mentions of Safaricom company. Finally, we do text analysis on this tweets to classify them whether they are positive or negative and assign a polarity/confidence score. Eventually, for tweets with a higher polarity score, we send notifications via e-mail to the personnel responsible for social monitoring from which, He can figure out the tweets to be flagged before they create an uproar on twitter if they are of effect to the company's reputation.



The flow of the system modules and code files, the classifier_trainer.py is trained against a dataset of positive and negative tweet data in the folder “Short reviews”. The trained data is saved in binary data as pickled files for each of the algorithm used into the folder “Picked algos file”. Consequently, the Classifier_twitanalysis.py uses the trained data and accesses twitter leveraging the twitter API. It then does sentiment analysis on the tweets containing only mentions of Safaricom. This is done by obtaining a stream of tweets for the mentions of Safaricom. This stream is classified into positive and negative tweets and they are assigned a polarity score. This data is then crawled into a database smm_sentiproject.sql, where it is stored for analysis. The analysis is done by the wordCloudtweeps.py. Which then writes the analyzed tweets to gword.js, which can later be viewed

in a browser by opening the twittercloud.htm file to see the accounts tweeting the most negative tweets with high polarity score.

4.4.1 Implementation

Six algorithms were used in the classifier which can be shown in the figure below with their accuracy percent, then lastly a voted classifier is used to find an average accuracy that is overall used as the accuracy percentage of the classifier. Using more than one algorithm is meant to improve the accuracy of the classifier.

Fig 4.3 the classifier algorithms and their accuracy percentages

```
In [11]: run classifier_trainer.py
10662
Original Naive Bayes Algo accuracy percent: 71.29909365558912
Most Informative Features
              engrossing = True              pos : neg    =     20.8 : 1.0
                 routine = True              neg : pos    =     15.8 : 1.0
                 generic = True              neg : pos    =     15.2 : 1.0
                    flat = True              neg : pos    =     14.3 : 1.0
              refreshing = True              pos : neg    =     13.5 : 1.0
               wonderful = True              pos : neg    =     12.1 : 1.0
                    warm = True              pos : neg    =     12.1 : 1.0
                mindless = True              neg : pos    =     11.8 : 1.0
               realistic = True              pos : neg    =     11.6 : 1.0
                   stale = True              neg : pos    =     10.4 : 1.0
                tiresome = True              neg : pos    =     10.4 : 1.0
                  stupid = True              neg : pos    =     10.3 : 1.0
          extraordinary = True              pos : neg    =     10.2 : 1.0
            mesmerizing = True              pos : neg    =     10.2 : 1.0
                    wry = True              pos : neg    =      9.6 : 1.0
MNB_classifier accuracy percent: 71.90332326283988
BernoulliNB_classifier accuracy percent: 71.45015105740181
LogisticRegression_classifier accuracy percent: 73.1117824773414
LinearSVC_classifier accuracy percent: 71.6012084592145
SGDClassifier accuracy percent: 69.18429003021149
voted_classifier accuracy percent: 71.75226586102718

In [12]:
```

```
twitterStream.filter(track=["safaricom"])
```

```
('She_united', 'RT @saint_makaveli: Dear safaricom\nI slept with 1.5 GB of data then i wake up to find " your data is below 0.8
 MB\' sms jeeez was i streamin…')
('iKinuthia_', 'RT @PorkReebz: In my pants. https://t.co/UtyLgQN7Pl')
('Timberwolf___', 'RT @saint_makaveli: Dear safaricom\nI slept with 1.5 GB of data then i wake up to find " your data is below
 0.8 MB\' sms jeeez was i streamin…')
('Paulo_Adoop', 'RT @saint_makaveli: Dear safaricom\nI slept with 1.5 GB of data then i wake up to find " your data is below 0.
8 MB\' sms jeeez was i streamin…')
('njuechris5', 'RT @SafaricomCare: Please find assistance from the official Safaricom customer care twitter account @Safaricom_
Care')
('kaptila44silas', '@SafaricomLtd Send the mpesa menu to my phone 0724158285. It just disappeared in my phone with the safarico
m toolkit.')
('OketchDerrick2', 'RT @saint_makaveli: Dear safaricom\nI slept with 1.5 GB of data then i wake up to find " your data is below
 0.8 MB\' sms jeeez was i streamin…')
('RobahRdm', 'RT @saint_makaveli: Dear safaricom\nI slept with 1.5 GB of data then i wake up to find " your data is below 0.8 M
B\' sms jeeez was i streamin…')
('gikuyu254', 'RT @PorkReebz: In my pants. https://t.co/UtyLgQN7Pl')
('Lets_B_Real', 'After Launch in Nairobi, Safaricom’s Little is going to Nigeria https://t.co/FNAgb4qCH0 viaafrica #business #e
ntrepreneur')
('innov8tivmag', 'After Launch in Nairobi, Safaricom’s Little is going to Nigeria https://t.co/NkcDpdQyRB viaafrica #business #
entrepreneur')
('IBOMLLC', 'After Launch in Nairobi, Safaricom’s Little is going to Nigeria https://t.co/tJoCyrmEr4 viaafrica #business #entre
preneur')
('GuruAfrica', 'After Launch in Nairobi, Safaricom’s Little is going to Nigeria https://t.co/mSqtsNWQxJ viaafrica #business...
https://t.co/ZP01vOL59w')
('NBITLO', 'After Launch in Nairobi, Safaricom’s Little is going to Nigeria https://t.co/RO0ktPkIHP viaafrica #business... http
s://t.co/BQmp9mgZfU')
('IBOMLLC', 'After Launch in Nairobi, Safaricom’s Little is going to Nigeria https://t.co/tJoCyrEfiC viaafrica #business... htt
ps://t.co/P2hRR0AXjz')
('Safaricom_Care', '@jmurrayth browsing session on the Safaricom 4G network and the 4GB Data Bundle will be sent to you. ^NJ')
('TeddyLumidi', 'Safaricom planning to Release holistic M-Pesa API for Developers \nhttps://t.co/FLuJPtCq9g via @techweez')
('donxut6', 'jmurrayth browsing session on the Safaricom 4G network and the 4GB Data Bundle will be sent to you. ^NJ')
('TimKanya', "Experience the thrilling life with Safaricom 4G. If you are a virgin don't wait!")
('BensonM00985352', 'RT @SmileInvestClub: To follow us via sms and be able to receive live updates from us: \nSend sms to 8988
 (Safaricom) or 4040 (Airtel) \nMess…')
('Nyaoks', "@SafaricomLtd why won't https://t.co/ort7jy8Au5 work!?")
('SafaricomLtd', '@zecky_obonyo Hi, DM your mobile number or login to your Selfcare account https://t.co/YdsIu4T5wG in order to
 view your billing.^JM')
('SafaricomLtd', '@obvin56 Hi. Please share your number we check and advice. You may also view billing on Selfcare here http
s://t.co/SETkSQf0BW  ^WP')
```

Fig: 4.4 classified tweets either positive or negative

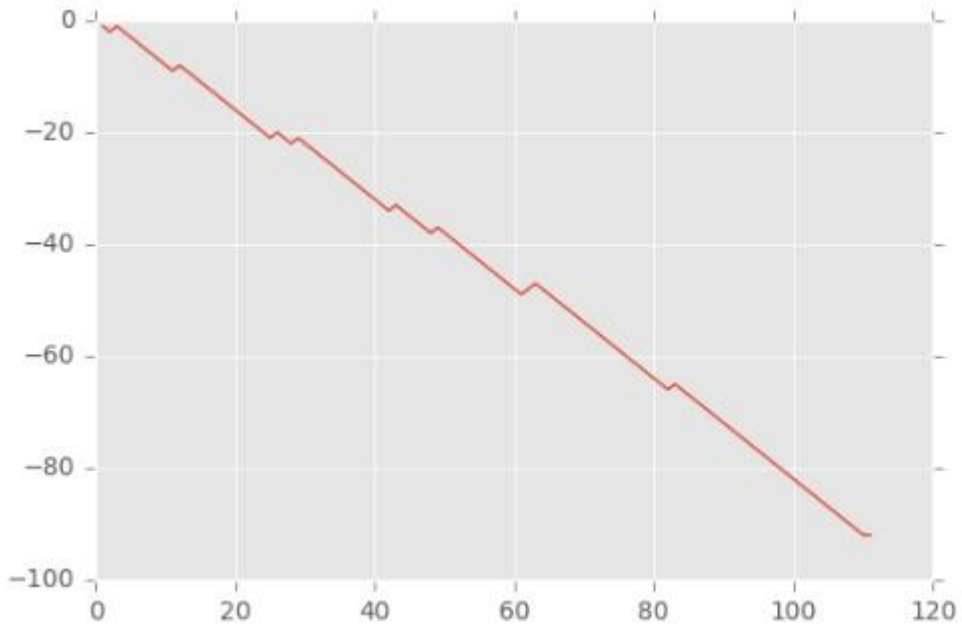


Fig 4.5 the graph of tweets with a polarity score higher than 80%

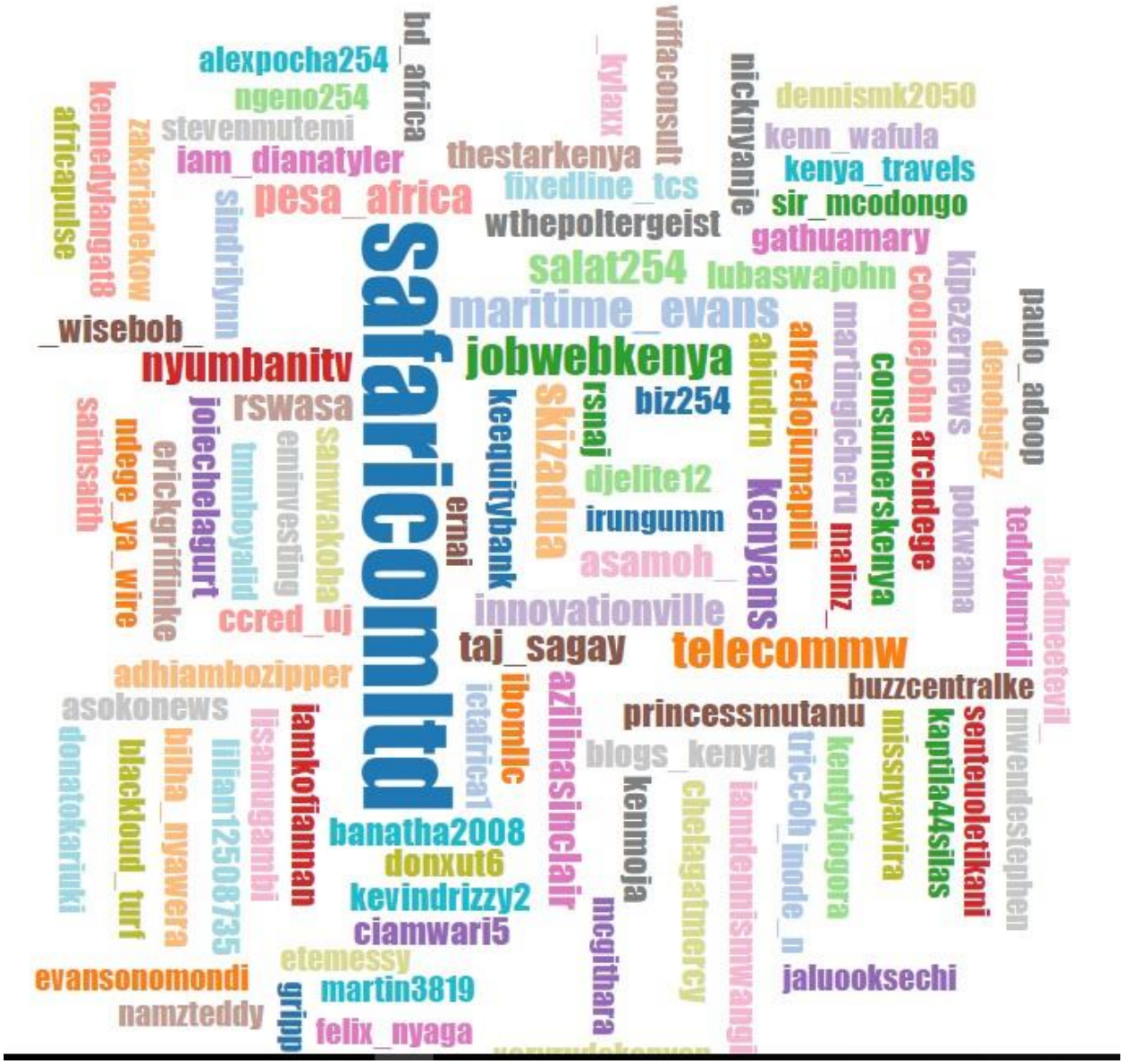


Fig 4.6 showing accounts tweeting more about Safaricom

4.4.5 Testing and validation

Here we describe how testing and validation tasks were performed, by describing the plans and strategies used in unit testing, integration testing and system testing. We also define the test plans and provide test procedures for testing the critical functions. Finally, we describe the test tools used

Chapter 5

5.0 Discussion and Conclusion

The project presents an approach for brand hate detection based on state-of-the-art natural language processing and a common sense knowledge base, which permits recognition over a broad user sentiments. We analyze comments made towards a particular brand using the Naïve Bayes algorithm to determine whether it is positive or negative and then assigning a polarity score. According to the classifier comments that are high in the polarity score, should then be flagged. Decisions can further be made on whether to send cease and desist letters. Or even blocking the individuals totally.

Statistical analysis of these brand hate messages can also help to identify a correlation, that is, whether an increase in brand hate/reputational damage by a given social audience is also statistically linked to an increase in losses for the company or causes hitches to the business process. In addition, it will help achieve the long-term goal of profiling these users whether they are insiders or outsiders.
The main goal of the project will be to help these enterprises mitigate some of the risks that arise from the use of social media such as social media squatting, cyber loafing, phishing and even loss of passwords. This is because by monitoring this sites they will be able to get insightful intelligence of whether someone outside there is squatting on their accounts and using it for malicious purposes. Furthermore, it will help them be prepared for any serious brand hate messages that when posted may result in an uproar in the larger online audience. It will thus help in analyzing insider and outsider threat of conceit. It will look at how the self-righteous activities of some of the employees and outsiders might affect the enterprise leading to such factors as flash mobs, boycotts, and reputation damage.

References

Adrian Bowes. (2016, October 28). *Set social media risk Managemnt policies by preparing for the worst.* Retrieved from techTarget: C:\Users\BUNDEX-PC\Downloads\Documents\http___searchcompliance_techtarget_com_tip_Set-social-media-risk-management-policies-by-preparing-for-the-worst.pdf

Cole, B. (2016, October 28). *Ways to mitigate risk with a corporate social media policy.* Retrieved from Techtarget.com: http___searchcompliance_techtarget_com_news_2240037516_Ways-to-mitigate-risk-with-a-corporate-social-media-policy.pdf

Cybersquatters-hit-e-commerce. (2016, November 31). Retrieved from http://www.businessdailyafrica.com/: http://www.businessdailyafrica.com/Cybersquatters-hit-e-commerce-/1248928-1498082-vbr0et/index.html

Eric Wainaina. (2016, November 31). *http://www.techweez.com/news*. Retrieved from http://www.techweez.com/: http://www.techweez.com/2016/07/21/uhuru-co-ke/

Gritzalis; Kandias; Stavrou; Mitrou;. (2012). *History of Information: The case of Privacy and Security in Social Media.* Athens: Information Security & Critical Infrastructure Protection Research Laboratory.

ISACA. (2016). Social Media: Business Benefits and Security, Governance and Assurance Perspectives. *An Isaca Emerging Technology White Paper* (p. 10). Rolling Meadows, IL 60008 USA: ISACA.

Kaigwa, mark; Madung, Odanga; Costelo, Samer;. (2015). *NENDO 2014/15 SOCIAL MEDIA TREND REPORT*. NAIROBI: NENDO.CO.KE.

Macharia Kihuro. (2015). *SOCIAL MEDIAAND THE INHERENT RISKS TO FINANCIAL SERVICES INDUSTRY*. NAIROBI: Shelter Afrique.

Maxwell . (2016, Nov 2). Reducing the Risks of Social Media to Your. *Security Policy and Social Media Use*, p. 28.

OSINT. (2016, Nov 1). Retrieved from OSINT: http://www.osint.org/

pwc. (2013). *2013 Information Security Breaches.* info security Europe.

Steven Bird, Ewan Klein, and Edward Loper (2009).

Natural Language Processing with Python. O'Reilly Media Inc.

http://nltk.org/book

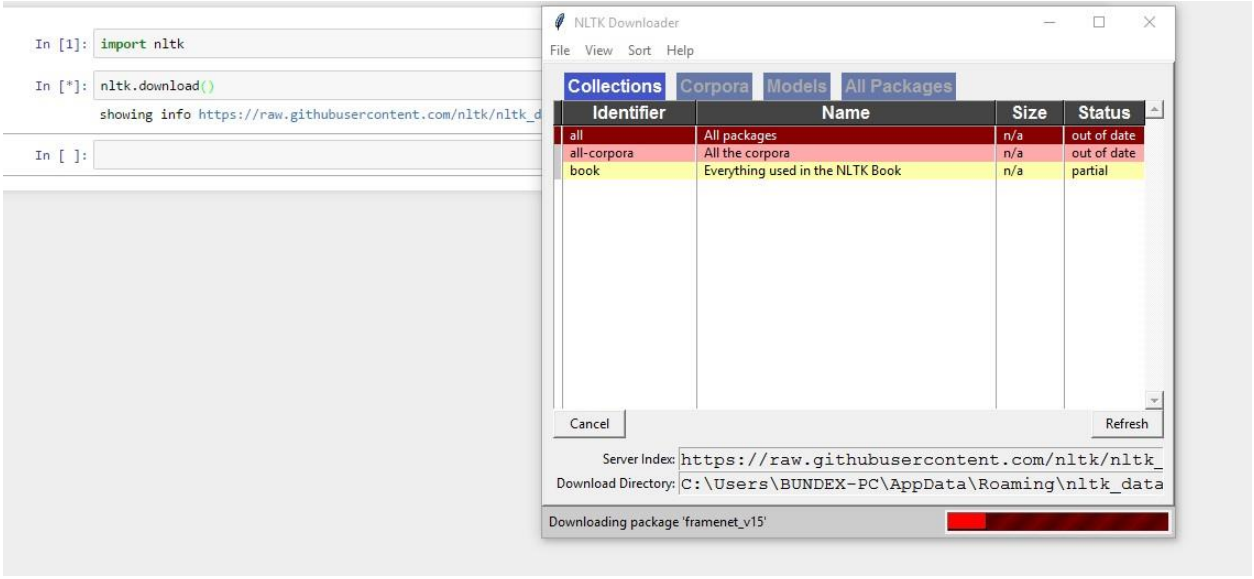
Six Converging Technology Trends. (2016). *Google.com*. Retrieved 26 November 2016, from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKEwiQosWIwsbQAhXDuhoKHfvoA_EQFggjMAA&url=https%3A%2F%2Fwww.kpmg.com%2FBE%2Fen%2FIssuesAndInsights%2FArticlesPublications%2FDocuments%2FSix-Converging-tech trends.pdf&usg=AFQjCNFjx4hZxCnEKGbf0jwo70SfvYPNnQ&sig2=89NsVv6HlWTQis9IQEEzKw

Appendixes

NLTK Installation and Setup

The NLTK module can be installed the same as other Python modules, either by downloading the package through the NLTK website directly or by using any number of third-party installers with the keyword “nltk”. After installing the module it’s preset text repositories are also downloaded. Some of the features can be tried out more easily by typing the following Python command line:

```
>>> import nltk
>>> nltk.download()
```



Twitter OAuth 1.0a Flow with Ipython Notebook

Twitter implements OAuth 1.0A as its standard authentication mechanism, and in order to use it to make requests to Twitter's API, you'll need to go to https://dev.twitter.com/apps and create a sample application. There are three items you'll need to note for an OAuth 1.0 A workflow, a consumer key and consumer secret that identify the application as well as the oauth_callback URL that tells Twitter where redirect back to after the user has authorized the application. One also needs an ordinary Twitter

account in order to log in, create an app, and get these credentials. For development purposes or for accessing your own account's data, one can simply use the OAuth token and OAuth token secret that are provided the application settings to authenticate as opposed to going through the steps here. The process of obtaining and the OAuth token and OAuth token secret is fairly straight forward (especially with the help of a good library

You must ensure that your browser is not blocking pop-ups in order for this script to work.

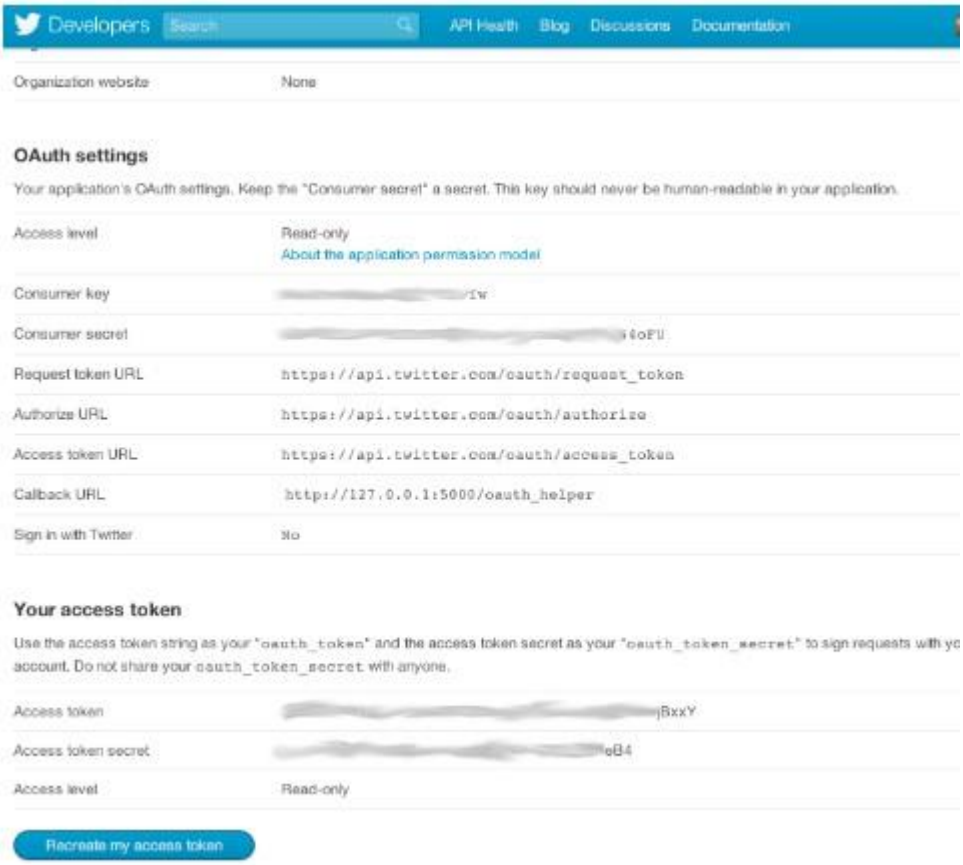


fig 1 the twitter OAuth flow

Anaconda installation

Anaconda is a FREE enterprise-ready Python distribution for data analytics, processing, and scientific computing. Anaconda comes with Python 2.7 and 100+ cross-platform tested and optimized Python packages. All of the usual Python ecosystem tools work with Anaconda.

Additionally, Anaconda can create custom environments that mix and match different Python versions (2.6, 2.7 or 3.3) and other packages into isolated environments and easily switch between them using conda, our innovative multi-platform package manager for Python and other languages.

For Detailed Anaconda Installation Instructions check out

http://docs.continuum.io/anaconda/install.html

INSTALLATION

System Requirements		
Linux	Windows	Mac OS X
32/64 bit x86 processor	32/64 bit x86 processor	64-bit x86 processor

Download Anaconda

Tweepy Installation
Tweepy supports OAuth authentication. Authentication is handled by the tweepy.AuthHandler class.

Tweepy can be installed from command line by this command

->Pip install tweepy

OAuth Authentication

Tweepy tries to make OAuth as painless as possible for you. To begin the process we need to register our client as in (Twitter OAuth flow above) application with Twitter. Create a new application and once you are done you should have your consumer token and secret. The next step is creating an OAuthHandler instance. Into this we pass our consumer token and secret which was given to us in the previous paragraph:

auth = tweepy.OAuthHandler(consumer_token, consumer_secret)

If you have a web application and are using a callback URL that needs to be supplied dynamically you would pass it in like so:

auth = tweepy.OAuthHandler(consumer_token, consumer_secret, callback_url)

If the callback URL will not be changing, it is best to just configure it statically on twitter.com when setting up your application's profile.

Unlike basic auth, we must do the OAuth "dance" before we can start using the API. We must complete the following steps:

1. Get a request token from twitter
2. Redirect user to twitter.com to authorize our application
3. If using a callback, twitter will redirect the user to us. Otherwise, the user must manually supply us with the verifier code.

4. Exchange the authorized request token for an access token.