Big Data, Internet Privacy and the Vulnerabilities of the African Regulatory Landscape

Dr. Mohammed Suleh-Yusuf
Post-Doctoral Researcher on Cyber Security, Data Privacy and AML/CFT Regulations.
Nasarawa State University, Keffi, Nigeria

Abstract

Social media generates massive amount of big data from users, the penetration of these platforms in Africa creates meaningful insights around customer needs and behaviour from the data. This helps to create new businesses that rebalance the technology and wealth gap in the continent. With every gigabytes of data generated brings about exploitation of customer data. Data Privacy becomes a focal point of concern. The global approaches to privacy for the users of social media platforms is still evolving but two jurisdictions have set a standard. The European General Data Protection Regulation (GDPR) and the Californian Consumer Privacy Act (CCPA) have provisions that are built on sustaining consumer consent and enabling consumers to be forgotten or have their consented data deleted at their own request. More so the exponential growth of the internet in Africa highlights the explosion of big data and there is a need to study its regulatory approaches in relation to the global best practices symbolized by the GDPR and the CCPA. The Paper reviews the African regulatory landscape and its approach to Big Data and possible vulnerable angles that exposes data of Africans on these social media platforms. It is clear that in spite of a continental treaty and a reasonable number of African countries with Data Privacy laws, these laws are in most of the countries either not built on strong legal grounds or lack an independent enforcement mechanism. Therefore the African approach leaves a lot of open issues and there is a need for a continental consensus on the best approach that will push through national legislations crept on a unifying continental model.

Keywords: Big Data, Privacy, Regulatory, Social Media, Data analytics, Guidelines, rights
DOI: 10.7176/EJBM/12-17-14
Publication date: June 30th 2020

Introduction

There has been a somewhat sudden move to Big Data and its many variables in the technological and connected world we live in. This move has created frontiers of opportunities as well as an ever expanding cache of challenges. This move has also escalated hitherto reasonable concerns on related matters such as data privacy to a scale that creates vulnerabilities unfathomable a decade ago. This Paper hesitates to define 'Big Data' in an academic form but will rather provide a functionality driven explanation of the concept and its implications. Big Data is an area of activity that indicates ways to analyze, systematically extract information in different formats, or otherwise utilize data sets that are too large or complex to be dealt with by traditional data-processing application software. Hence over time the concept of Big Data has been associated with volume of collected data, the variables of such collected data and the velocity of its accumulation. Surprisingly these characteristics have expanded to include method of storage of data, ownership and privacy of collected data and more importantly the regulatory environment that ensures 'Big Data' stays hygienic and within boundaries of consent and sustainable protection from misuse.

Why is Big Data important? In the world of today big data analytics and its mesmerizing ingredients have pushed organisations, businesses and governments into a level that helps them harness data and use it to identify new opportunities, explore marketing angles and govern their countries. That, in turn, leads to smarter business strategies, cutting edge efficiency in operations, soaring profits and smart urban planning, higher level of security of countries and better social services. Therefore Big Data is an important element of the future, in a hyperbolic hype it can even be construed as the future itself. The manner companies, organisations and governments rely on Big Data analysis and its outputs will continue to shape business ideas, organizational mapping and projections for expansion and growth. Certainly it does not refer to a specific amount of data, but rather describes a dataset that cannot be stored or processed using traditional database software. Good examples of big data include the Google search index, the database of Facebook user profiles, and Amazon.com's product list. There has been a growing realization of the unique influence that Big Data can exert in providing priceless insights to any company, organization or government in planning and even implementation of business concepts or programmes that can rely on the accurate outputs of the analytics.

This Paper will look at Big Data generated from three primary sources: social data, machine data and transactional data. The Social data comes from the Likes, Tweets & Retweets, Comments, Video Uploads, and general media that are uploaded and shared through the world's expanding social media platforms. This form of data provides invaluable insights into consumer behavior and can be enormously useful in marketing analytics, product planning and sales projections. The internet is also a good source of social data, and tools like Google

Trends can be used to good effect to increase the accumulation of Big Data or even Google Map. While machine data is defined as information generated by equipment, sensors that are installed in buildings or devices and public cameras and surveillance machines. This type of data is expected to grow exponentially as the internet of things grows ever more pervasive and expands around the world. The Sensors may be related on medical devices, smart electricity meters, road cameras installed by municipal authorities and the rapidly growing Internet Of Things (IOTs)[1]. This form of data can be delivered at high velocity, value, volume and varieties in the very near future. The last form of Big Data is the transactional data that is generated from daily activities that take place online (this Paper restricts itself to online activities but acknowledges that there are substantial number of activities still offline), digitalized Invoices and, payment vouchers, public storage records, online subscriptions, banking and credit card activities etc.

This Paper acknowledges that Big Data is now an integral part of a technologically-driven world and an inevitable byproduct of billions of people accessing the internet and leaving behind bits and pieces of their personal information as well as transactional footprints. Yet we cannot accept this somewhat inevitability without interrogating its impact on data privacy, data protection and cumulative impact on lives of human beings. Thus the speed of accumulation of Big Data cannot negate the principles that will ensure personal information are protected and their rights to privacy guaranteed. The not too recent rise in mobility and participation in social networks, the increasing willingness to share more and more data by both discerning and non-discerning individuals is quite worrisome. Likewise the exponential deployment of new technology that captures more data, and the growing commercialization of Big Data only leads to the erosion of data privacy.

The reaction of Europe through the General Data Protection Regulation (GDPR)[2] is not a fleeting counter punch to Big Data and issues related to gathering of data of its citizens. It is indicative of the growing concern of governments across the globe that traditional data privacy and protection principles are showing patent weaknesses in meeting what can be concluded as an existential threat to the rights to privacy and data protection. Another good illustration is the Californian Consumer Privacy Act[3] in the United States and its attempt to upscale the regulatory landscape to meet the beclouding challenges of digitalization of data and its impact on lives of persons.

The focus of this Paper is to review the African regulatory landscape and its reaction to the issues around Big Data and possible impact on its citizenry. Clearly the African continent is at the knife edge part of the new internet world and until it reacts appropriately (not necessarily in the form of a GDPR) it will remain a fertile ground for deployment of applications and platforms that will exploit mountains of data generated. This exposure will further exacerbate the challenges the continent faces in relation to universal access to the internet, high prices of services and not-existing innovation outputs. To review the regulatory landscape of the continent will be restricted to matters related to Big Data collated through the internet, either through social media platforms or search engines. Hence the review, while acknowledging its existence, will not look at data collated offline and those that are not digitalized.

The Internet and the Big Data World

The Mckinsey Global Institute[4] defined Big Data as “datasets whose size is beyond the ability of typical database software to capture, store, manage or analyse”[5]. In its Report the Mckinsey Global Institute states that all the biggest internet companies; such as Google, Facebook, Amazon, eBay, Microsoft and Yahoo, are engaged in the collection and utilization of Big Data. It further states that these companies use Big Data as a major asset and a source of value creation in their business modelling. It used Google as a good example and concludes that the company has used Big Data in tuning its search engine an algorithm to deploy data intensive services such as voice recognition, translation and location-based services. The conclusion of this Report is also in line with the postulation of Zheng, Vasilakos et al[6], where they state that there is a proliferation of devices like smart phones, mobile tablets, sensors, smart metres and smart appliances that feed from an increasing mobile data traffic on the internet. To them Big Data is all about its volume, variety, velocity and value; all of which cumulatively create a valuable asset for companies like Google.

Therefore, Big Data has become a priceless commodity for businesses, organisations and governments, with most of its valuation tied to its use rather than its storage. This Paper is looking at Big Data derived from internet activities and utilized for marketing, market profiling and socio-economic planning by companies and

[1] We must acknowledge that these technologies are not widespread in Africa but there have been promising penetration of new technologies

[2] https://www.researchgate.net/deref/https%3A%2F%2Feur-ex.europa.eu%2Feli%2Freg%2F2016%2F679%2Foj%2Feng <accessed on 9 March 2020>

[3] Californian Consumer Privacy Act 2018

[4] The Economic and Business arm of McKinsey Consulting

[5] MGI Report on Big Data 2016 https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/big-data-analytics-should-be-driven-by-business-needs-not-technology <accessed on 22 March 2020>

[6] Dai, H. N., Wong, R. C. W., Wang, H., Zheng, Z., & Vasilakos, A. V. (2019). Big data analytics for large-scale wireless networks: Challenges and opportunities. *ACM Computing Surveys (CSUR)*, *52*(5), 1-36.

governments. N. Abdul Ghani and S. Hamid[1] published the outcome of a survey of social media and Big Data analytics; they stated that Big Data analytics has emerged as a key source of research information due integration of background details and daily activities of its users. The same conclusion was made by Betty Jane and Ganesh because in their opinion the increase in growth of new and innovative technologies and rising reliance on Internet of Things (IoTs) has produced a large amount of data. These opinions have shown that data gathering and utilization from the internet and other online platforms is now a market of itself.

But why is Big Data important? The Big Data analytics market is set to reach $103 Billion by 2023[2] and in 2020 every internet user will generate at least 1.7 megabytes of data every second. In fact internet users generate about 2.5 Quintillion bytes of data per day[3]. It is quite clear that Big Data has impact on business performance and delivery of government programmes; with increased profits and implementation efficiency. R. Mihet[4] et al points out that Big Data has features that makes it a fuel for increased business performance. In the Paper they outlined the important features of Big Data as it relates to business performance. First, Big Data is a byproduct of an economic activity. Secondly, that companies use Big Data to increase their efficiency. Thirdly, Big Data is a collection of information that is distinct from technology or the platform producing it. Lastly, that such accumulated Big Data is a valuable asset. The Paper, though restrictively focused on United States companies, highlights the importance of Big Data to companies harvesting them online.

There are several avenues of collating Big Data and utilizing them for analysis, mapping and planning; these includes data collated from consumers by companies. The focus of this Paper is the Big Data accumulated through the internet, particularly through social media platforms. The collation of data through these media throws up several data privacy challenges and more importantly the quantum of the data is both astounding and impressive. There are billions of people accessing these platforms, leaving behind data footprints and zillions of bytes of personal information. Clearly the speed of the growth of the internet and explosion of medium such as Facebook and Twitter has sped pass the traditional buffers protecting personal data and related information of individuals.

To illustrate the magnitude of Big Data on the internet, we can easily review statistics of two giants in the Big Data market. Google, a search engine own by Alphabet Inc., receives 63,000 searches per second and owns 90.46% of the search engine segment. Google has an advertisement revenue of $95.46 Billion in 2017[5], with 2.5 billion users on Android and 500 million on Google photos. Google is also dominant in the deployment of applications and technologies, with more than three million mobile applications on Google Play Store in 2018[6]. It is noteworthy that Google owns YouTube, a video streaming service, with more than 1.9 billion users every month. While Facebook has 2.6 billion users monthly as at 1st Quarter of 2020 and generates 4 Petabytes of data per second (that is one million gigabytes), incidentally Facebook owns WhatsApp and it has two billion persons accessing it every month[7]. The amount of Big Data the two companies generate is immense and it will increase as new data gathering advancements are made and deployed.

Where do all these users come from? The internet has expanded beyond any projection and it has created avenues and mechanism of engagements; that generate data and information that is being used for advertisement and commercial mapping by companies such as Google and Facebook. It will help to look at the statistics of internet users as it will also pinpoint where the most data can be generated. The table below shows internet users by region:

[1] Ghani, N. A., Hamid, S., Hashem, I. A. T., & Ahmed, E. (2019). Social media big data analytics: A survey. *Computers in Human Behavior*, *101*, 417-428.
[2] https://www.internetworldstats.com/stats.htm <accessed on 7 April 2020>
[3] ibid
[4] Farboodi, M., Mihet, R., Philippon, T., & Veldkamp, L. (2019, May). Big data and firm dynamics. In *AEA papers and proceedings* (Vol. 109, pp. 38-42).
[5] https://techjury.net/stats-about/big-data-statistics/#gref <accessed on 11 April 2020>
[6] ibid
[7] ibid

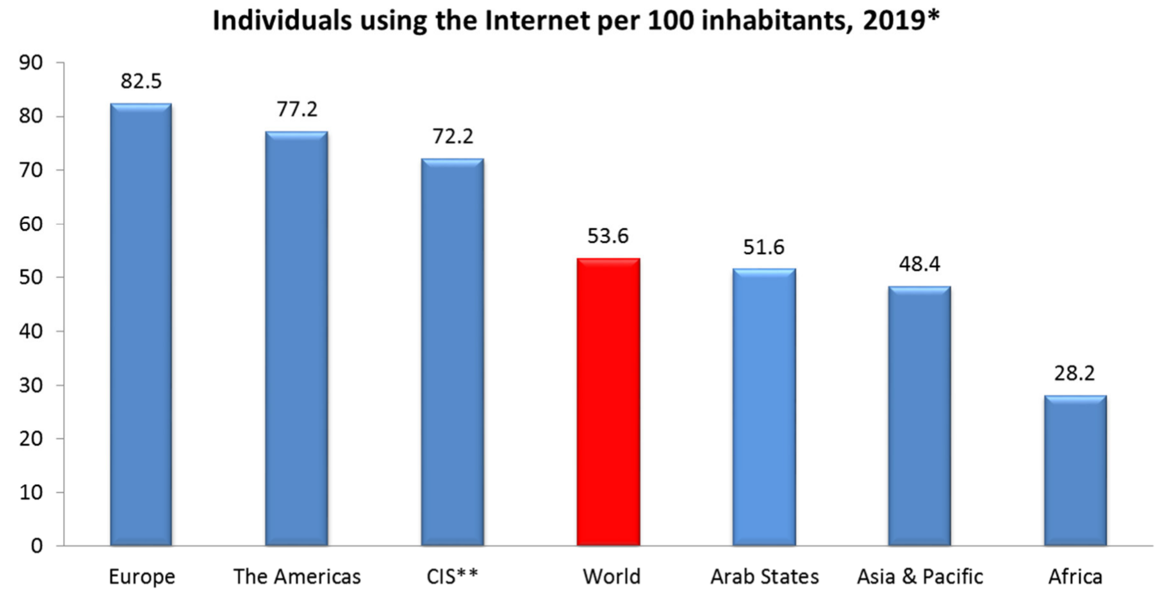

Table 1: Individuals using the Internet per 100 inhabitants 2019[1]

The Chart clearly shows that there are more internet users in Europe and the least is in Africa, the differences in the statistics highlights or even mirrors the economic and development levels of each Region. We must though admit that the chart leaves open other issues for inquiry. The condensed nature of the statistics cannot provide clarity as to whether there are individual countries with substantial internet users contrary to their regional composite data.

Table 2 below provide internet usage per country and selects ten countries randomly across all the regions covered by the composite statistics in Table 1 above.

Table 2: Statistics of Internet Usage per countries[2]

Country	**Region**	**Population**	**Number of Internet Users**
United Kingdom	Europe	66,574, 000	63,544,106
Germany	Europe	82,293,000	79, 127, 551
South Korea	Asia	51,164,000	49, 484, 000
China	Asia	1, 415, 046,000	854,000,000
Saudi Arabia	Middle East	33,554,000	27, 048, 961
Morocco	North Africa	36, 192, 000	22, 072, 765
South Africa	Southern Africa	57, 398, 000	31, 858, 027
Egypt	North Africa	99, 376, 000	49, 231, 493
Nigeria	West Africa	195, 875,000	126, 078, 999

While Table 3 shows the statistics of smart phone users in the same countries captured in Table 2 above. This further illustrates not just internet penetration but more importantly it points to those who surf the internet through smart mobile devices that make them easy ‘preys’ for data collection

[1] https://www.itu.int/en/ITU-D/Statistics/Pages/default.aspx <accessed March 18 2020>
[2] ibid

Table 3: Statistics of Smart phone usage in selected countries[1]

Country	Region	Population	Internet Penetration Rate	Smart Phone Users
United Kingdom	Europe	66,574, 000	82.2%	54, 713,000
Germany	Europe	82,293,000	78.8%	64,830,000
South Korea	Asia	51,164,000	68%	34,562,000
China	Asia	1, 415, 046,000	55.3%	782, 848,000
Saudi Arabia	Middle East	33,554,000	46%	15, 449, 000
Morocco	North Africa	36, 192, 000	37.9%	13, 707, 000
South Africa	Southern Africa	57, 398, 000	35.5%	20, 371, 000
Egypt	North Africa	99, 376, 000	28%	27, 852, 000
Nigeria	West Africa	195, 875,000	13%	25, 552, 000

This large population of internet users generates billions of bytes of data that feed the companies that rely on Big Data to harvest patterns, trends and customer behaviour. Therefore to understand the impact (or possible impact) of Big Data on individuals, particularly on data privacy and protection, we must view the level of exposures.

To further illustrate how the Big Data on the internet is gathered, the Table below highlights the daily and monthly usage of a select group of social media platforms as at the first Quarter of 2020.

Table 4: Social media activity statistics[2]

Platform	Daily Users	Monthly Users
Facebook	2.26 Billion	2.50 Billion
Snapchat	218 Million	360 Million
Twitter	152 Million	330 Million
Instagram	500 Million	1 Billion
YouTube	30 videos watched daily	2 Billion Videos watched monthly

It is amazing that these numbers of persons access these platforms on daily basis and leaving behind data and information that can be used for Big Data analytics. A good illustration is that there are 50 Billion pictures shared on Instagram as at January 2020[3], with a daily upload of 100 Million videos and pictures. While on Twitter there are about 500 Million tweets daily from 152 Million users[4]. Another huge number is the 3 Billion snapchats being created every day by 218 Million users[5].

Data Privacy, Regulatory Approaches and the Internet

There has been a growing concern on the erosion of privacy on the internet and the attendant issues of data ownership and protection. With the rapid generation of information on the internet and the Internet of Things (IoT), companies and platforms have an unprecedented capacity to collect, analyse and share data. At the same time, there is a new awareness that such massive data accumulation provides a huge opportunity for information discovery and possible compromise of the privacy of the data owners. Therefore, there is a nagging question on how to protect the data privacy on the internet even though the collection methods and speed differs from traditional data collection processes. In this paper, the “information rush" which is characterizing the current phase of the Big Data age calls for actions aimed at enforcing the citizens' right to privacy. Clearly the entire data life-cycle is being driven by technologies, leading to cropping of data through the deployment of algorithms that collect Big Data continuously and consistently.

Technology has progressed and is giving room for varied techniques for collecting data and manipulating them. In fact most of the data individuals generate on the internet and all the activities that they undertake during their life are available. These data are collected by the digital devices and applications that have become indispensable in our daily lives, good examples are: the Internet itself, social media/network platforms, email services, game, fitness, lifestyle and several others. Furthermore, with the growth of Social Networks individuals are now voluntarily sharing personal information and the platforms are using data analytics algorithms to create the Big Data that is now a prized asset for the global business world. This means that individuals will never be able to know which of their personal information is known somewhere and what analysis they may have gone through.

Incidentally Europe is playing a key role through the General Data Protection Regulation (GDPR) that came into force on 25th May, 2018[6]. This regulation aims to protect the privacy rights of individuals, and to protect

[1] ibid
[2] https://techjury.net/stats-about/big-data-statistics/#gref <accessed on 7 March 2020>
[3] ibid
[4] ibid
[5] ibid
[6] It has set the standard globally and a key reference document for regulators across the world.

these rights from erosion by business models that thrive on cropping these data for modelling and other commercial activities. This paper sees the GDPR, not as a perfect Regulation but as a benchmark of an approach that understands the role technology and the internet is playing in relation to privacy rights of individuals. The GDPR applies to most application firms and networks within Europe as it clearly states that it applies to any organization that "processes personal data" of any individual that is a citizen of any country within the purview of the GDPR[1]. Importantly, the GDPR defines "personal data" broadly, as any information that might identify a consumer ("data subject"), while it states that "processing" as any operation that is performed on personal data, whether or not by automated means. The GDPR refers to some organisations/companies as 'Data Controllers'[2] and places the burden of responsibilities on them to determine the purposes and the means of processing of personal data, while organizations/companies that process personal data on behalf of the Data Controllers have to comply with a considerable portion of the GDPR. The GDPR has a wide territorial scope and applies to organisations/companies that have no physical presence in the European Economic Area (EEA)[3].

A key and impressive aspect of the GDPR is that consent must be freely given, specific, informed, unambiguous, and revocable by the data subject. This provision weakens the ability of organisations/companies of using lengthy and inaccessible consent processes to obtain personal data. This provision referred to as the 'purpose limitation principle', clearly specifies that personal data shall be "collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes[4]." The second key prescription is the 'data minimization principle' that specifies that personal data should be "adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed".

The GDPR in Article 4 (11) requires that consent must be 'unambiguous' rather than 'explicit' and such explicit consent is required only for processing sensitive personal data. The Regulation is quite clear in Article 9 (2) that nothing short of 'opt in' will suffice in meeting this requirement. While it provides that for information considered as non-sensitive data, then 'unambiguous' consent suffices and this permits the reliance on implied consent where the individual's actions are sufficiently indicative of their agreement to processing of their data. This demarcation by the GDPR is key in looking at Big Data accumulated on the internet on social media platforms, as the basis for consent has to be clear for personal data.

A key distinction of the GDPR is its expansion of the territorial scope of EU data protection law and applies to both controllers and processors of data of subjects. It provides three important basis of determining its applicability to an organisation or company. First, Article 3 3 (11) states that the Regulation applies where processing takes place in the context of activities of an establishment of a controller or a processor within the EU, regardless of whether or not the processing takes places in the EU); secondly, Article 3 (2) (a) further states that it applies to the processing of personal data of EU based data subjects by a controller or processor that is not established in the EU, where processing relates to the offering of goods and services, whether a payment of the data subject is required; lastly, Article 3 (2) (b) of the GDPR applies to the processing of personal data of EU based data subjects by a controller or processor not established in the EU, where processing relates to the monitoring of their behaviour as far as their behaviour takes place within the EU. Therefore organisations and companies should assess whether any of their EU-based group entities process personal data (as processors or controllers), as these entities will be captured under the GDPR. While Non-EU organisations and companies will now have direct statutory obligations for their activities where they undertake processing activities (either as a controller or a processor) related to the offering of goods or services to data subjects within the EU or monitoring the behaviour of European data subjects. This creates compliance challenges and provides adequate ammunition to push through data privacy and protection of impacted data subjects.

Article 17 of the GDPR that refers to the 'right to be forgotten' as the 'right of erasure' is an impressive part of the instrument. However, contrary to the well-established right to opt-out of direct marketing, it is not an absolute right. Hence as outlined in Article 17 (1) (a) organisations and companies may continue to process data where the data remain necessary for the purposes for which they were originally collected. Based on the provisions of Article 17 (1) (b) organisations and companies retain a legal ground for processing the data under Article 6 and Article 9. Another dimension of the GDPR is its provision in Article 22 (1) that relates to consent requirement for profiling activity that 'produces legal effects' or 'significantly affects' a data subject. The obvious implication of this provision is that the GDPR does not generally mandate consent for the profiling activities of advertising technology companies.

This second key legislation is the California Consumer Privacy Act (CCPA)[5] that came into effect on January 1, 2020, pursuant to a 2018 California ballot initiative responding to the public's desire to protect private

[1] See Article 5 of the GDPR

[2] See Article b24 of the GDPR

[3] See Article 3 of the GDPR

[4] See Article 7 of the GDPR

[5] Although a State Law in the State of California, it has created quite a stir due to the nature of the Californian business climate and the tendency of the State to influence national legislative nehaviour.

information. The CCPA requires the California Attorney General to implement regulations providing more guidance by July 1, 2020. Noteworthy the CCPA grants California consumers various rights regarding their personal data held by businesses[1]. The Act also defines a consumer as "a natural person who is a California resident," this sets the parametre of applicability to any Californian defined as such under state income tax law. Key provisions of the CCPA include granting consumers (1) the right to know what personal information is obtained by companies[2], (2) the right to delete information companies obtain[3], (3) the ability to opt out from the sale of their personal information[4], and (4) the promise that consumers will not be discriminated against for following through with any of these options. Broadly, the CCPA aims to grant California consumers various rights with regard to their personal information held by businesses: the right to know, the right to be forgotten, the right to opt out, the right to equal service and price, and the right to pursue a civil remedy if compliance is not followed by businesses. The main purpose of the CCPA is to give Californians more control over their personal information, by granting them a number of fundamental rights: to know what personal information is being collected about them; to access this information; to know whether it is sold and to whom; to ask that their personal data be anonymized.

While not nearly as comprehensive as the impressive GDPR, the CCPA has a robust data stewardship process hinged on two key factors. The first factor sets the parametres for consideration of the circumstances that may warrant a company to collect personal data[5]. This is important because the primary ways that most jurisdictions allow for collection of personal data is through established privacy policies, which obtain the consent of the consumer before personal data is collected and used. While this approach may be standard across many countries, it lays out a dilemma when reviewing personal data collected on the internet. Practically most internet sites have created a 'no-option' approach that leaves the consumer no choice if he wishes to access some services. While the second factor relates the method of storage of personal data and how it is protected by the collecting organisation or company. There is an implicit duty on the collecting organisation or company to properly manage and protect any personal data in its possession (separate from how it may or may not use that data for any data processing). When this is balanced against the first factor and its weaknesses, there are serious risks beyond just the collection process and it extends to the protection and management of collected data.

It is noteworthy that the CCPA contains no provisions which defines or outlines the actual circumstances under which a company can collect personal data and it also failed to define specific criteria that must be met in collecting personal data. This is a clear difference with the GDPR, which sets and allows for multiple circumstances under which an organisation or company can become legally authorized to collect personal data. While the CCPA concentrates on data stewardship factors that seeks definition of more than mere conditions for legal collection of data.

The CCPA also empowers consumers to seek redress and allows consumers to not only see what data is held by companies[6], but also whom the data has been sold to[7] and the use of the data. This sweeping scope of the CCPA provisions allow consumers to determine which companies have purchased, collected, or used their personal data, at a relatively low cost and within a reasonable period of time. The intendment is to serve as a basis for consumers to ensure the safety and privacy of their data collected by any organisation or company. More so the CCPA also provides two additional means of redress. First, the CPPA grants consumers the right to request companies to delete their personal data, although we must admit that the nine exception provided by the CCPA has weakened this key redress approach. After the right to delete, the CCPA also guarantees a right for consumers to "opt-out"[8] of sale of their information to third parties and this right has not been weakened by the exceptions that invariably impact on the right to delete.

The review of the GDPR and the CCPA has shown that there are four determinants that affect those that access the internet and get their personal data collected. First, whether they have consented, how they consented and whether they can even access those services without consenting. The second determinant is the nature, quantum and storage of the collected data and whether the 'consent' permits access for other users or multiple usage of the collected data. Thirdly, what the data is used for and whether other parties get access to the primary data or analytics extracted from the stored data. Lastly, does the consumer have a right to be forgotten or have his collected and stored data deleted. These determinants are key when building any regulatory framework on data privacy and protection. Yet in our review of the African landscape we must be conscious of the fact that most of the national legislations lack the push to ensure sustainable consent of consumers as well as provide them with the

[11] S. 1798.140 (o) (1)

[2] ibid

[3] S. 1798.105

[4] S. 1798.120 (a)

[5] S. 1798.100 (b)

[6] S. 1798.110 (a) (5)

[7] S. 1798.115 (d)

[8] S. 1798.120 (a)

option of being forgotten and their data deleted. This can be traced from the fact that most of the Big Data companies operating Africa are not even within the regulatory purview of these countries,

The Current Regulatory Approaches to Big Data and Data Privacy in Africa: How Vulnerable is Africa on the Internet?

Incidentally in 2014 the African Union (AU) members adopted the African Union Convention on Cyber Security and Personal Data Protection ("the Convention")[1]. The Member-States through their Ministers affirmed their commitment to the Convention in the African Union Specialized Technical Committee on Communication and ICT Ministerial Declaration (AU/CCICT-2)[2]. This Declaration is by far the most serious commitment by the African continent to set the rules and manage matters related to cybersecurity and personal data protection as it relates to their respective countries. The Declaration called on the African Union Commission (AUC) to develop guidelines on personal data protection for the continent. As part of the implementation strategies and engagements the AUC further requested the Internet Society to help develop the Privacy and Personal Data Protection Guidelines for Africa ("the Guidelines")[3].

While reviewing the Guidelines, we must focus on its prescriptions on managing and protecting personal data collected through online services that are driven by the new digital economy. The Guidelines focuses on building trust in these online services, not just as a factor but rather as a facilitating element in further deepening internet access and related services. The Guidelines set out 18 recommendations related to matters such as trust, privacy and responsible use of personal data. There are also eight recommendations on the roles of Governments and policymakers, Data Protection Authorities (DPAs) and Data controllers and data processors The Guidelines further made another eight recommendations on building multi-stakeholder solutions, wellbeing of the digital citizen and enabling and sustaining measures for data privacy and protection in the continent. The Guidelines sets out for the DPAs roles related increasing legal and regulatory certainty through passage and implementation/enforcement of data protection laws, review of activities of those that collect data and imposing statutory and regulatory sanctions for breaches and possible violations. Meanwhile the Guidelines still expects the DPAs to work closely with stakeholder groups and other DPAs.

The Guidelines expects Data Controllers (DCs) to act responsibly through building of sustainable practices in handling personal data, this responsibility extends to protecting the data subject's interests as well as those of the data controller and partners. This responsibility will create trust in the DCs by the citizen/customer/user, because such trust enhances reputation of the DCs and strengthens consent of the citizen/customer/user. Yet that still leaves open the need for all citizen/customer/user to understand the risks involved in online life, get enlightened on their rights relating to personal data, privacy and autonomy.

However, the AU Guidelines has not currently taken effect as it has, to date, not been ratified by 15 out of the 54 AU member jurisdictions. Nonetheless, we must accept the fact that the AU Guidelines provides a framework for personal data protection in the continent. It also provides a skeletal structure for African countries to make national legislations and ensure a standardized approach that will make the continent more competitive in the online space. Yet the Guidelines has not pushed out national legislations to meet the expectations of the provisions of the Guidelines. This Paper in its review has discovered that only 24 African countries, out of 53, adopted laws and regulations to protect personal data, although the number is slowly rising. The GDPR has been more of an influence for the African legislative action that the AU Guidelines. A clear illustration is the adoption by Nigeria of its first Data Protection Regulation in early 2019. Sadly, the Nigerian Data Protection Regulation (NDPR) has failed the test because there are justifiable doubts on the mandate of the National Information Technology Development Agency (NITDA), to make such a Regulation. NITDA is an Agency of the Nigerian government and its core mandate is the expansion of a 'regulated' digital market.

There are currently 17 countries in Africa that have enacted comprehensive personal data protection legislation, namely Angola, Benin, Burkina Faso, Cape Verde, Gabon, Ghana, Ivory Coast, Lesotho, Madagascar, Mali, Mauritius, Morocco, Senegal, Seychelles, South Africa, Tunisia and Western Sahara1 . In this regard, there are three countries, namely Kenya, Uganda and Zimbabwe, which have already enacted personal data protection legislation, the promulgation of which has not yet been made effective, as the laws are still in the form of bills. Tanzania is another country which is currently in the process of enacting personal data protection legislation.

There are further data on the making of data protection laws and regulations in Africa, if we use a more flexible categorization the situation in Africa is not as bad as the statistics in the last paragraph portray. This Paper in its review of the African continent has discovered that thirty-two of the fifty-five countries in the continent have enacted one form of data privacy laws or the other as at February 2020[4]. A few illustration will help us understand the spread of these laws across the continent. The following countries have Data Privacy laws in place;

[1] https://www.internetsociety.org/resources/doc/2018/personal-data-protection-guidelines-for-africa/#_ftn1 <accessed on 8 March 2020>
[2] https://www.internetsociety.org/resources/doc/2018/personal-data-protection-guidelines-for-africa/#_ftn2 <accessed on 9 March 2020>
[3] The Guidelines were jointly issued by the AUC and the Internet Society in May 2018
[4] This statistics combined both primary and secondary legislations and did not review the implementability of the instruments

Table 5: Statistics of select countries with Data Protection Laws/Regulations[1]

Country	First Passed	Revision/Amendments
Cape Verde	2001	2013
Seychelles	2003	-
Burkina Faso	2004	-
Mauritius	2004	2017
Tunisia	2004	-
Senegal	2008	-
Benin	2009	2017
Ghana	2012	-
Cote d'Ivoire	2013	-
Mali	2013	-
South Africa	2013	-
Nigeria	2019	-
Uganda	2019	
Kenya	2019	-
Egypt	2020	-

There might have been quite a few comparative studies of African regulatory approaches to data privacy issues but one readily comes to mind; a Paper presented to the International Journal of Data Privacy by Graham Greenleaf and Bertil Cottier and titled Comparing African data privacy laws: International, African and regional commitments[2]. The paper states that the most important development in Africa is the adoption of the African Union Convention on Cyber-security and Personal Data Protection, at the African Union's Summit in Malabo, Equatorial Guinea on 27 June 2014[3]. The paper states that that the African Union (AU), requires that State parties who accede to and ratify the Convention should be committed to 'establishing a legal framework' based on its provisions, although this is stated to be 'without prejudice to the free flow of personal data'. The paper further states that Africa is now the first region (in fact a continent) outside Europe to adopt a data protection Convention as a matter of international law, but it will require accession by fifteen states before it is in force. Unfortunately as of February 2020, only five countries have ratified this treaty (Senegal, Mauritius, Guinea (Conakry), Ghana and Namibia) and thirteen more countries have signed but not ratified (Benin, Chad, Comoros, Congo-Brazzaville, Guinea-Bissau, Mauritania, Mozambique, Rwanda, Sierra Leone, Sao Tome & Principe, Togo, Tunisia and Zambia). Ironically Six countries without laws have signed (Comoros, Guinea-Bissau, Mozambique, Sierra Leone, Rwanda, and Zambia), and one without either a law or Bill (Namibia) has ratified. Fifteen ratifications are required for the Convention to enter into force as stated in Article 36, so the current eighteen signatures/ratifications indicates that it is feasible that the Convention may enter into force.

It is clear that African countries understand the challenges of data privacy but it is worrisome that none of them has focused on data mining and Big Data harvesting of their citizens' personal information. Google and Facebook are not incorporated in any African country and operate out of Ireland with footprints all over the African region. This business model is also being used by Facebook to set up projects and even partner governments across the region. This means that they are not within the regulatory purview of the region, therefore national legislations and even continental conventions cannot tailor or influence their behaviour. Thus all the national laws identified by this Paper and the AUC Guidelines lack the teeth to enforce its own prescriptions on the major Big Data players. This is both a vulnerability and regulatory inadequacy.

We can also deduct from our review of the global best practices in Europe and California that the African regulatory landscape has gaping holes in its formulation and enforcement of data privacy regulations. The AUC Guidelines have failed to provide the unifying impetus to push out national legislations as expected; rather several countries are adopting different approaches and processes. A good example is the case of the two West African nations of Nigeria and Ghana. While both are Anglophone they approached the data privacy issue in a different manner. While Ghana passed a Data Privacy Act that sets up a Data Protection Agency; Nigeria pushed out an attempted copy of the GDPR and arrogated its enforcement to an obscure agency mandated to regulate IT. Hence the African approach is disparate and lacks the cohesion envisaged by the framers of the AUC Guidelines. Another key deficiency is the nature of the legislations in Africa have not met the existential threats of the social media and still approaches Big Data as the traditional data protection activity, that focuses on data storage and access.

Conclusion

We have highlighted the gaps in Africa and how vulnerable it remains to data exploitation by companies working

[1] This data is a combination of laws passed by Legislatures and secondary legislations such as the NDPR in Nigeria

[2] Pre-print version submitted to the International Data Privacy Law (OUP) on 22 April 2020

[3] ibid

from other jurisdictions. As a result the privacy of individual Africans have largely been ignored, with governments and businesses gathering up and processing as much data as they can find, to use for whatever purposes they wish. Clearly, except for few exceptions, African governments and parliaments have failed to review, scrutinize and create laws and regulatory arrangements to ensure data privacy and its protection in the Continent. Consequently data and information gathered in Africa are exposed to commercial misuse by Big Data harvesters and their third party partners. Thus Big Data companies like Facebook and Google, located in distant jurisdictions, have gobbled up data, based on consents obtained from customers using obscure terms and conditions, sometimes in the second or third language of the customers, before transferring them to the USA for processing, knowing they are unprotected there. In some cases the data are sold or made available to third party businesses. There are some African countries that have laws on data protection as well as regulatory outfits for enforcement, however, there is remarkably little evidence that the few operational data protection authorities are effective. The rule of law is poorly observed, with the courts offering little, if any, protection against violations of the right.

Despite the push for economic growth from the widespread use of the Internet and the consequent generation of personal data, there has been little interest in creating laws in African countries that would govern the collection of data, its use or protection. There has been very limited recognition of the importance of the right to privacy or of the risks created by data processing, nor any appreciation of the need for consumer confidence in their adoption of e-commerce and digital government. One area where there should have been special concern is the rapid rise of in the use of social media platforms across the continent. This ought to have raised concerns about privacy and security, since data from these platforms can be used for all manner of analysis and marketing. Yet the limited number of data protection agencies and the use of conventional agencies to also enforce data protection laws, has left the door wide open. There is no doubt that there is a need for all African regulators of the communications sectors to work on a regulatory framework for Big Data as quickly as possible, otherwise Africa will remain a game in the Safari of the new digital world.

References

1. Zhuo, R., Huffaker, B., & Greenstein, S. (2019). *The Impact of the General Data Protection Regulation on Internet Interconnection* (No. w26481). National Bureau of Economic Research.
2. Determann, L. (2019). Letter re California Consumer Privacy Law Corrections.
3. Dai, H. N., Wong, R. C. W., Wang, H., Zheng, Z., & Vasilakos, A. V. (2019). Big data analytics for large-scale wireless networks: Challenges and opportunities. *ACM Computing Surveys (CSUR)*, *52*(5), 1-36.
4. Ghani, N. A., Hamid, S., Hashem, I. A. T., & Ahmed, E. (2019). Social media big data analytics: A survey. *Computers in Human Behavior*, *101*, 417-428.
5. Farboodi, M., Mihet, R., Philippon, T., & Veldkamp, L. (2019, May). Big data and firm dynamics. In *AEA papers and proceedings* (Vol. 109, pp. 38-42.
6. Farboodi, M., Mihet, R., Philippon, T., & Veldkamp, L. (2019). Replication data for: Big Data and Firm Dynamics.
7. Mihet, R., & Philippon, T. (2019). The Economics of Big Data and Artificial Intelligence. *Disruptive Innovation in Business and Finance in the Digital World (International Finance Review, Vol. 20), Emerald Publishing Limited*, 29-43.
8. Hossain, E., Khan, I., Un-Noor, F., Sikander, S. S., & Sunny, M. S. H. (2019). Application of big data and machine learning in smart grid, and associated security concerns: A review. *IEEE Access*, *7*, 13960-13988.
9. Veldkamp, L., & Chung, C. (2019, October). Data and the Aggregate Economy. In *Annual Meeting Plenary* (No. 2019-1). Society for Economic Dynamics.
10. Ramadorai, T., Uettwiller, A., & Walther, A. (2019). The market for data privacy. *Available at SSRN 3352175*.
11. arrière-Swallow, Y., & Haksar, V. (2019). The Economics and Implications of Data.