

Big Data Storage in Zimbabwean Financial Institutions: A Review of the Need for Big Data

Hillary Marufu^{1*}, Newton Mavu^{2**}

* Faculty of Information Technology, Zimbabwe Open University

DOI: 10.29322/IJSRP.13.03.2023.p13533

<http://dx.doi.org/10.29322/IJSRP.13.03.2023.p13533>

Paper Received Date: 15th February 2023

Paper Acceptance Date: 15th March 2023

Paper Publication Date: 24th March 2023

Abstract- Data in this millennium is growing rapidly and exponentially each and every year. It is increasingly difficult to handle the complexity and humongous amounts of data being produced. Data generated is in various types from social media to company records. This is the basis of big data. Hadoop framework plays a pivotal role in processing and storing Big Data. It provides solutions that are speedy and cost-effective for Big Data and is used in various sectors like health care, education, telecommunication, insurance and social media. This is a very big advantage. Zimbabwe is a country in the southern part of African. It is not spared from the rapid growth in data and the need to process and store large amount of data. This paper focuses on the review of the need for Big Data in Zimbabwe over its Financial Institutions.

Index Terms- Big Data, Mobile Money, Ecocash, 3Vs.

I. INTRODUCTION

Enormous measures of data are being prepared and dissected day by day. Challenges are created as the data amount being processed and analysed increases. As the amount of data increases and so does the complexity due to its diversity. Data will come in the form of structured, semi-structured and unstructured data (michaelgramlich website). Such data is described as Big data. Big Data allows an organization to create, collect, retrieve, manage, analyse data and make decisions that are remarkable (Kanyan at el 2018). Volume, velocity and variety are known as the 3 Vs. Increasing organisational adaptation is fuelled by big data.

II. BIG DATA

Data comes in gigantic sums. Estimated in petabytes. As it increases exponentially, this data will soon be measured in zettabytes. Data change frequently and drastically through multiple sources like legacy data and streamed data from sensor,

III. PROBLEM STATEMENT AND OBJECTIVES

There is a lot of unprocessed data being thrown away or not being utilised by financial institutions. This data could be used to give

social media, traditional file records, cellular technology and many more. (Kanyan 2018). All this data need to be analysed for the purposes of decision making. Decisions can either be descriptive or predictive.

Data in Zimbabwe and worldwide is growing quickly each and every year. It becomes very problematic to process this complicated and enormous volume of data using traditional applications/tools (Kanyan 2018). Big data is the solution to such challenges brought by attempting to process complex and huge data. Big Data is an assortment of huge amounts of historical and imperative data. Real fact can be used to make decisions rather than the use of perceptions. This becomes the most important asset of every organisation and can be used intelligently for business (Bathal 2019).

The data to be either created, collected, stored, retrieved, managed or analysed for decision making is processed in Hadoop. Hadoop framework plays a pivotal part in storing and processing Big Data. It provides a speedy and economical solution for Big Data (Kumar 2018). This allows the huge amounts of data to be processed cost-effectively and in a speedy manner. Due to the nature of data being produced, big data need to be able to process most, if not all types of data. Data comes in a lot of variety. It can either be structured, semi-structured or unstructured data. The 3Vs are three defining aspects or dimensions of big data. Volume refers to the amount of data, variety refers to the number of types of data and velocity refers to the speed of data processing (Wigmore 2020).

Big data allows businesses several types of analysis on their data ranging from descriptive analysis to predictive analysis. All these allow businesses to make informed decisions because the organisation can understand where the business is coming from and where it can be in the future. Businesses in Zimbabwe can harness this technology to allow them to make decisions that affect the organisation. Most organisation that needs big data are still using traditional applications/tools to process data like a database. These applications/tools are barely functional for they only provide a fraction of the advantages offered with big data.

institutions competitive edge or benefit by giving clients service that fit their needs. Data is not processed enough to give financial institutions this advantage.

Objectives

1. to establish the reasons why financial institutions are not using Big Data
2. to find out if financial institutions need Big data
3. to find out why there is a big gap between knowing and implementing big data

IV. DISCUSSION

A. Financial institutions

A financial institution is a company involved in the business of dealing with financial and monetary transactions such as deposits, loans, investments, and currency exchange (Investopedia website). A financial institution is an arbitrator amongst clients and the capital or the debt markets providing banking and investment services (myaccountingcourse website). Financial institutions are organisations that process monetary transactions, including business and private loans, customer deposits, and investments (study website).

Financial institutions are basically sub-divided into two types: banking financial institutions and non-banking financial institutions (fbs website). Financial institutions include commercial banks and central banks whose primary role is to accept deposits and make loans. Non-banking financial institutions include investment banks, insurance companies, finance firms, leasing companies, etc. Banks are one of the most typical and well known types of financial institutions.

Zimbabwean financial Institutions have several platforms they use. Each online platform has certain clients it services. Unstructured Supplementary Service Data (USSD) platform is used to service clients who stay in rural areas or areas which do not have internet services. USSD is a protocol that communicates with Service providers' servers which is used by GSM cellular phones. Clients without smartphones or data/internet can transact using this type of banking.

B. Tools used by Financial Institution

Zimbabwean financial Institutions now have mobile phone applications (apps) that are open every time of the day to assist clients. The most common being android based apps. These assist in the day to day running of the said institutions.

Web services are utilised but they could either be social media or applications designed by banks to fit this purpose. Clients are given unlimited access to their accounts using Internet Banking. This is an electronic solution which is very secure. Internet Banking offers a safe and secure innovation for clients to do their real time banking any time.

The fact is social media platforms like Whatsapp, Twitter, Instagram and Facebook are evolving into professional tools and critical business tools as evident by a lot of banks moving to introduce Whatsapp and Facebook Banking. The ZB WhatsApp Platform for banking is a user friendly and easy to use service that lets customers to do their everyday banking on their WhatsApp chat application. Alisa is a first step in the use of Artificial Intelligence and is our first ever innovative WhatsApp banking chatbot that uses machine learning and cognitive computing technologies to provide you with various banking services and transaction capabilities (First capital bank website). FBC WhatsApp Banking is an easy-to-use secure channel which allows you to conduct basic banking functions on the go (FBC website).

Zimbabwe's financial sector is made up of Banks and Mobile Money services like EcoCash, One Money and TeleCash. Each generates a lot of data which can be used to give each institution an advantage. Just a few months after launch, the results are impressive. 2.3 million Zimbabweans have registered for EcoCash mobile money accounts, outnumbering all of Zimbabwe's traditional bank accounts combined (Levin 2014). More than 1 million of the said accounts were active and pushed about US\$200 million over the EcoCash platform monthly. The amount symbolises an equivalent to 22% of Zimbabwe's GDP if annualised. EcoCash essentially changed the financial landscape in Zimbabwe because of ambition coupled with effective implementation. All these translate to data being generated.

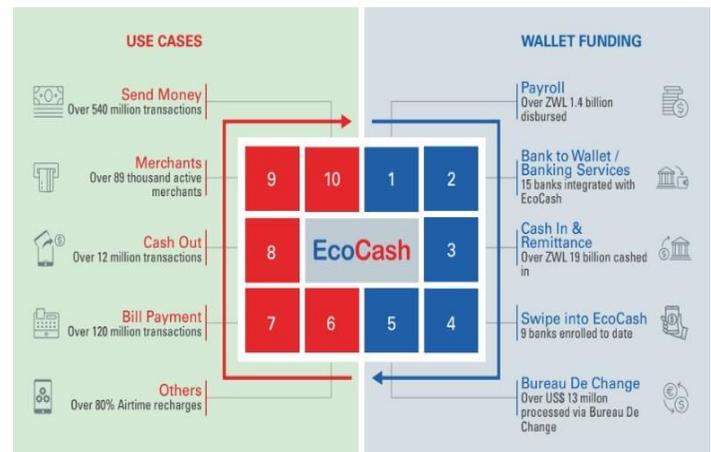


Fig.1 Ecocash transaction (levin 2014)

The above diagram shows the total number of transactions, total amount of money transferred through its platform. From the diagram, it is evident that there is a lot of data and patterns being produced. Huge sums of data are generated from several services offered by EcoCash.

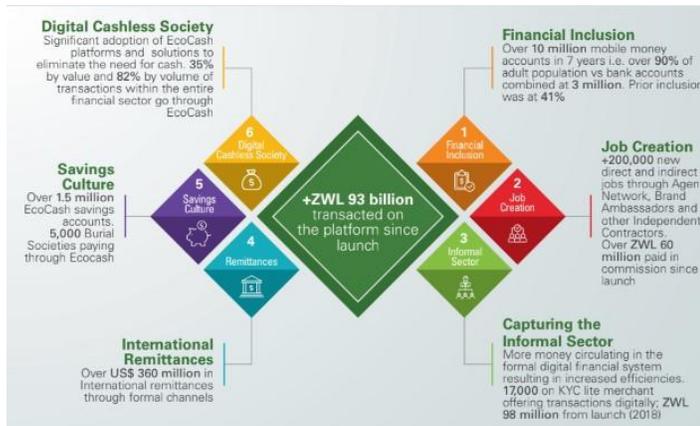


Fig 2. Total Amount of money thought EcoCash (Techzim)

The above diagram shows the total amount of money that has been transferred to and from the EcoCash platform. From the diagram, we can easily notice the amount of data that EcoCash alone is generating.

The worth of Big Data to the monetary administration industry is colossal. Quick expense decreases openings lie in fraud and approvals administration, while account the executives can be improved by upgraded client understanding. Taking a more extended term view offers Banks the capability of critical new income streams. Companies that incorporate data into their operations show productivity rates much higher than those of their peers (December 2012 | Jonathan Gordon, Manish Goyal, Tim McGuire, and Dennis Spillecke)

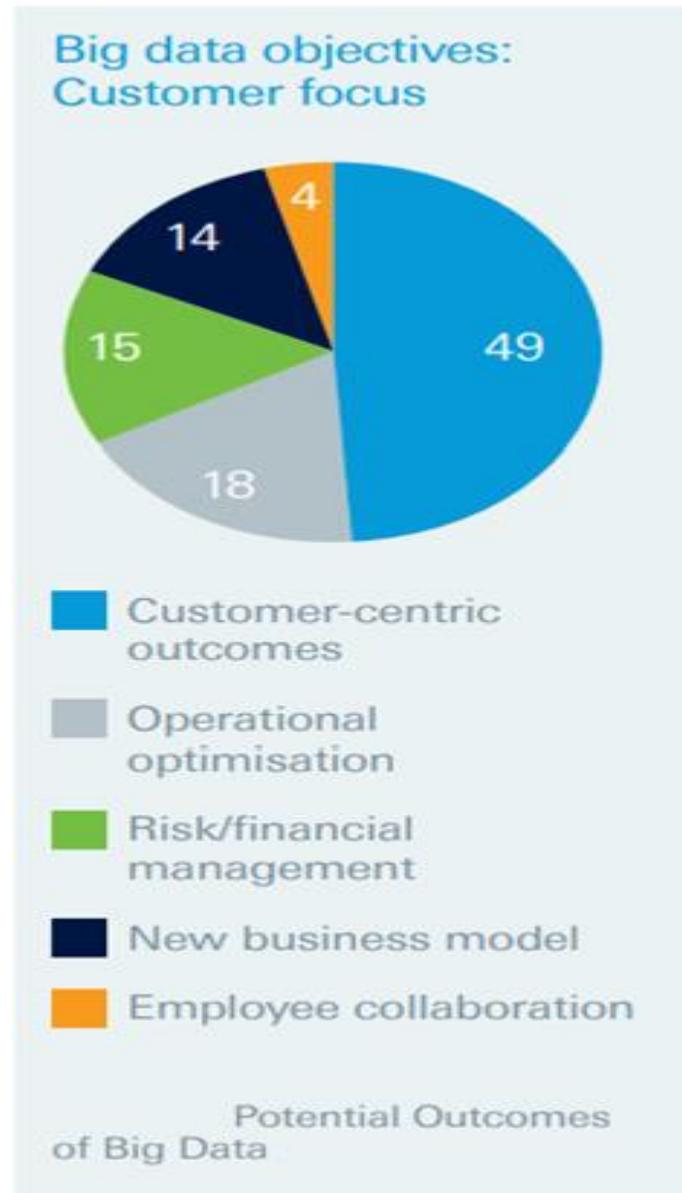


Fig 3 Potential Outcomes of Big Data (Jonathan 2012 at el)

C. Big Data advantages

Big Data offer these following advantages to financial Institution all over the world:

Overall, managing and analysing big data is essential to bank success as indicated by the perceptions of 62% of the banks. However, only 29% report that they are currently extracting enough commercial value from data.

According to Oracle, customers are searching for personalised and tailor made experience. 84% of executives surveyed confirmed with this notion. Offering clients such tailor made solutions can increase the annual revenue by 18% according to the report. Due to the nature of the technology era, the role of banks is changing. Important issues can now be focused on. Clients powered by mobile application, can in turn have more continuous online

access to their accounts. With the use of smart phones, clients can now perform any operation.

Marketing budget is reduced by 15-20% if decisions are made using big data. This also means that decision made using big data are better. Marketing overall budget is about 8% of the total budget spent by banks. Harnessing big data seems like a very ideal chance to save as well as creating marketing strategies targeted revenue generation.

In addition to optimizing internal processes, big data and AI are able to identify and prevent fraud amongst the organisation's employees. The financial institutions process huge amounts of data to analyse individual client behavioural patterns. Once a pattern has been established, then any behaviour different from the known pattern is flagged and reported. This identifies and prevents fraudulent activities.

Even though big data has the potential to progress the manner in which financial services are delivered and provided, challenges are presented for the successful use of big data in financial institutions. If these challenges are not dealt with, they could have adverse effects to the client and society as a whole. Every other sector is not spared from these challenges since they have the potential to affect each and every sector.

The use of big data in financial institutions includes building huge data sets which incorporate private data. Examples of this private data include bank card numbers, salaries or income, asset details, spending details and so much more. Outsider can used by financial institutions, for their benefit, to gather customer information with the aim of expanding their own data. Financial institutions may be tempted to sell their privately collected or proprietary data to companies, like retails and marketing. Such companies may be interested because they want understand spending behaviours of various social groups as well as demographics groups. Data privacy becomes very crucial. Financial institution should be straight forward and immediately uncover what data they have on purchasers and what they indent to do with it. In certain conditions, purchasers may be allowed the ability to restrict companies from selling explicit data.

Financial institutions have to protect the data once it is gathered. Once it lands in the wrong hands, it may be used for extortions, scams, identity theft as well as fraud. An increase in the budget for IT security need to be effected due to this concern. More and more data breaches have occured hence security need to also improve. On average, 25 % of all financial institutions report a data breach as of 2018. 50 percent of the 25% confirm that the breaches have occurred at least once before. Steps need to be taken by financial institutions to improve their data security. This can be archived by either encrypting, masking, strengthening and or access monitoring the available data.

D. Big Data Challenges

However, Big Data comes with many challenges, relating mainly to its complexity. A portion of the essential works of art identifies

with how banks can comprehend and utilize Big Data streams that arrive in an unstructured organization, for example, text or video or how they can catch the main data as it occurs and convey the information to the ideal individuals continuously. The other Big Data challenge identifies with how banks will actually want to store, dissect and recover Big Data given its size contrasted with the restricted computational limit of most foundations' present frameworks. Added to these intricacies are various difficulties, going from legitimate issues and worries around information protection, security of access and sending of frameworks according to iobz

To address the difficulties confronting them, the financial services players need another arrangement of bound together data platforms that join all information into an intertwined data texture that will permit them to store, oversee, apply and investigate data at the right speed, at an adequately enormous scope, and with a lot of unwavering quality.

Today banks presently know this, and accordingly, they are putting progressively more noteworthy accentuation on building merged stages that are fit for running all way of insightful applications. Harnessing artificial intelligence, empowering the learning patterns and afterwards determining experiences autonomously, these intelligent applications immensely increment an opportunity to market of new items. They also support current financial modelling and analysis programs and other products currently existing in the market

According to iobz

Cloud computing also poses benefits and risks. Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction (Mell 2011). Cloud computing is the current hype of the Information Technology (IT) industry. Cloud computing is accessing on demand computing resources via the internet, such as software, storage and even infrastructure (Goundar 2012). This cloud model is composed of five essential characteristics, three service models, and four deployment models (Mell 2011).

Even though there are plenty of benefits associated with the use of big data, financial institutions opt to store their data in house because they argue that keeping data online increases the risk of hacking. If a reputable hosting cloud computing company is used then risk of hacking reduced not eliminated. This has a hefty cost associated with. So here has to be a cost-benefit analysis to weigh the pros and cons of cloud visa vee local servers.

If big data is used to make decision without appropriate due diligence, it may lead to bias. This type of bias called algorithm bias. There are various stages which can be presented by human inclination which may arise due to the process of analysing big data with the use of algorithms. These stages range from picking

the factor in the model and their general significance to the manner in which samples are collected and results decoded.

The models being used can introduce bias inadvertently even when they are not programmed to do as such. For example, models created and trained using historical data under different conditions or discrimination can create involuntary bias (loan applications that where by applications are discriminated due to race or gender).

Financial institutions may not just apply Big Data to offer items that purchasers may interest in, however could likewise tailor costs to separate the most extreme benefit from every customer. For example, greater interest rates may be charged to minorities after establishing that these sectors barely use cost comparisons by algorithm –based lending. Algorithm-based lending may result in higher interest rate if groups live in areas where there is limited financial services (little to no competition). Identifying vulnerable groups, by financial institutions using big data that are in great need for some quick money and approve them financially risky products, such as pay day loans.

The bigger the data, the greater the risk. It is evident that financial institutions need to ensure the collected user data is always safe and secure. According to Information System Audit and control Association (ISACA), only 38% of all organisations across the globe are ready and capable of handling threats if they arise. For this reason, cyber security remains one of the most burning and important issues in financial institutions.

Better analysis is not necessarily brought about by more data. It may become so difficult to know where to start and what to look for if the data set available is too large. Useful and needed data may be hidden beneath tons of irrelevant and useless data. Even though big data analytical tools are good and establishing relationships between variable, they may fair well distinguishing correlation from actual causation. Natural experiments and randomised control trials are traditional techniques that may infer causation. In addition, big data is not accurate all the time. Its heavily depended on the provided data.

According to Experian.com, costs for a system development can be overwhelming, ranging from \$5 million - \$100 million. These costs are too great to be incurred by a Financial Institution in Zimbabwe. The investment will cost more than it is meant to be beneficial.

Countries in general, would greatly benefit from the development of an information system. Such information system should permit financial and economic data collection but various involve agencies and stake holders. This data must be matched and more systematically. Building these systems may be challenging because they take some time and cost a lot to develop but the all of it might be worth it. The developmental benefit are eventually realised regardless of the cost. The financial sector understanding, may improve the value and quest for building a more stable, competitive and effective financial sector.

<ul style="list-style-type: none"> Lack of early stage funding Limited VCs and growth investors Thin capital markets Few linkages with formal FSPs 	<ul style="list-style-type: none"> No focal point Limited international access Limited exposure to market opportunities 	<ul style="list-style-type: none"> Unstable currency Bank financing gap deficit Inflation and BoP deficit
Training and skills	Hubs, incubation and acceleration	Policy environment
<ul style="list-style-type: none"> Limited fintech-specific education Lack of entrepreneurship training Few mentorship opportunities 	<ul style="list-style-type: none"> Lack of dedicated incubators Limited acceleration facilities Hubs lack scale and international links 	<ul style="list-style-type: none"> Tendency towards risk aversion Regulatory frameworks ill-suited to fintech Transaction tax on mobile money
Financial sector	Infrastructure and data	Access
<ul style="list-style-type: none"> Capital concentrated in banks Dominance of single player in mobile money Limited innovation or collaboration 	<ul style="list-style-type: none"> Little culture of sharing through APIs Hubs lack scale and international links Lack of publicly available financial data 	<ul style="list-style-type: none"> Limited smartphone access High data costs and poor connectivity

Fig 4. Challenges faced by Zimbabwean Financial Technology

V. CONCLUSION

Big Data offers a lot of advantages which any business can harness. There are also many challenges associated with the use of Big Data in Zimbabwean financial institutions. Most of them revolve around infrastructure and investment. With the proper investment in infrastructure, Big Data can be fully implemented. At the moment institutions are making do with the use of ordinary databases which are not giving them the best competitive advantage. Although the advantages outweigh the disadvantages in terms of number but they out way the Advantages in terms of magnitude. Due to the cost infrastructure required for setting up and running Big Data Technology Financial institutions have managed to do without Big Data.

REFERENCES

- [1] Big Data: Hadoop framework vulnerabilities, security issues and attacks Gurjit Singh Bhathal *, Amardeep Singh 2019
- [2] Big Data Storage in Hadoop: A Review of Security Issues and Threats Kavita K. Kanyan1, Er. Ritika Mehra2 2018
- [3] IJSER © 2017 <http://www.ijser.org> Challenges and Security Issues in Implementation of Hadoop Technology in Current Digital Era Dr. Vinay Kumar, Ms. Arpana Chaturvedi 2018
- [4] <https://www.economicshelp.org/blog/390/inflation/hyper-inflation-in-zimbabwe/>
- [5] Big Data Storage in Hadoop: A Review of Security Issues and Threats Kavita K. Kanyan1, Er. Ritika Mehra2
- [6] Impact of Big Data Analytics on Banking Sector: Learning for Indian Banks Utkarsh Srivastavaa , Santosh Gopalkrishnanb * 2015
- [7] <https://www.techzim.co.zw/2021/02/ecocash-clocks-zw193-billion-in-lifetime-transactions-according-to-cassava-2020-report/>
- [8] <https://www.techzim.co.zw/2014/03/opportunities-for-big-data-in-telecoms/> <https://www.iobz.co.zw/big-data-new-frontier-finance>
- [9] <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/505891573224492672/using-big-data-to-expand-financial-services-benefits-and-risks>
- [10] Using Big Data to Expand Financial Services: Benefits and Risks Facundo Abraham Sergio L. Schmukler José Tessada No. 26, Nov 2019
- [11] Forbes. 2015. “Turkey’s Garanti Bank Captures Elusive Millennials.” June 15.
- [12] Carrière-Swallow, Y., and V. Haksar. 2019. “The Economics and Implications of Data: An Integrated Perspective.” Departmental Paper No. 19/16, International Monetary Fund.
- [13] <https://www.coursehero.com/file/73967409/8-Application-of-Predictive-Analytics-in-Financial-Institutionspdf/>
- [14] <https://www.sciencedirect.com/science/article/pii/S2590005619300025>
- [15] <https://www.irjet.net/archives/V5/i4/IRJET-V5I41074.pdf>
- [16] <https://www.ebusinessweekly.co.zw/big-data-new-frontier-of-finance/>

- [17] <https://documents.worldbank.org/curated/en/405521588701013612/text/Development-Digest-Navigating-the-New-Normal.txt>
- [18] <https://olc.worldbank.org/system/files/Using-Big-Data-to-Expand-Financial-Services-Benefits-and-Risks.pdf>
- [19] Big ambition meets effective execution: How EcoCash is altering Zimbabwe's financial landscape Author: Phil Levin
- [20] <https://www.oracle.com/us/industries/oracle-era-ready-anything-2969053.pdf>
- [21] Marketing & Sales Big Data, Analytics, and the Future of Marketing & Sales March 2015
- [22] <https://easternpeak.com/blog/big-data-in-the-banking-industry-the-main-challenges-and-use-cases/>
- [23] Bigdata : how it can become a differentiator. Deutsche Bank Global Transaction Banking
- [24] "Big Data: Big Opportunity In Banking... Or Big B.S.?" The Financial Brand Marketing Insights for Banks Credit Unions. N.p., n.d. Web. 19 Aug. 2014.
- [25] <https://www.finextra.com/blogposting/20446/the-role-of-big-data-in-banking--how-do-modern-banks-use-big-data>
- [26] <https://www.business2community.com/cybersecurity/5-alarming-cyber-threat-statistics-vulnerable-business-02029063>
- [27] <https://www.experian.com/assets/consumer-information/white-papers/cis-bigdata-tl.pdf>
- [28] <https://www.investopedia.com/terms/f/financialinstitution.asp>
- [29] <https://www.zimshoppingmalls.com/whatsapp-banking-in-zimbabwe/>
- [30] The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology Peter Mell Timothy Grance
- [31] <https://www.michael-gramlich.com/what-is-structured-semi-structured-and-unstructured-data/>
- [32] https://www.mckinsey.com/assets/dotcom/newsletters/topten/2015_toptenBusinessTechnology.html
- [33] <https://www.myaccountingcourse.com/accounting-dictionary/financial-institution>
- [34] <https://fbs.com/glossary/financial-institution-fi-30>
- [35] **Evolution Of Financial Institutions In Zimbabwe** Kelvin K Tafirenyika
- [36] <https://study.com/academy/lesson/types-of-financial-institutions-definition-examples-roles.html>
- [37] <https://www.fbc.co.zw/banking/faqs/fbc-whatsapp-banking-noku>
- [38] <https://www.zb.co.zw/whatsappbanking>
- [39] <https://firstcapitalbank.co.zw/personal/whatsapp-banking/#>
- [40] https://onlinefirstcapitalbank.co.zw/corp/AuthenticationController?_STAR_T_TRAN_FLAG_=Y&FORMSGROUP_ID_=AuthenticationFG&_EVENT_ID_=LOAD&FG_BUTTONS_=LOAD&ACTION.LOAD=Y&AuthenticationFG.LOGIN_FLAG=1&BANK_ID=01&LANGUAGE_ID=001
- [41] https://www.researchgate.net/publication/332979373_Chapter_4_-_Understanding_Cloud_Computing [accessed Aug 25 2021].
<https://whatis.techtarget.com/definition/3Vs>

AUTHORS

Hillary Marufu – Master of Science in Information Systems, National University of Science and Technology, Zimbabwe.

Mavu Newton – Master of Science in Information Systems Management, Midlands State University, Zimbabwe.