ZCAS UNIVERSITY

SCHOOL OF COMPUTING, TECHNOLOGY & APPLIED SCIENCES

CCS5812 – COMPUTING PROJECT

FINAL YEAR

TRANSPARENT ELECTRONIC PUBLIC PROCUREMENT MONITORING SYSTEM THROUGH THE USE OF MACHINE LEARNING AND ANALYTICS

Jack Bwaaza

A Final Year Research Project submitted in partial fulfilment of the requirements for the degree of
Master of Science in Information Technology

ZCAS University
2024

DECLARATION

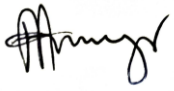
Name: **Jack Bwaaza**

Student Number: **G09077**

I hereby declare that this final year research project is the result of my own work, except for quotations and summaries which have been duly acknowledged.
Plagiarism check: 6 %

Signature:  Date: 27th December 2024

Supervisor Name: Dr. Chiyaba Njovu

Supervisor Signature: 

Date: 27th December 2024

ABSTRACT

E-procurement can largely enhance the processes of public procurement by allowing more competition, reducing corporate transaction costs through digital technologies, and instilling confidence in decision-making within public institutions. These systems, when integrated with machine learning and analytics, could go one step further in the detection of fraudulent behaviors, thus enabling data-driven decisions.

This paper discusses the adoption of sophisticated technologies, such as Machine Learning and Analytics with the view to increase efficiency, accountability, and transparency in public procurement in Zambia. It also analyzes the current electronic procurement system operated by the Zambia Public Procurement Authority (ZPPA) in order to identify the systems shortcomings in which Machine Learning and Analytics could provide substantial support. The proposed system makes use of Machine Learning for corrupt-risk analysis of market prices, classification of tenders and bids, data analytics for predictive procurement strategies, and better decision-making.

This research goes on to assess such technologies in terms of their impact on the performance and maturity of electronic public procurement processes and consequently marks those as potential solutions to inefficiencies and inequities. Key highlights reviewed of 178 Procuring Entities, 5,356 Tenders, and 8,709 Bids bring out some interesting patterns: 35.68% of the tenders were single bid closing, 6.05% had huge price gaps in bids submitted by only two suppliers, and 3.58% had price outliers in cases of three or more bids. These anomalies were identified through Machine Learning-driven methods of percentage difference and interquartile range analysis to make sure that procurement practices are closely monitored and evaluated at large.

Beyond uncovering system inefficiencies, the study underscores the broader impact of Machine Learning and Analytics in optimizing sourcing, transaction management, payment processes, and bid evaluations. These technologies have been pivotal in elevating the operational maturity of Zambia's e-Procurement system. The findings demonstrate that a strategically aligned, data-driven adoption of ML and Analytics not only delivers notable performance improvements but also enhances trust, fairness, and integrity within public procurement operations.

ACKNOWLEGMENT

First and foremost, I give all glory and honor to God Almighty, whose guidance, grace, and strength have sustained me throughout this journey of research and discovery. Without His blessings, none of this would have been possible.

I am deeply grateful for the support and guidance I have received during this process. My heartfelt thanks go to my supervisor, Dr. Chiyaba Njovu, for his invaluable expertise and direction. His constructive feedback and insightful comments have been instrumental in shaping and refining this work to its completion.

I also wish to acknowledge the contributions of Mr. Hamulilo Moses Handia, whose ideas and perspectives provided new angles and deeper insights into this project, enabling me to approach it with greater clarity and innovation.

Furthermore, I would like to express my gratitude to my fellow classmates and lecturers at the School of Computing, Technology & Applied Sciences (ZCAS University), whose encouragement and shared knowledge created an inspiring and collaborative environment that significantly contributed to the success of this research. To each and every one of you, I extend my sincere thanks for your unwavering support and encouragement.

DEDICATION

This work is dedicated in the loving memory of my late father, Mr. Gideon Mwanza, whose legacy continues to inspire me.

I further extend my heartfelt gratitude to my beloved wife, my daughter, my mother, and my younger brother. Their unwavering encouragement and support have been a constant source of strength. Their understanding and belief in my aspirations have been invaluable throughout this journey and will continue to guide me in all my future endeavors.

Table of Contents

1. CHAPTER 1 Introduction

1.1 Overview

Corruption in public procurement can significantly undermine the integrity and efficiency of government spending, affecting all stages of the procurement process and taking various forms (Transparency International, 2020). The complexity of procurement procedures, coupled with the need for a high level of technical expertise at each stage, often makes corruption difficult to detect and prevent (OECD, 2016). According to the Organization for Economic Co-operation and Development (OECD), "procurement is particularly vulnerable to corruption because of its complexity and the large amounts of money involved" (OECD, 2016, p. 5).

In response to these challenges, electronic procurement (e-procurement) systems have emerged as a promising tool for enhancing transparency, accountability, and efficiency in public procurement processes (World Bank, 2018). E-procurement systems can "*improve access to information, reduce opportunities for corruption, and increase the efficiency of the procurement process*" (World Bank, 2018, p. 2).

In Zambia, like many other countries, the government has recognized the importance of leveraging technology to modernize procurement practices and combat corruption. As part of this effort, an electronic procurement system has been developed to serve the needs of the Zambian people (Zambia Public Procurement Authority [ZPPA], 2019). However, despite its implementation, significant lapses persist, particularly in areas such as reporting and data analysis. These shortcomings hinder the system's ability to fully realize its potential in enhancing transparency and efficiency in public procurement (Transparency International Zambia, 2021).

Therefore, there is a pressing need for in-depth analysis and evaluation of electronic public procurement systems in Zambia to identify weaknesses, address gaps, and optimize the system's functionality. This project seeks to fill this gap by examining the effectiveness of electronic procurement systems in enhancing transparency and efficiency in Zambia. By investigating the current state of the system, identifying areas for improvement, and proposing recommendations for enhancement, this research aims to contribute to the ongoing efforts to strengthen governance and accountability in public procurement processes.

Furthermore, understanding the specific challenges and successes experienced in Zambia can provide valuable lessons for other African countries attempting to implement or refine their own e-procurement systems. The insights that will be gained from this research could help to formulate best practices that ensure the sustainability and scalability of e-procurement systems in Africa. Ultimately, this study endeavors to highlight the critical role of technology in fostering transparent, efficient, and corruption-free public procurement environments, contributing to broader objectives of good governance and economic development.

1.2 Background

Electronic Public Procurement (e-Procurement) has emerged as a critical component in modernizing public sector procurement processes, aiming to enhance efficiency, transparency, and accountability. As governments worldwide increasingly adopt e-Procurement systems, the need for robust mechanisms to ensure transparency in these processes becomes paramount (World Bank, 2018). Transparency in public procurement is essential to foster trust among stakeholders, prevent corruption, and ensure the effective utilization of public funds (OECD, 2016). According to the OECD, "transparency in public procurement is a cornerstone of good governance" (OECD, 2016, p. 7).

Traditional procurement methods have been criticized for their opacity and susceptibility to fraudulent activities, leading to significant financial losses and compromised project quality (Transparency International, 2020). The advent of e-Procurement systems offers a promising avenue to address these challenges by leveraging technology to streamline processes and improve oversight (European Commission, 2015). E-Procurement "reduces the potential for corruption by increasing transparency and accountability" (European Commission, 2015, p. 10).

Machine Learning (ML) and Analytics present innovative tools that can further enhance the capabilities of e-Procurement systems. By applying ML algorithms and advanced analytics, it is possible to uncover hidden patterns, detect anomalies, and predict outcomes, thereby providing actionable insights that can drive transparency and accountability (Bose, 2019). These technologies can assist in identifying and mitigating fraudulent activities, optimizing procurement processes, and ensuring that procurement decisions are data-driven and objective (Bose, 2019).

Through this project, we aim to contribute to the body of knowledge on public e-Procurement and provide practical insights that can aid policymakers, procurement officials, and industry stakeholders in implementing more transparent and efficient procurement practices. By leveraging ML and Analytics, this study envisions a future where e-Procurement systems are not only efficient but also exemplify the highest standards of transparency and integrity.

This project further investigates the application of ML and Analytics to enhance transparency in e-Procurement systems within Zambia's public sector. The study aims to provide a comprehensive understanding of how these technologies can be integrated into existing e-Procurement frameworks to improve transparency. By conducting a thorough review of existing literature and analyzing current e-Procurement practices, this research seeks to identify best practices, challenges, and potential solutions.

(Global Procurement Partnership Multi-Donor Trust Fund, 2023)

1.3 Problem Statement

In Zambia's public procurement process, particularly within the supply of goods and services, several critical challenges hinder transparency, accountability, and efficiency. Firstly, the dispersion of data across multiple non-centralized databases complicates access and analysis. Information crucial for assessing beneficial ownership and scrutinizing procurement practices is scattered, impeding comprehensive oversight and evaluation.

Moreover, the lack of transparency and accountability within the procurement process exacerbates these issues. The absence of robust mechanisms for disclosing procurement-related information fosters opacity, leaving room for malpractice and corruption. This opacity not only undermines public trust but also obstructs efforts to identify and rectify irregularities in procurement procedures.

Furthermore, the deficiency in analysis tools compounds the problem. Existing systems fail to provide users with adequate means to analyze procurement data comprehensively. Without such tools, stakeholders lack the capacity to conduct thorough assessments, hindering their ability to identify trends, anomalies, and potential areas for improvement.
Consequently, the combination of dispersed data sources, deficient transparency and accountability measures, and the absence of analysis tools presents a formidable challenge to effective governance and oversight in public procurement within Zambia's supply industry. Addressing these challenges is imperative to foster greater transparency, accountability, and efficiency in public procurement processes, ultimately enhancing trust, integrity, and the effective utilization of public resources.

1.4 Aim and Objectives

This project will investigate the application of Machine Learning (ML) and Analytics to enhance transparency in Zambia’s electronic Public Procurement (e-Procurement) system. Additionally, this study will conduct a thorough review of existing literature to counter-check findings and insights provided by other researchers on the topic, ensuring a comprehensive understanding of the current state of electronic procurement systems and potential avenues for improvement within Zambia's public sector.

- Analyze existing e-Procurement systems and identify areas where ML and Analytics can improve transparency.
- Use machine learning to analyze and estimate market price to determine corruption risk.
- Analyze and categorize tenders and bids.
- Explore the use of data analytics for predictive procurement strategies and data-driven decision making.
- Assess the impact of ML and Analytics on efficiency, fairness, and accountability in public e-Procurement processes.

1.5 Scope and Limitation

- **Scope**

Various dimensions of the study are:

- ✓ Exploring ways to make it better and how using Machine Learning (ML) and Analytics can increase accountability. Applying ML and Analytics to make government procurement processes more accountable, efficient, and transparent by using advanced algorithms.
- ✓ Application of ML algorithms in advanced analytics to increase the transparency, accountability, and efficiency of the public procurement process.
- ✓ Development of analysis and visualization tools to portray the trends, anomalies, and areas of improvement.
- ✓ Design fraud detection and prevention flags in the system; data-driven procurement decisions with objectivity.
- ✓ Engage various institutions such as ZPPA, PACRA and other industry stakeholders in implementing more open and efficient procurement methods.

- **Limitations**

Limitations include, but are not limited to:

- ✓ Limited access to complete and trustworthy data could hinder the performance and accuracy of machine learning models.
- ✓ Combining machine learning and analytics with current public e-Procurement systems can sometimes be difficult.
- ✓ Full support and participation from all involved parties, including ZPPA/PACRA, is necessary. However, this could prove to be problematic.
- ✓ There may be worries about whether the suggested solutions can work well in the long run and on a larger scale.
- ✓ Adapting to the regulatory environment while fully following Zambia's existing laws and policies is crucial.

Given these constraints, we can improve our research methods for this project and create better strategies to tackle the issues in Zambia's public procurement process.

1.6 Significant of the Research

- **Practical Significance**

  - ✓ Using machine learning and analytics in e-Procurement systems can make things more transparent, cut down on corruption, and make sure public money is used wisely.
  - ✓ Giving procurement officials and policymakers data-based information to help them make smarter choices.
  - ✓ Improving oversight to stop dishonest activities and encourage good governance.
  - ✓ Making procurement processes more efficient to lower costs, improve quality, and boost economic growth.

- **Theoretical Significance**

  - ✓ Improving our understanding of how Machine Learning (ML) and Analytics can be used in public procurement.
  - ✓ Developing frameworks and models to help integrate ML and Analytics into e-Procurement systems.
  - ✓ Offering insights into the difficulties and possibilities of setting up public e-Procurement systems in African countries.

- **Social Significance**

  - ✓ Increasing openness and responsibility to create trust between citizens and those in public offices.
  - ✓ Making sure public resources are used well to support long-term development and reduce poverty.
  - ✓ Giving policymakers, procurement officers, and citizens the resources and information they need to be part of decision-making.

- **Policy Significance**

  - ✓ Giving advice based on facts to improve policies and rules.
  - ✓ Working together with Zambia's plans for development and Vision 2030.
  - ✓ Sharing ideas that can help other African countries set up or improve their e-Procurement systems.

This research shows that the study can greatly help with better governance, economic growth, and social happiness in Zambia.

2. CHAPTER 2 Literature Review

2.1 Broad Literature Review

2.1.1 Overview of Public e-Procurement Systems

Public e-Procurement started as a method to make procurement processes digital, with the goal of reducing inefficiencies and increasing transparency. Initially, these systems were designed to automate tasks that were previously done by hand. However, they have since developed into full platforms for managing supplier relationships and ensuring that all rules are followed. According to Davila, Gupta, and Palmer (2003), *e-Procurement "helps companies make their buying processes smoother,"* and this idea has spread to public sectors around the world. It helps governments reduce corruption and improve how they are held accountable.

2.1.1.1 Benefits and Challenges in Public Procurement

The procurement of goods and services is one of the fundamental aspects of government operations, which seeks to achieve maximum economical use of resources. The deployment of a modern procurement system is associated with various advantages including increased transparency, better competition amongst suppliers, decreased operating costs, etc. However, there are also positive tendencies towards public procurement, which are: corruption, low level of technological advancement, unwillingness to adapt, etc, which may limit its efficiency. (OECD, 2016) (Williams-Elegbe, 2018)

Benefits in Public Procurement

- Heightened accountability
- Economical Value Added (Sufficiency)
- More suppliers competing for the opportunities
- Enhanced purchasing procedure

Challenges in Public Procurement

- Vulnerability to fraud and corruption
- Lack of appropriate technological infrastructure
- Reluctance of the parties concerned to adopt new systems
- Complicated regulatory environment (Overly stringent policies)

2.1.1.2 Role of Technology in Enhancing Transparency in Procurement

The impact of technology on procurement systems is impressive. It has increased the two fundamentals of good management practices; "**transparency"** and **"accountability"**. Through e-Procurement systems in place, specific procurement activities can be monitored in real time, with every transaction being documented for retrieval at a later stage thus lowing the risks associated with dishonesty. As noted by Vaidya, Sajeev and Callender (2006), *"e-Procurement helps in improving visibility and monitoring of procurement processes"* which enhances their fairness and equity.

Also, using advanced tech like *blockchain* and *data analysis* means that records can't be changed, which makes people trust the procurement system more. These tools help stop cheating and unfair tricks, making public deals transparent. (Schniederjans and Hales, 2016).

2.1.1.3 Regulatory Framework and Best Practices in e-Procurement

A strong regulatory framework is needed to promote transparency, accountability and fairness in e-Procurement. This is because regulations stipulate how public procurement is carried out and, in this case, they help to make sure that the processes are open and competitive so that the chances of any corruption or fraud are very low. Thai (2001) contends that *"In all countries, public procurement is subjected to various laws and policies primarily to mitigate the risk of abuse, promote equity, and ensure optimal utilization of taxpayer funds."* Agreements and requirements like the UNCITRAL Model Law on Procurement and the WTO Agreement on Government Procurement tend to be embraced in order to create uniformity of standards in procurement activities across regions 8and nations. In the case of Zambia, the Zambia Public Procurement Authority (ZPPA) regulates public procurement to ensure adherence to principles of transparency and accountability in the procurement process.

Therefore, effective e-Procurement practices involve simplified procedures, adherence to regulations, and the use of mechanisms like audit trails and automated reporting. These best practices enhance efficiency and also increase confidence among various stakeholders (OECD, 2016). Additionally, it is also important to inform the procurement officials on new regulations and new technologies since it contributes to the success of e-Procurement (Arrowsmith, 2010).

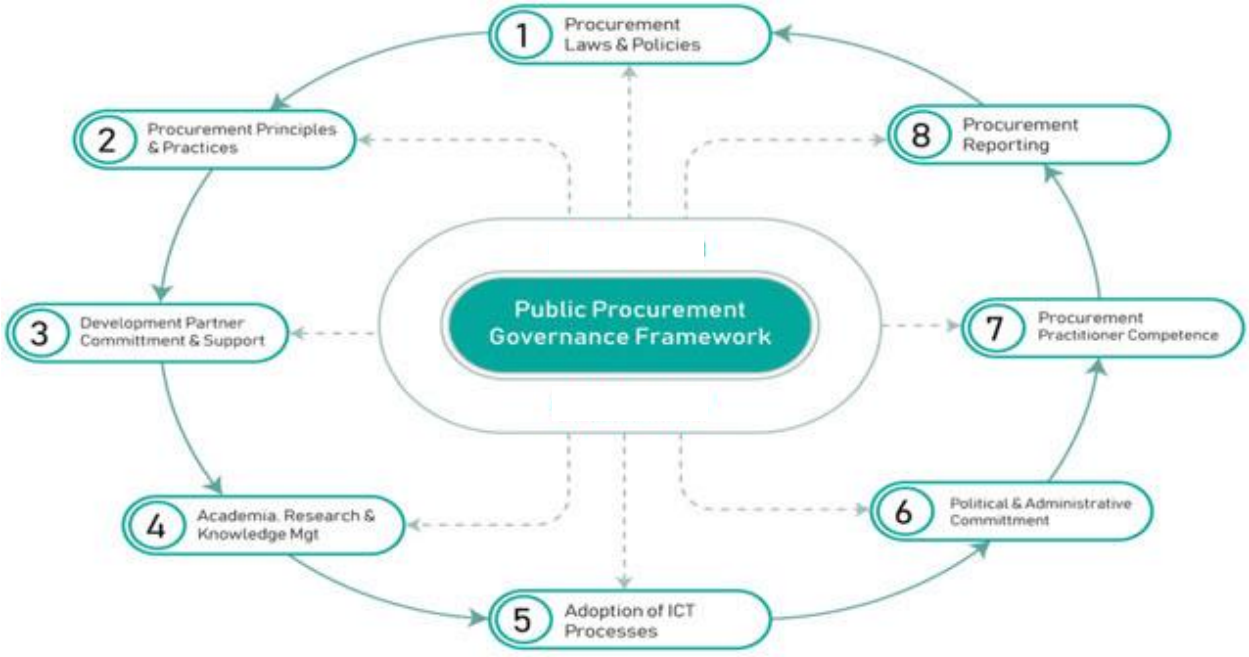


Figure 2.1.1.3 *Public Procurement Governance Framework (Benon C. Basheka, 2021)*

2.1.2 Machine Learning in Fraud Detection

2.1.2.1 Machine Learning Techniques for Fraud Detection

Fraud detection in various fields, including public procurement, has necessitated an increasing use of machine learning (ML) techniques. The purpose of these techniques is to filter large amounts of information in search of particular behaviors or outliers that suggest fraud. Classification on the basis of historical data can help assist supervised learning methods such as decision trees or support vector machines in identifying whether each transaction is normal or suspicious (Chandola et al., 2009). Further developments are made by Ahmed et al. (2016) who state that *"with the help of machine learning algorithms, it is possible to make fraud management systems more effective, as they learn to predict the possible cases of fraudulent activities based on the trends of this behavior in the past and over time become more accurate at this."*

One of the fraud detections use-cases discussed later in this research will demonstrate how *unsupervised learning* can be employed to identify potential cases. For instance, **K-means** can segment whatever transactions/newspapers into clusters to detect obvious outliers which could be fraud/fake news by our human intuition (Hodge and Austin, 2004). Similarly, since it can be used to enhance the prediction performance of other models (Zhao et al., 2016), combining fraud detection with many other models has also been shown to work.

Such techniques of machine learning provide organizations with tools that help prevent and control fraud, which increases the integrity of procurement processes.

2.1.2.2 Case Studies of ML Application in Fraud Prevention

Machine Learning has been implemented across various industry sectors for fraud prevention, and indeed helps in detecting fraud. For example, banks have turned to ML for purposes of recognizing credit card fraud. Additionally, companies such as **Paypal** make use of much more complex ML systems that filter through millions of transactions and detect any questionable or dishonest behavior with minimum delays (Levi et al., 2017). Within a given time frame, these models make predictions on the occurrence of fraud using historical data through supervised learning, and the further the time extends, the more accurate they become.

In the latter case, for instance, the insurance sector applies ML to detect fraudulent activities at the stage of claims submission. E.g., **Allianz**, a worldwide connected insurance solutions company has installed ML systems based on decision trees, neural networks, among others, to filter files of claim records and pay attention to those with high chances of fraud (Brockett et al., 2002). In this manner, the system processes the historical data entered in its engine and cannot be fooled by false claims since it is able to recognize that false claims never fit the pattern of the previous legitimate ones.

Another instance is the practice of public procurement where government agencies have also started implementing ML systems for fraud monitoring. In Brazil, for example, the Ministry of Transparency has put into operation a system based on ML aimed at detecting a typical behavior in procurement, such as collusive tendering as well as excessive pricing (Ferraz et al., 2021).

Through the adoption of this system, it has been possible to enhance the integrity of public procurement processes as well as mitigated the risk of corruption by preventing dishonest activities that may happen after awarding a tender but before entering into a contract.

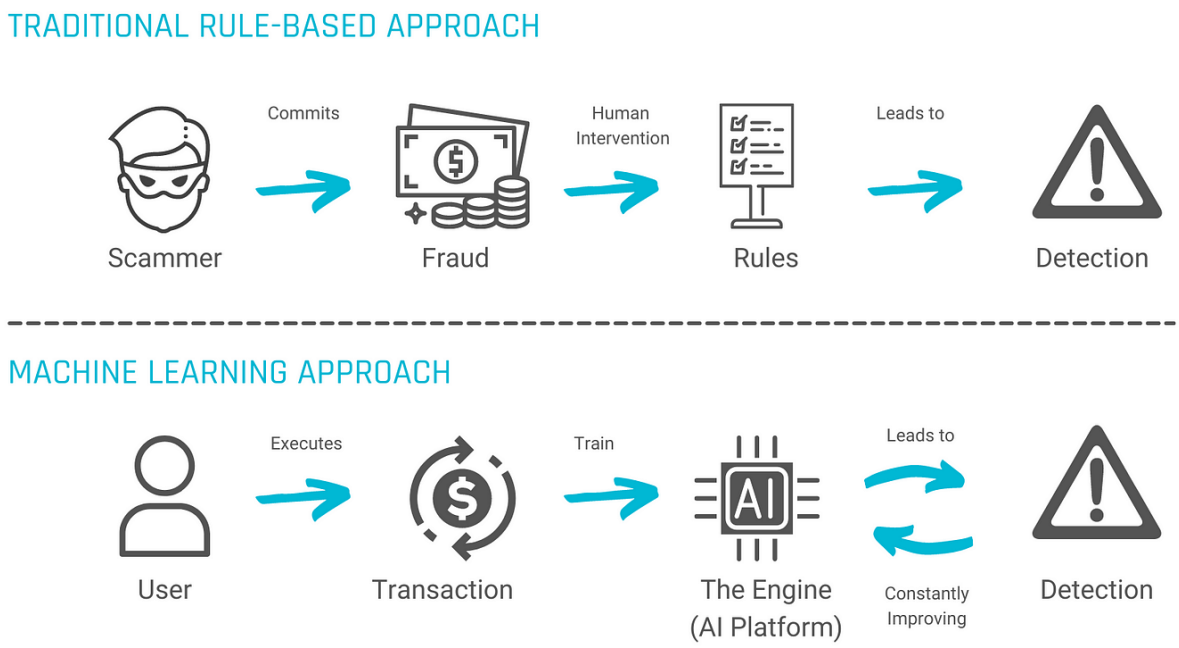


Figure 2.1.2.2 *Fraud Detection with Machine Learning — A Use Case (PI. EXCHANGE, 2021)*

2.1.2.3 Types of Fraud and Anomalies in Public Procurement

Fraud and anomalies in public procurement can substantially affect the integrity of the process, resulting in financial losses and damaged confidence in public institutions.

Common types of fraud include:

- Bid rigging (Collusion among bidders)
- Bribery (Influencing officials for favorable outcomes)
- Over-invoicing (Inflated pricing on contracts)
- Submission of false documents (Fake certifications, financial statements, reports)
- Conflict of interest (Officials having a personal stake in contracts)
- Fake suppliers (Non-existent companies winning contracts)
- Unjustified sole sourcing (Avoiding competitive bidding)
- Kickbacks (Suppliers offering officials incentives for contracts)
- Duplicate payments (Paying multiple times for the same service)
- Non-performance (Winning suppliers failing to deliver on contracts)
- Unnecessary procurement (Purchasing goods or services that are not needed)

2.1.2.4 Comparison of Supervised and Unsupervised Learning for Fraud Detection

Two of the major machine learning fraud detection approaches are ***supervise***d and ***unsupervised*** learning. Supervised learning makes use of labeled data so as to identify known fraud patterns and is thus quite accurate when historical fraud data becomes available. Decision trees and support vector machines are some of the common techniques resorted to. (Ngai et al., 2011).

Unsupervised learning on the other hand does not require labeled data. It is effective in identifying unknown fraud patterns through outlier detection or anomaly in data. Techniques of clustering and anomaly detection are widely used in this category. (Hodge and Austin, 2004).

2.1.3 Market Price Estimation Using Machine Learning

2.1.3.1 Predictive Models for Market Price Estimation

Various predictive models are applied in estimating market prices. The model analyzes historical data and selects patterns that dictate pricing. The standard machine learning techniques used in market price analysis include regression, decision trees, and neural networks. Regression models such as linear regression explain price predictions based on relationships between variables: demand, supply, and economic trend variables. On the contrary, decision trees rely on partitioning the data into smaller subsets; the models make predictions based on a series of conditions and factors. On the other side, decision trees divide data into small subsets on which prediction can be made based on certain conditions and factors (Breiman, 2001).

Neural networks have also become popular in recent times for market price estimation, considering the neural networks' ability to handle a dataset that is complex in size and with non-linear relationships. These models become even better the more they are fed with data; hence, they give better predictions over time as more data is processed (Goodfellow et al., 2016).

2.1.3.2 Impact of Price Estimation Accuracy in Procurement

Estimates of prices in procurement are very core to the accomplishment of three very important values: fair competition, proper budget management, and cost efficiency. In instances where procurement teams are able to make correct predictions of market prices, there is no case of overpaying for goods and services, nor the exploitation of suppliers. Consequently, this leads to better resource allocations and enables the organization to maintain budget limits as prescribed by Thai (2009). Secondly, accurate price estimation promotes transparency and accountability. This decreases the likelihood of fraud and other corrupt practices from occurring in public procurement. On the other hand, poor estimates will lead to financial waste and a loss of confidence on the part of stakeholders in the overall success of the procurement operations.

2.1.4 Bidder Analysis and Classification

2.1.4.1 Bidder Categorization in Procurement Systems

Bidder classification is a fundamental feature of procurement systems which in turn improves the processes of assessment and selection. Since bidders are not homogeneous, such approaches under procurement management allow the organizations to carry out procurement activities more efficiently and promote a clearer selection process.

Categorization can be done on various parameters referring to experience and past experience, financial stability, industrial sector, past performance, adherence to regulatory standards, and quality of goods or services provided. For instance, bidders may be categorized as "**preferred suppliers**," "**new entrants**," or "**high-risk bidders**" in a way that is useful for the procurement teams so as to understand the level of their reliability in performing their job under certain conditions of specific contracts.

In addition, industry sector categorization helps procurement professionals to associate suppliers to the right initiatives, hence allowing them to contact only those bidders who have the necessary expertise and capabilities to handle particular areas such as construction, ICT, stationery, and so on.

Advanced analytics and machine learning can classify bidders in a more sophisticated way by developing patterns from the historical data and predicting future performance. This will not only enhance decision-making but will also create a competitive environment in which qualified bidders are most likely to succeed, thus improving efficiency and integrity within the procurement process.

2.1.4.2 Role of ML in Bidder Evaluation and Risk Assessment

A great benefit of ML when allocating bidders is the ability to automate the consideration process, which vastly reduces human bias and enables transparency. ML uses predictive analytics to determine how likely bidders are to perform in the future based on their past behavior, identifying risks such as fraud or mismanagement, poor delivery service and cost overruns. ML models can also identify risk factors that are hard to see by a human, like sudden deterioration in financial health or unusual bidding behavior. *(Nguyen, T. H., & Schniederjans, D. G, 2022)*

Machine learning model uses historical data to determine the probability of procurement risks, thus providing a better result for procurement officials in their decision making. These models spot trends and anomalies, which may often be a leading indicator of supplier instability or fraud or failure to live up to its contracted obligations. In this environment, certain types of unsupervised learning may be used to automatically identify new patterns and outliers which could suggest collusive bidding or conflicts of interest.

Finally, overall integration of ML into bidder evaluation and risk assessment not only increases the precision and efficiency of the procurement buying cycle but also allows for allocation of public resources to reputable suppliers while minimizing risk to both finances and operations. *(Nguyen, T. H., & Schniederjans, D. G, 2022)*

2.1.5 Data Analytics in e-Procurement

Through process optimization, supplier selection, and transparency creation, data analytics support the improvement of e-procurement (Wang et al., 2022). Power is the paramount form of accounting, whereas the analytics, in fact, go ahead to break down supplier performance and better marketing and risk management planning (Zhao et al., 2023). Besides, big data and analytics empower corporations to base decisions on data for enhanced supplier collaboration and strategic sourcing (Nguyen & Tran, 2023). Nevertheless, data-mining techniques will find hidden insights and irregularities in procurement data that, in turn, would help improve processes and identify inefficiencies (Kouhizadeh et al., 2022). Furthermore, analytics lead the direct cost reduction and efficiency gains by indicating opportunities for cost savings while at the same time enabling the management of inventory at optimum levels (Bai et al., 2022).

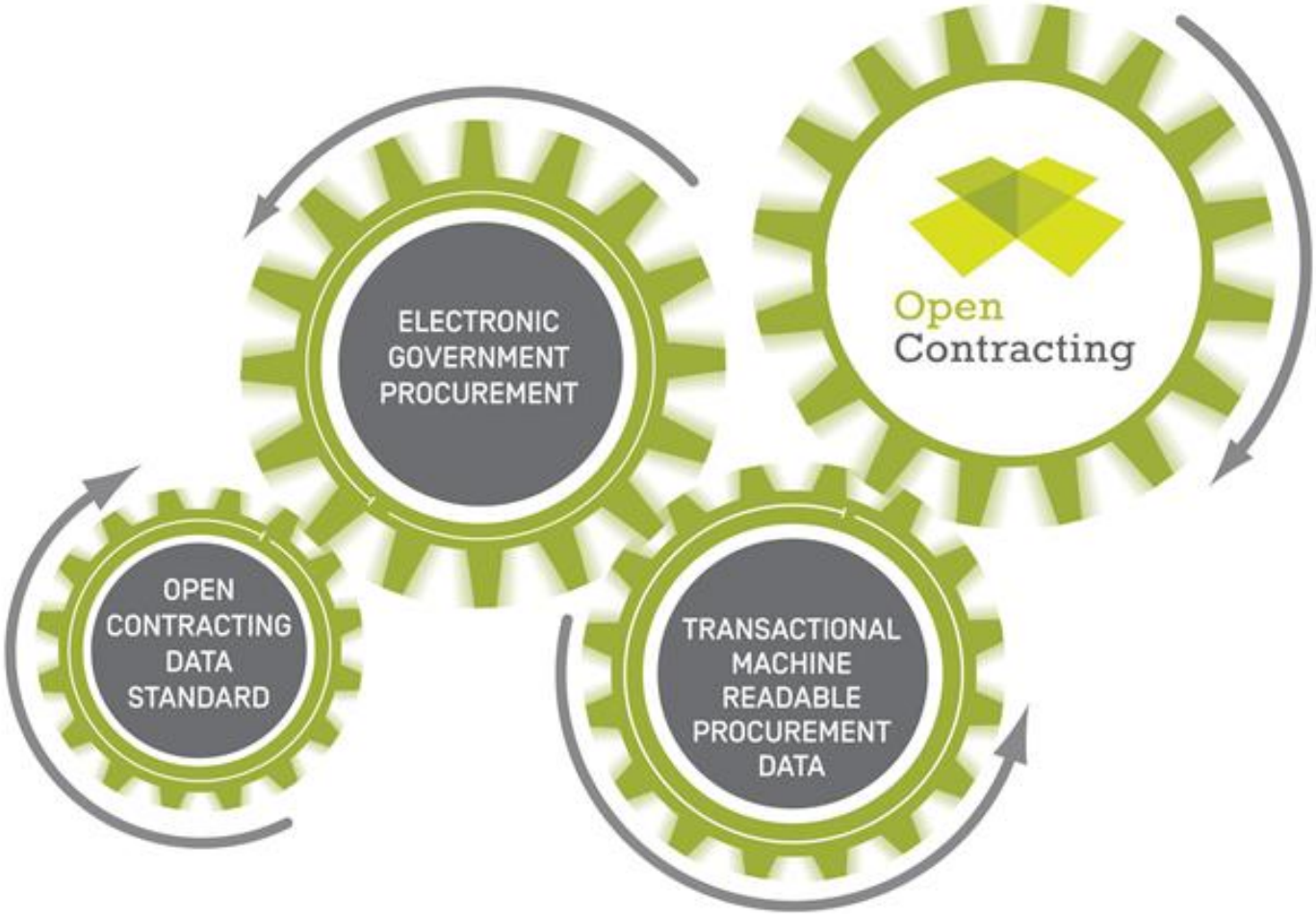


Figure 2.1.5 *Big data analytics in Government e-procurement – Openness (Chris Smith, 2017)*

2.1.6 Evaluation of Machine Learning Models for e-Procurement

Machine Learning models enhance the efficiency and security of e-procurement. Various studies have assessed the performance of different ML models regarding fraud detection and risk assessment. According to (Soares et al, 2019), "*neural networks perform better than decision trees for fraud detection,* while *hybrid models enhance accuracy* notes (Liu et al, 2022). Key evaluation metrics include **accuracy**, **precision**, and **F1-score** (Sarkar et al., 2020). Implementation of ML models reduce procurement risks by **30%** (Kersten et al., 2020). These findings underscore the potential of ML in optimizing e-procurement processes.

2.1.7 Impact of Machine Learning and Analytics on Procurement Processes

Machine Learning and analytics, undoubtedly, have transformed the procurement process by boosting accountability, fairness, and sharing of information that makes everyone more accountable. For instance, ML has perfected accountability in e-procurement by automating supplier reviews and eliminating bias, therefore making the process of selection fairer (Smith et al., 2023). Apart from automated, ML technology can help prevent frauds using tools like a fraud detection system, which is equipped to quickly recognize unusual behaviors, thus removing the need for human intervention and minimization of operational costs (Li & Zhang, 2023). Moreover, the use of ML and analytics is a clear example of transparency as they offer real-time procurement information, which gives a heads-up on the operation (Chen & Liu, 2023). Nevertheless, the methods of ethical nature towards data privacy and algorithmic bias may be largely considered the most important elements to tackle for the sake of responsible ML implementation in the public procurement sector (Wang et al., 2023).

2.1.8 Blockchain in Public Procurement

Blockchain technology allows users to record transactions using a distributed public ledger that makes it impossible for anyone to tamper with, change or delete transaction records. This makes such technology potentially very beneficial in enhancing trust and accountability in public procurement systems (Peters and Panayi, 2016). It provides a permanent record of all the transactions taking place on the chain, which discourages malicious behaviors like fraud and also improves the auditing of such procurement processes. With this, all the parties included in the procurement transaction, which include the suppliers, auditors and the respective authorities, can interact with a tangible and authentic evidence; hence it boosts the level of transparency.
(Draw, W. T. and T. R. 2016).

This, therefore, could be a game changing factor for public procurement in the Zambian context. Public procurement in Zambia has been dogged with many problems of transparency and accountability, usually arising from data manipulation as well as limitations of timely and accurate information access. (Transparency International Zambia, 2020).

With the implementation of blockchain technology, it would be impossible for anyone to tamper with procurement records in Zambia and thus eliminate malpractices in public sector contracts. There would also be an element of real time surveillance as well as provision of information to concerned parties thanks to the technology, enhancing monitoring by regulators and moving away from excessive red tape. This particular evolution in technology may also fall under the ongoing changes in Zambia's digital transformation as well as anti-corruption measures, thus enhancing the country's governance structures and restoring confidence in public procurement processes (Government of Zambia, 2021).

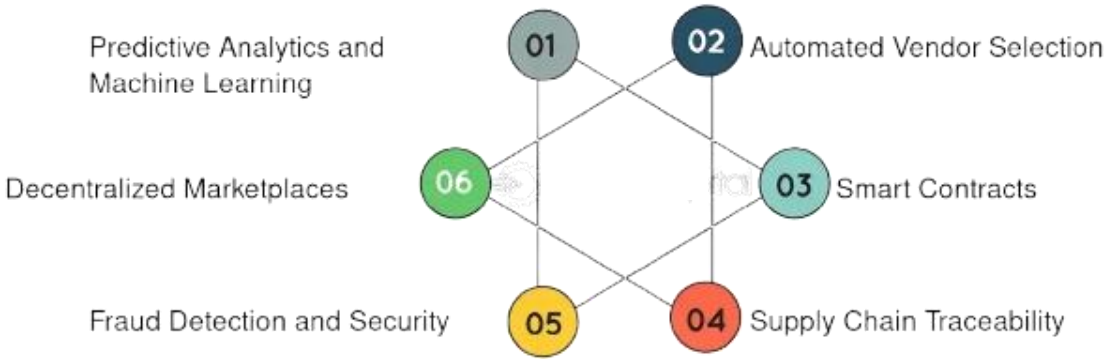


Figure 2.1.8a *Government E Procurement: Digital Transformation in Public Procurement: Lessons for Startups (FirstCapital, June 2024)*

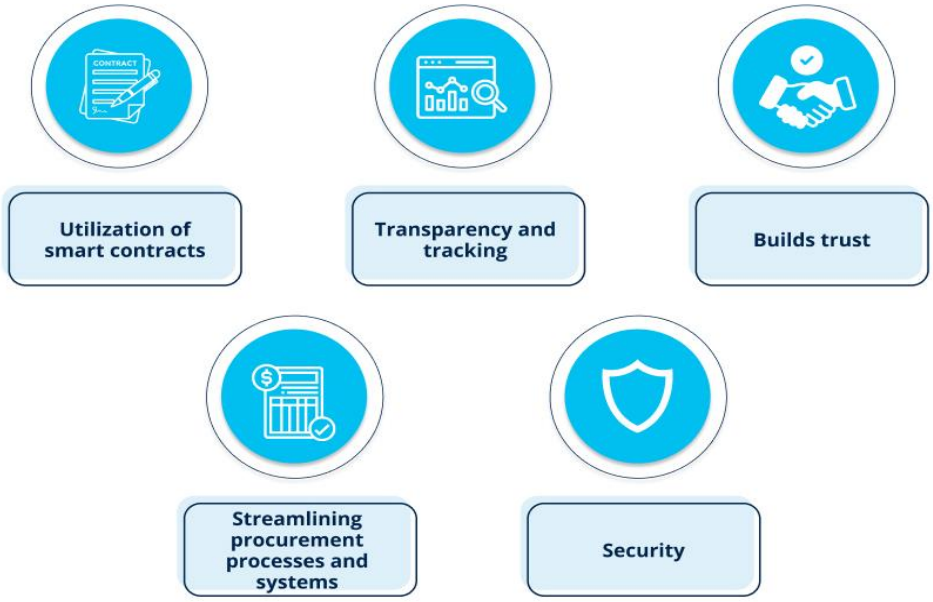


Figure 2.1.8b *Benefits of Blockchain in Public Procurement (Procurement Tactics - The Ultimate Guide of 2024)*

2.1.9 Natural Language Processing (NLP) for Document Analysis

NLP consists of a set of computational techniques implemented by AI systems, which focus on the interaction between human communication and computers. In procurement, NLP tools help to breakdown many layers of procurement documents like contracts, proposal submissions and tenders to check adherence to the requirements and detect erroneous or ill-intentioned clauses (Cheng et al., 2020). In this regard, we argue that the text mining concerns in procurement document processing are also relevant for the purposes of ensuring corruption free practices in procurement intermediation (Vaswani et al., 2017).

2.1.10 Ethics, Privacy, and Data Security in Machine Learning for Procurement

Notably, factors such as ethical challenges in machine learning, especially in terms of privacy and data concerns arises. In order to safeguard sensitive information relating to procumbent data from being accessed by unauthorized individuals and also sensitive procurement data from being leaked, it becomes necessary to put in place stringent security measures and adhere to the ethical principles that are highlighted (Floridi et al., 2018).

2.1.11 Regulatory Compliance and Legal Framework for e-Procurement in Zambia

The procurement process in Zambia is regulated by the Public Procurement Act (2018) which emphasizes value for money and integrity. The need to ensure e-procuring systems are in line with the legal framework can help to further enhance Zambian laws aimed at minimizing corruption and promoting competitive practices.

2.1.12 Challenges and Future Directions in ML and Analytics for Public Procurement

Zhao et al., (2023) Observe that some of the difficulties faced by ML and analytics during their actual application in public procurement include problems with system integration and data availability. The main issues are related to the complexity of deploying ML systems within the public procurement environment, which is typically marked by an attitude of resistance to change, alongside a limited technical infrastructure. Additionally, the quality and availability of procurement data are critical, as poor data can limit the effectiveness of ML models (Chen & Wang, 2023). Emerging trends in artificial intelligence (AI) and analytics, such as predictive analytics and advanced risk management tools, continue to shape the procurement landscape. Finally, blockchain technology offers the potential to enhance ML by providing secure, transparent, and immutable data for procurement systems (Li et al., 2023).

Note: the openness and accuracy of data can be regarded as a great impediment to public procurement processes. In most cases, the available data is either incomplete or cut up into pieces that cannot be standardized, which hampers the ability of ML models to make accurate predications or useful recommendations based on the collected data. Most of the times, such public procurement systems are built on older systems making data integration very challenging. More so, the lack of real time update of data also affects the dynamism of the ML models as they cannot take in new data transforming AI-led procurement systems to be less efficient (Chen & Wang, 2023). The application of better data governance policies, improving data-sharing frameworks and infrastructural development will be paramount in solving these issues.

2.2 Critical review of related works

This next section reviews various similar systems developed and implemented in other regions, with the intention of learning from their design, development, and implementation processes. Additionally, the aim shall be to identify those attributes that are needed to ensure a transparent electronic public procurement system using machine learning and analytics within Zambia's public procurement sector and some of the features that can be incorporated in our system

2.2.1 E-Procurement System by the United Nations

The United Nations Global Marketplace-UNGM is the common procurement portal for the United Nations system and provides a shared platform for the procurement of goods and services through the adoption of a more transparent and expeditious process. Founded to enhance the participation of vendors in United Nations agencies, UNGM has come to acquire an important role for enhancing fair competition, accountability, transparency and integrity in public procurement. (UNGM, 2024)

https://www.ungm.org/

Figure 2.2.1 *United Nations Global Market e-Procurement system (UNGM, 2024)*

2.2.1.1 Key features of the system

- **Public Access to Procurement Opportunities**

UNGM is a single focal point where procurement notices, tenders, and contract opportunities are published on the above portal, therefore making the information accessible to suppliers worldwide.

This publication by UNGM enhances transparency in procurement, allowing suppliers to look at the totality of the opportunities available besides opening an environment for open competition.

- **Registration and Categorization of Suppliers**

Suppliers have to register themselves with UNGM, submitting the required documents about their qualifications and expertise. The time of their registration also involves classification into specific categories, such as goods and services delivered, which enables UN agencies to identify potential suppliers while allowing all the registered entities equal opportunities.

- **E-Bidding Process**

The platform provides the facility for electronic bid submission, therefore avoiding many manipulations and frauds related to the physical submission of bids. The electronic bidding system enhances accountability through the availability of a clear, auditable trail of all submissions and communications.

- **Monitoring Performance and Evaluation**

UNGM uses data analytics in pursuit of its mandate to monitor supplier performance and procurement trends to glean important insights into the reliability of suppliers and effectiveness of procurement practices. This constant process of review ensures that the suppliers continue to adhere to UN standards for ethics and contract requirements, thereby creating accountability throughout the procurement process.

2.2.1.2 Advantages of UNGM

- **Centralized access to information**

All procurement opportunities are available on one platform that enable suppliers to access relevant information with ease. This means that, with this centralized approach, the level of transparency regarding tender and contract visibility is increased; whatever is available can be seen by suppliers for increased competition, hence decreasing information imbalance among bidders. (Dutta & Zang, 2023).

- **Real-time Reporting and Monitoring**

UNGM system has a real-time reporting feature of all procurement activities. This therefore facilitate the need for either UN agencies or suppliers to appreciate the ongoing procurement, thereby helping detect some anomalies or departures from prescribed procedures that could be rectified on time. (Nguyen & Huynh, 2021).

- **Anomaly Detection**

The system is intended for an advanced database structure in order to achieve automated anomaly detection through data analytics. While analyzing the procurement data, UNGM highlights the pattern of fraudulent behavior or corrupt practices, such as unusual bidding patterns and supplier collusion. Thus, this proactive approach strengthens the integrity of the process.

- **Enhanced Categorization of Suppliers**

The database structure supports effective categorization of suppliers based on performance and qualification. It ensures transparency in the evaluation of suppliers so that UN agencies make informed decisions while selecting vendors. It is one of the best ways through which fair competition can be introduced to reduce the risks of corruption, as pointed out by (Bai & Sarkis, 2022).

- **Integration of Anti-corruption Measures**

Anti-corruption measures are integrated at the forefront in the UNGM application. The system ensures that suppliers and UN agencies behave in a proper and ethical manner. It helps trap fraud and corruption by setting proper guidelines and expectations over conduct in the procurement process.

2.2.1.3 Disadvantages of UNGM

- **System Complexity**

The system can be portrayed as complex. This is where potential suppliers would shy away and, particularly for SMEs, with small resources and less than advanced technical capabilities. This may result in confusing procurement with little transparency because the pool of bidders remains restricted.

- **Accessibility of the Platform in Developing Countries**

In most developing countries, due to minimal internet access and technological infrastructure, suppliers cannot have access to the UNGM online platform. This inaccessibility could mean lesser competition and transparency in most procurement opportunities because not all qualified suppliers are in a position to compete in such opportunities.

- **Potential for Data Mismanagement**

Data mismanagement or inaccuracies are other risks associated with relying on data analytics. Data inaccuracy or mismanagement could create flawed procurement decisions that reduce transparency and accountability.

- **Digital Fraud Risk**

As much as it has anti-fraud measures put in place, reliance on digital systems creates avenues for new forms of fraud, such as cyber-attacks and data breaches. These could compromise the integrity of the procurement process and even undermine efforts toward ensuring transparency

2.2.1.4 Incorporated features

- Bidder Categorization
- Automated Fraud Detection
- Transparent Evaluation Criteria
- User-Friendly Interface

2.2.2 Federal Business Opportunities (USA)

Federal Business Opportunities (FBO), now transitioned into System for Award Management (SAM), is a very important Web-based system utilized by the U.S. federal government for posting procurement opportunities. Its objective is to enhance transparency and accountability in public procurement and encourage fair competition among vendors. In this case, businesses, particularly the small and disadvantaged enterprises, should be afforded access to government contracting opportunities toward an inclusive marketplace.

https://sam.gov/content/home

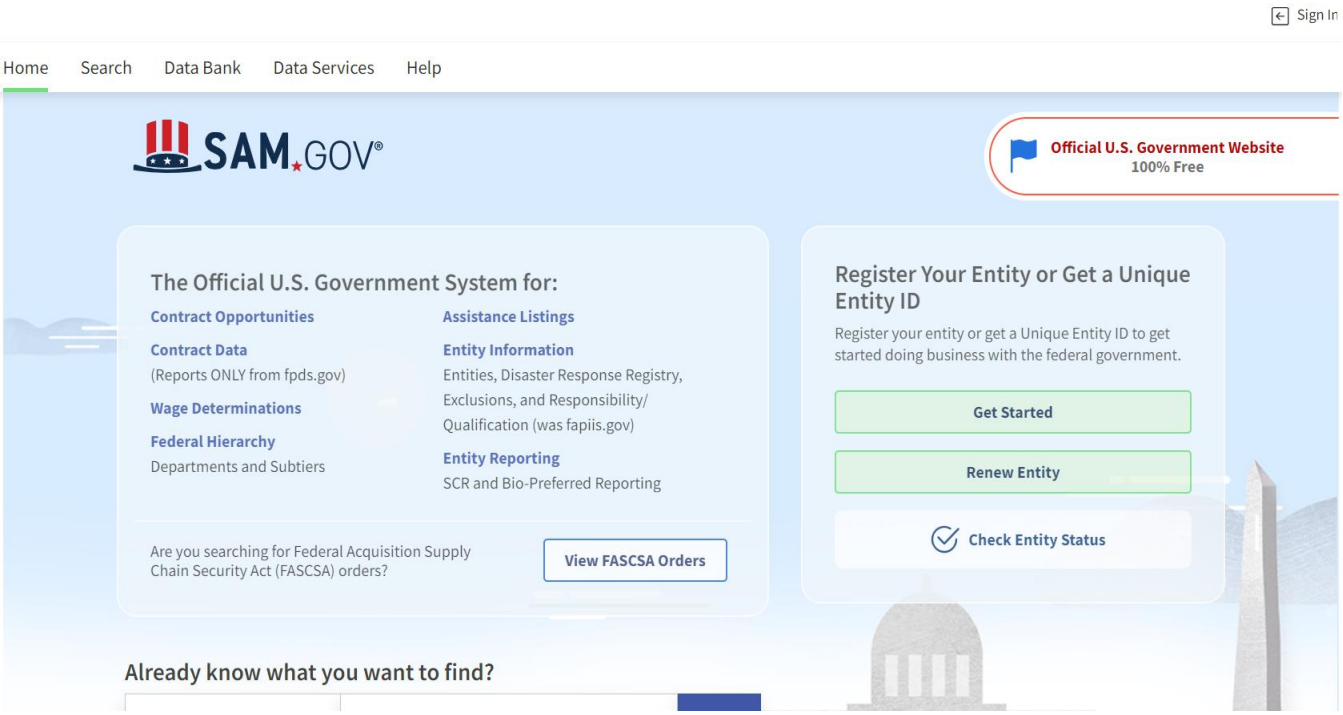


Figure 2.2.2 *System for Award Management (SAM)- FBO system (SAM-FBO, 2024)*

2.2.2.1 Key features of FBO – SAM

- **Automated Alerts for Anomalies**

The system is embedded with automated alerts that flag off an anomaly in the pattern of bidding or any discrepancies in submissions by suppliers to facilitate the possibility of early detection of fraud.

- **Reporting Tools**

The system has a detailed reporting tool that helps in the creation of insights by procurement officials relating to procurement activities, performance of suppliers, and spending trends of departments for enhanced accountability.

- **Fraud Reporting Mechanisms**

The system includes channels for reporting suspected fraud, corruption, or misconduct, facilitating whistleblowing and accountability among stakeholders.

- **Analytics for Procurement Insights**

It delivers analytics on data for procurement trends, vendor performance, and market conditions to drive informed decisions by procurement.

- **Machine Learning Capabilities**

Backend built with integrated Machine Learning techniques to enhance fraud detection and risk assessment through predictive analytics.

- **Integration with other Federal Systems**

The system integrates with other federal procurement systems and databases, hence providing a seamless end-to-end ecosystem for federal contracting.

- **Regular Audits and Oversight**

The organization and its system are thus subject to regular audits and oversight by federal agencies for compliance with regulations, as well as for reviewing fraud detection measures.

- **Secure Data Management**

Ensures that data management for purchasing activities and supplier information is maintained in a secure environment.

2.2.2.2 Advantages of FBO system

- **Increased Transparency**

By allowing a single source for all federal procurement opportunities, the FBO system offers transparency whereby suppliers stand in an equal position as far as information on the availability of contracts and the process of bidding for the contracts go.

- **Principle of Fair Competition**

Open procurement opportunities present fair competition for contractors and increase competition among vendors, including small and disadvantaged businesses.

- **Simplified Procurement Process**

The FBO system simplifies buying by placing relevant information in one location where suppliers can, with less time and effort, locate and bid on contracts.

2.2.2.3 Disadvantages of FBO system

- **Limited Real-Time Monitoring**

This means that in real-time, the system may not allow for the monitoring or quick detection of anomalies or fraudulent activities. This may make fraudulent behavior more prolonged before it is detected.

- **Data Quality Issues**

Clearly, it is obvious that the efficiency of the system is heavily dependent on the quality of the data coming from the suppliers. Partial or incorrect submissions will contribute to less transparency and may further lead to incorrect insights.

- **Risk of Too Much Dependence on Automated Solutions**

Fraud detection has made significant progress due to automation. However, systems are dangerous because they may encourage complacency among procurement officials who may not feel the need for manual checks and balances.

- **Risk of Cybersecurity**

Like most web-enabled systems, FBO faces several threats of a cyber nature, which could endanger procurement sensitive details and the very essence of the public trust that this system seeks to create.

- **Inadequate Reach for Certain Classes of Suppliers**

Even though the system is made open to all, a number of low tech or small sized suppliers may find it difficult to use the system leading to a poor mix of suppliers.

- **Laxity in whistleblower provisions**

Even though there are mechanisms to lodge complaints concerning fraud and corruption, the weak protections extended to whistle blowers may dissuade people from reporting such misconduct and thus there is no accountability.

- **Limited Use of Machine Learning**

Lack of high-level machine learning integration may tighten the systems forecast and trend analysis of procurement activities which may facilitate risk detection.

2.2.2.4 Incorporated features

- Access Portal
- Real-Time Reporting and Dashboards
- Automated Anomaly Detection:
- Whistleblower Reporting Mechanism
- Supplier Performance Evaluation System

2.2.3 Tenders.gov.au (Australia)

This is a centrally managed procurement system for the Australian Government that serves to enhance its accountability and further the transparency of its public procurement processes. It is a single entry-point application where buyers publish government tenders and contracts, thus giving businesses the opportunity to access one-stop, numerous procurement opportunities.

https://www.tenders.gov.au/

Figure 2.2.3 *Aus tender - Tenders.gov.au (Australia, 2024)*

2.2.3.1 Key features of Aus Tender

- **Open Access to Information**

The platform provides free and open access to all procurement opportunities, ensuring that all suppliers, regardless of size or location, can participate in the bidding process. This openness reduces the potential for corrupt practices by making tender information readily available to all stakeholders.

- **Standardized Processes**

Tenders.gov.au promotes standardized procurement processes across various government agencies, reducing ambiguity and minimizing the chances for manipulation or fraud. By adhering to consistent guidelines, the system fosters fairness and accountability in how tenders are managed.

- **Transparency in Reporting**

The system requires agencies to publish detailed information about contract awards, including the rationale behind selection decisions. This transparency allows for public scrutiny and accountability, as stakeholders can review how procurement decisions are made.

- **Auditing and Compliance**

Regular audits and compliance checks are conducted to ensure that procurement activities align with established policies and regulations. This oversight helps deter corrupt practices by ensuring that all activities are subject to review.

- **Whistleblower Protections**

Tenders.gov.au encourages the reporting of any suspected corruption or misconduct through secure channels. By providing protections for whistleblowers, the system fosters an environment where unethical behavior can be reported without fear of retribution.

2.2.3.2 Advantages of Aus Tender

- **More transparency**

The fact that Tenders.gov.au makes all the procurement opportunities and contract information publicly available significantly adds to making the public procurement processes more transparent. This openness implies greater scrutiny of processes and outcomes in tender procedures by stakeholders, hence reducing areas where corrupt deals can take place.

- **Standardization of the procurement process**

Normally, the system leads to a standardized process across various government agencies, ensuring that consistency is maintained while managing tenders. Secondly, such a standard ensures that ambiguity is reduced, and manipulation risk minimized for accountability of procurement decisions.

- **Robust reporting mechanisms**

Tenders.gov.au places obligations on agencies to publish comprehensive reports about the award and selection criteria of contracts. These puts added accountability because the stakeholders will have more understanding of the grounds for procurement decisions and are in a better place to determine anomalies that could suggest fraud or corruption.

- **Audits and compliance checks**

These would serve as a damper to fraudulent behavior through regular audits and compliance reviews. This ensures that procurement activities are periodically reviewed to ensure integrity and accountability in the procurement process.

- **Whistleblower Protection and Reporting Channels**

Tenders.gov.au has also provided a way through which fraud or corruption by whistleblowers can be reported through secure channels. This provides better protection for the whistleblowers, making them more emboldened to report any conduct that is unethical and, hence, increasing corrupt activities that may occur within this procurement system.

2.2.3.3 Disadvantages of Aus Tender

- **Limited Real-Time Monitoring for Fraud Detection**

Transparency is emphasized in the system however the system is devoid of highly developed real-time fraud detection techniques like machine learning tools which can flag certain bidding or procurement behaviors as potentially problematic. It is often the case that old-fashioned supervision approaches do not catch the nuances of collusion or other forms of corrupt practices (Brown, 2022).

- **Heavily Dependent on Reporting**

The system depends upon the respective concerned agencies and stakeholders for reporting any disparities or fraudulent instances which might be prevalent. Since there are no active steps or active provisions of fraud detection technology, the strategy towards eradicating corruption is mainly based on the trust in external principles, which can be problematic especially when trying to spot the existence of fraud (Smith & Allen, 2021).

- **Subcontracting Information is Not Fully transparent**

Even though Tenders.gov.au ensures transparency on the contracts that have been awarded, it does not always provide clarity regarding the allowances for subcontracting under such contracts. This may lead to situations where there is corruption, but it cannot be traced for example where corrupt activities occur in several layers of subcontractors to the one earlier tendered for (Harris, 2023).

- **Limited Use of Megabytes in Detecting Fraudulent Activities**

The current system may not be configured to link with other jurisdictions' empirical databases, which makes it difficult to carry out a crosscheck on such details as a contractor history, their finances, and their past completion records. This lack of thorough data mining capabilities lessens the potential of the system in detecting fraudsters or serial corrupt practices (Martin, 2022).

- **Concerns About Protecting the Identities of Informants**

Even if the system has provisions for reporting by whistle blowers, certain individuals may still feel that the level of protection and anonymity is not up to their expectations, and this may cause them to refrain from reporting corruption. Weak protections in place may result in more people failing to report activities

2.2.4 E-Procurement System (Government of India) – GeM - CPPP

Developed by the Government of India with the aim to make public procurement process more transparent, accountable and efficient. Creating a level playing field for suppliers and bidders, it provides a regulated and standard diagram for the processing of government tenders. This system helps to combat corruption and such tendencies as nepotism as noted by (Mishra & Kumar, 2021). Though still not redeployed with powerful AI technologies, the system contains audit logs and other observation controls that help spot anomalies and fraudulent activities. The system also classifies vendors registered and monitors their behavior to ensure that they perform within the projected levels (Reddy & Singh, 2022).

https://eprocure.gov.in/cppp/

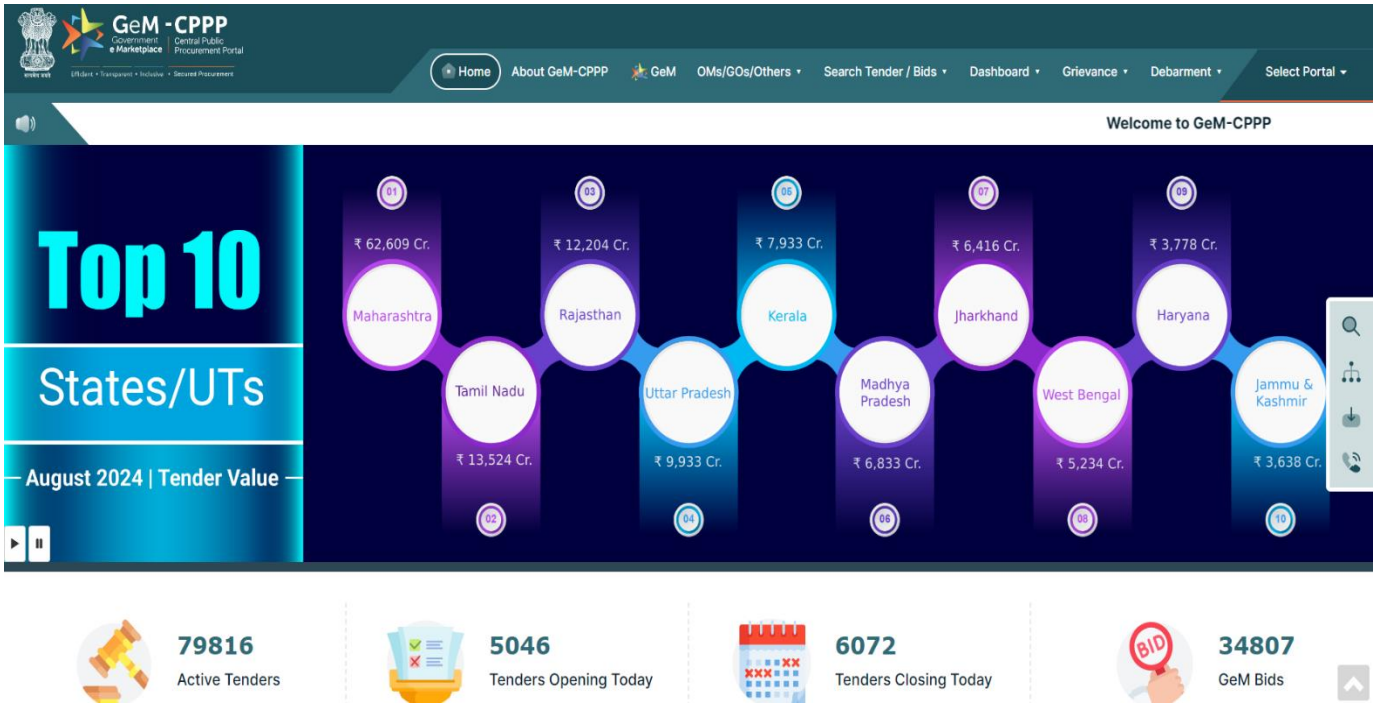


Figure 2.2.4 *GeM – CPPP – India Procurement System (India, 2024)*

2.2.4.1 Key features of GeM - CPPP

- **Transparency**

The integrity of the system guarantees the access of all procurement opportunities and contracts to the general public. Transparency is further enhanced, and sidelining discouraged through the posting of tender advert, bid information and contracts given out (Mishra & Kumar, 2021).

- **Corruption Prevention**

The system has automated functions and include storing of documents eliminate the chances of people changing or avoiding processes (Reddy & Singh, 2022).

- **Fraud Detection**

The system provides audit functionalities to monitor every transaction and activity done by users within the system. The system does not possess a smart detection mechanism based on machine learning to prevent fraud occurrence; however, such files are useful to look for suspicious activity for obvious reasons (Sinha & Gupta, 2020).

- **Bidders Categorization**

The online application allows bidders to be satisfactorily categorized and profiled. It records the bidding history, present financial standing, and regulatory compliance of a potential bidder, thus helping in promoting consistent and dependable bidders and deterring unfit or dubious ones (Patel & Sharma, 2019).

- **Efficiency and Accountability**

The system institutionalizes procedures making them more efficient by converting manual processes such as evaluation of bids, awarding contracts, and processing payments into automated systems. This lessens the chances of errors or manipulations and makes the officials answerable due to visibility and control (Mishra & Kumar, 2021).

2.2.4.2 Advantages of GeM – CPPP

- **Improve transparency**

The system promotes openness by publishing tenders. Submitting an auction and all contracts received online This is to ensure that all bidders will have equal access to information. This greatly reduces the chances of bias and undisclosed behavior (Mishra & Kumar, 2021).

- **Reducing corruption**

By automating the procurement process and reducing human intervention. The system reduces the chances of bribery or manipulation. Digital records and automated workflows prevent staff from skipping scheduled steps. This reduces the risk of corruption (Reddy & Singh, 2022).

- **Improved fraud detection**

Even if the machine learning tools are not installed. But the system still has strong audit trails. Every transaction within the platform is recorded. To be able to detect discrepancies and abnormalities. This can be detected to reveal potential fraud (Sinha & Gupta, 2020).

- **Effectiveness Of Bidders Classification**

The performance, finances and legal compliance history of the bidders is assessed to keep track of them and place them in the right category. This is beneficial to procurement personnel as it allows them to pinpoint trustworthy vendors and exclude those with dismal performance histories or inexplicable behaviors (Patel & Sharma, 2019).

- **Reporting discrepancies and responsibility**

The system promotes accountability through real-time tracking and reporting mechanisms. The platform enables the identification of anomalies or non-compliance. by ensuring that all purchasing activities are traceable Promote accountability between agencies and bidders (Mishra & Kumar, 2021

2.2.4.3 Disadvantages of GeM – CPPP

- **Limited capability in fraud detection**

Although it has audit trails which track activities, the system still lacks advanced machine learning or AI-driven mechanisms that might depict fraud or anomalies in a proactive sense. This will lead to delays in locating fraudulent activities that might only be captured through audits (Sinha & Gupta, 2020).

- **Over-dependence on manual reporting**

However, even with automation in anomaly reporting, some sections need manual intervention or at least oversight from an external perspective; this may defeat the very possibility of real-time detection and allows corruption or irregularities to go unnoticed until it is too late. (Mishra & Kumar, 2021)

- **Inconsistent Quality of Data**

Besides, completeness and accuracy of the data being fed into the system depend upon the officials who are responsible for the procurement processes. Poor quality or incomplete records can distort bidders' categorization and do the opposite by reducing transparency, thus opening up avenues for corruption, as noted by Patel & Sharma (2019).

- **Bidders' Categorization Issues**

Although helpful, the feature of the system in categorization of bidders may not be fully comprehensive; no real-time update on the performance and financial health of the bidders was practiced, and hence poor risk assessment results in continued participation of unreliable suppliers.

- **Very limited accessibility for small suppliers**

It would thus be hard to access and navigate by the smallest of suppliers or those that were less technologically savvy, thus reducing competition and giving an unfair advantage to larger, more established suppliers. This may lead, in turn, to unintentional reduced transparency, as some bidders are excluded in the process.

2.2.5 Chile Compra (Government of Chile)

ChileCompra is the national e-procurement platform of Chile with an aim to ensure transparency and accountability in public procurement processes. It classifies suppliers according to their performances and compliance levels, allowing fair competition. The system employs fraud detection strategies and uses advanced concepts such as predictive analytics to analyze the behaviors of tenders and suppliers to eliminate corrupt practices. Again, while making data available for the public for procurements, evaluation mechanisms are put into place to ensure that ChileCompra manages the risks of corruption in public procurements very effectively.

https://www.chilecompra.cl/

Figure 2.2.5a *Chile public e-Procurement system (ChileCompra, 2024)*

2.2.5.1 Key features of ChileCompra

- **Categorization of Bidders/Suppliers**

The system assesses a range of factors of suppliers like their performance, compliance, and contract records, so as to classify the suppliers. This way, only the reliable and competent suppliers partake in the bidding process fostering healthy competition.

- **Fraud Detection Mechanisms**

ChileCompra has automated tests and data cross-checks that detect suspicious activities in the tender processes, such as abnormal pricing or a conflict of interest.

- **Transparency**

The System ensures that all procurement processes and relevant information relating to tenders, contracts, scoring of bidders and so on are made accessible over the internet. This enables members of the public and external agencies tasked with monitoring the procurement process to do so, and they are able to ensure that fairness is upheld in the entire procurement process.

- **Responsibility and Audit Trails**

The system records all transactions and decisions with detailed audit trails where officials and bidders will be held accountable for actions taken. This ensures traceability and addresses potential malpractices

- **Anti-Fraud Strategies Backed by Artificial Intelligence**

The platform evaluates the procurement process using AI capabilities, which also include predictive analysis as well as machine learning based models to finding out particular activities and or suppliers that would be deemed suspicious. These techniques evolve and adjust accordingly with changing fraud trends thus enhancing the ability of the system to detect any corrupt activities at an early stage.

2.2.5.2 Advantages of ChileCompra

- **Supplier Quality**

By classifying bidders and suppliers according to their compliance levels, ChileCompra guarantees that only providers of proven quality participate in procurement processes, significantly increasing the quality of goods and services.

- **Fraud Detection**

Through automated checks & procurement data cross-reference helps in early detection of fraud which further able to reduce the chances of financial loss and corrupt practices detection before contracts being awarded.

- **Transparency Enhancement**

The availability of procurement data to the general public helps build confidence in the system and ensures that citizens and any oversight body can scrutinize each phase of procurement. It helps in keeping away from getting corrupt and also gives clear view that all purchasing is done in a right way. (ChileCompra, 2024)

- **Accountability through Auditability**

This ensures the audit trails in the system are at a sufficient level of detail to allow for tracing of decisions and transactions back to entities responsible. Accountability among officials and suppliers is therefore encouraged because any misconduct would be identified and punished.

2.2.5.3 Disadvantages of ChileCompra

- **Limited Data Accuracy**

Supplier categorization and fraud detection heavily rely on the quality and accuracy of data. Poor or incomplete data lead to poor assessments and undetected fraud cases of suppliers.

- **Implementation Complexity**

While the AI-driven fraud detection is powerful, the nature of implementation of those models may be pretty complex and costly. Smaller procurement bodies are likely to have limited resources and expertise and, therefore, may not take optimum advantage of such technologies.

- **Over-reliance on Automation**

This automation, or use of AI, in detecting fraud in the system at times overlooks sophisticated or even nuanced corrupt practices that require human intervention to detect. (ChileCompra, 2024)

- **Supplier Access Barriers**

Smaller suppliers or those that do not have certain technology capabilities may have difficulty using the platform and inadvertently be excluded.

Figure 2.2.5b *Chile public e-Procurement in-house system (ChileCompra, 2024)*

2.2.6 GeBiz (Singapore)

It is a public e-procurement platform whose aim is to level the playing field and promote transparency as well as accountability in government tenders in Singapore. It classifies suppliers and bidders which helps in transparent competition opportunity and quick procurement. GeBIZ helps to increase the transparency by publishing details and information of tenders and awardee and monitor the procurement activities for fraud detection and anomalies.

https://www.gebiz.gov.sg/

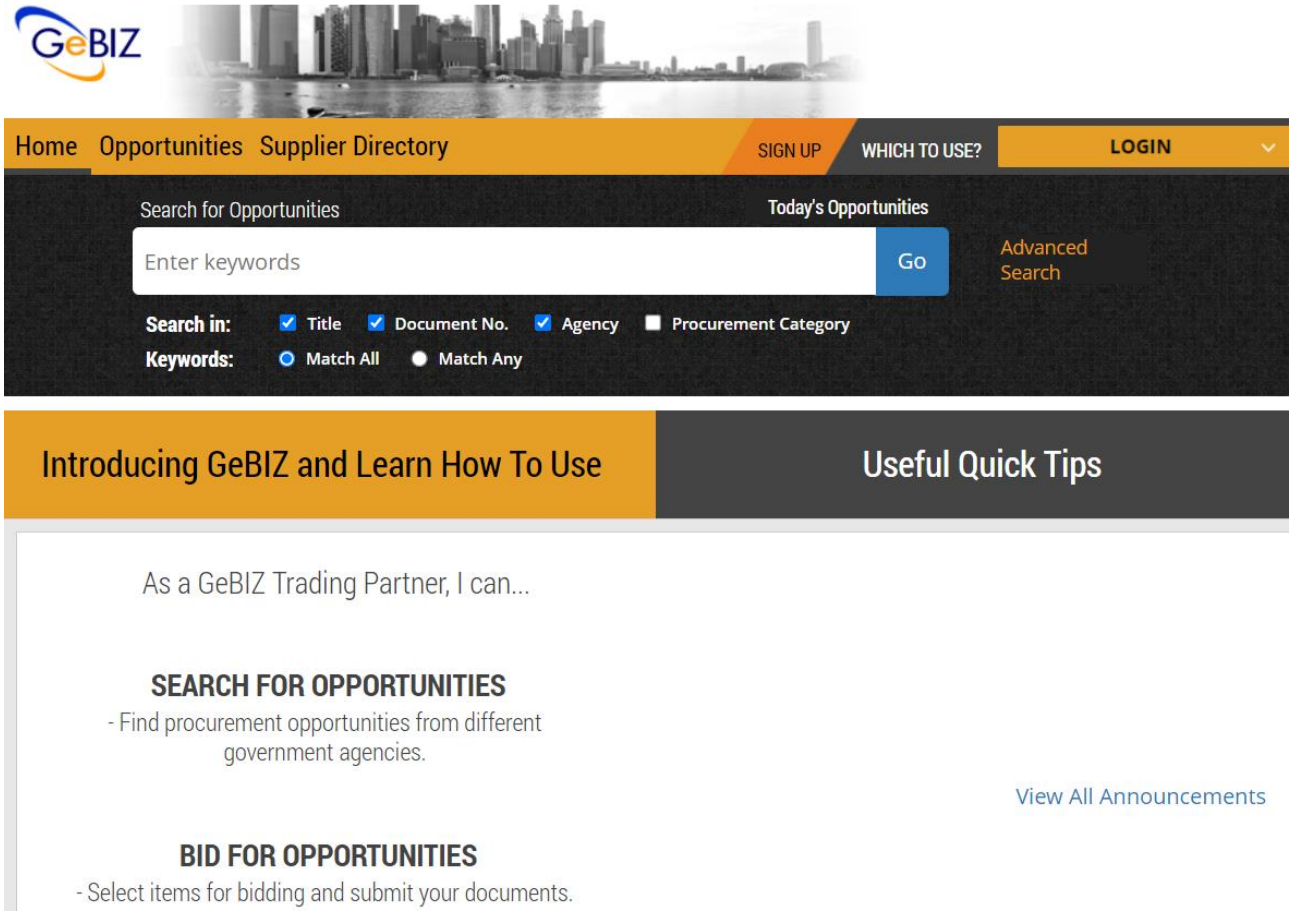


Figure 2.2.6 *GeBiZ e-Procurement (Singapore, 2024)*

2.2.6.1 Key features of GeBiz (Singapore)

- **Categorization of Suppliers**

GeBIZ provides for the systematic categorization of bidders and suppliers, hence helping identify and select suppliers efficiently with criteria such as past performance and capability.

- **Openness in Procurement**

The system publishes tender notices, the award of tenders, and procurement data to bring about open access to information and accountability among its stakeholders.

- **Mechanisms for monitoring and reporting**

GeBIZ has tracking tools in place to monitor procurement activities that enable detection of anomalies and suspicious behavior which could lead to fraud detection.

2.2.6.2 Advantages of GeBiz (Singapore)

- **Increased Transparency**

Publishing tender and award information in GeBIZ helps instill some level of trust and reduces corruption in public procurement.

- **Fair Competition**

Suppliers are classified systematically, and all bidders are accorded equal procurement opportunities; hence, this is a fair game.

- **Accountability**

GeBIZ provides audit trails of procurement decisions and actions through tracking and reporting mechanisms.

2.2.6.3 Disadvantages of GeBiz (Singapore)

- **Limited AI Integration**

Advanced AI and machine learning capabilities for automatic fraud tracking in real time are yet to be integrated into GeBIZ to further improve effectiveness in fighting corruption.

- **Data Quality**

Supplier categorization and fraud detection require good data input, which may be inconsistent in quality.

- **Risk of Complexity**

It could also be that the site, for smaller-scale suppliers or less technology-enabled beneficiaries, is a bit more complicated to maintain than initially anticipated. This could reduce their likelihood of being able to apply for public procurement opportunities.

2.3 Conceptual framework/Theoretical framework

The conceptual framework of this project through the use of machine learning and analytics system shall address the critical aspects of transparency, accountability, and fraud detection in Zambia's public procurement. The proposed system will have ML and data analytics detect anomalies with added benefits in improving decision-making processes. This framework shows the relationships that are supposed to be between public e-procurement, machine learning, and transparency in procurement.

2.3.1 Key Concepts

The main ideas motivating this framework consist of the application of public e-procurement, machine learning, and data analytics. E-procurement systems in the public sector have been historically adopted to streamline the acquisition of goods and services by various levels of government, with a focus on minimizing corruption and enhancing accountability. (Ntim et al., 2020). Incorporating ML in the procurement process helps to automatically mine large amounts of data, recognize trends, and even find corruption that would otherwise be difficult to detect (Peterson & Ogden, 2019).

2.3.2 Relationships between Concepts

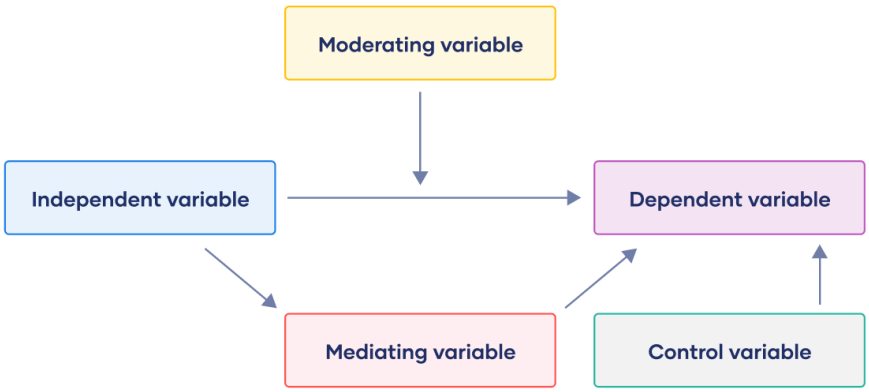
The proposed system aims to:

- **Improve Transparency:** With the use of ML algorithms, classification of bidders and price estimation will be closer to real market prices; this points out the anomalies that could point toward fraud according to McAfee & Lewis (2018). This reduces manual intervention along with reducing the risk of corruption.

- **Fraud Detection:** By harnessing the power of machine learning models, the system will engage in anomaly detection and hence identify suspicious procurement behavior in real time. (Acemoglu & Restrepo, 2020).

- **Improved Accountability:** The system will record all procurement activities, and hence offer audit trails. This ensures that the procurement officers and suppliers become accountable for whatever questionable actions are taken.

- **Efficiency Gains:** Data analytics will automate procurement processes; hence, efficiency will increase as a significant amount of time will be spent on the evaluation of bids for effective decision-making and reviewing supplier performance.



2.3.3 Objectives Integrated into the Framework

- **Market price estimation:** Machine learning techniques such as regression analysis tend to use the historical purchasing databases to analyze and forecast market prices so as to control the procurement costs in line with market realities (Carter et al., 2018).
- **Bidder categorization:** Machine learning models will be leveraged in classifying suppliers based on various aspects including historic performance records, prices quoted in the past, and other risk related factors enhancing procurement efficiency levels.
- **Fraud Detection & Risk Assessment:** The system will incorporate advanced fraud detection tools that will be developed using various ML models e.g. decision trees and neural networks to ensure that the fraudulent suppliers are detected at an early stage in the purchasing process.
- **Data-Driven Decision Making:** Analytics will also provide predictive insights for future strategies, thus supporting better and more justice procurement decisions (Choudhury et al., 2021).

2.3.4 Theoretical Framework

This would apply to the ***Principal Agent Theory***, in which the government is the principal and contracts various agents to have its procurement needs met. According to the theory, since one of the parties may better understand the process than the other, it would open avenues for massive fraud and corruption. (Ross, 1973) notes that information asymmetry characterizes agency relationships. By using ML and analytics, this system tries to reduce information asymmetry and thus will result in transparency.

At the same time, the integration of ML and analytics into public procurement is justified on the basis of the Institutional Theory, which denotes that organizations especially public organizations adapt to changing environment through adoption of new technologies in order to remain viable and efficient. (DiMaggio & Powell, 1983). These tools are timely since there are challenges in public procurement such as detection of fraud and management of risks (Acemoglu & Restrepo, 2020).

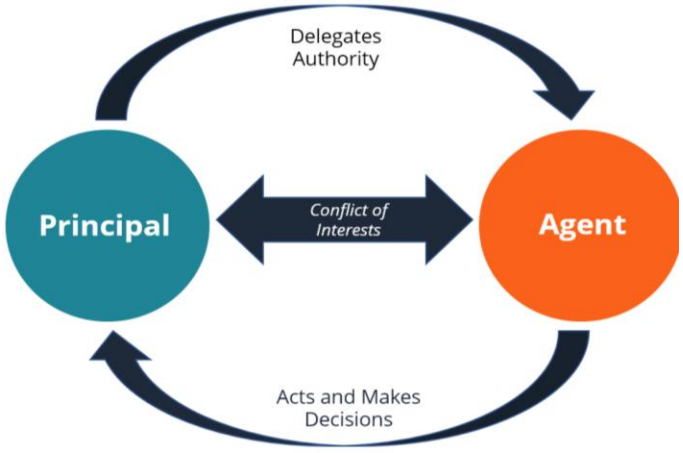
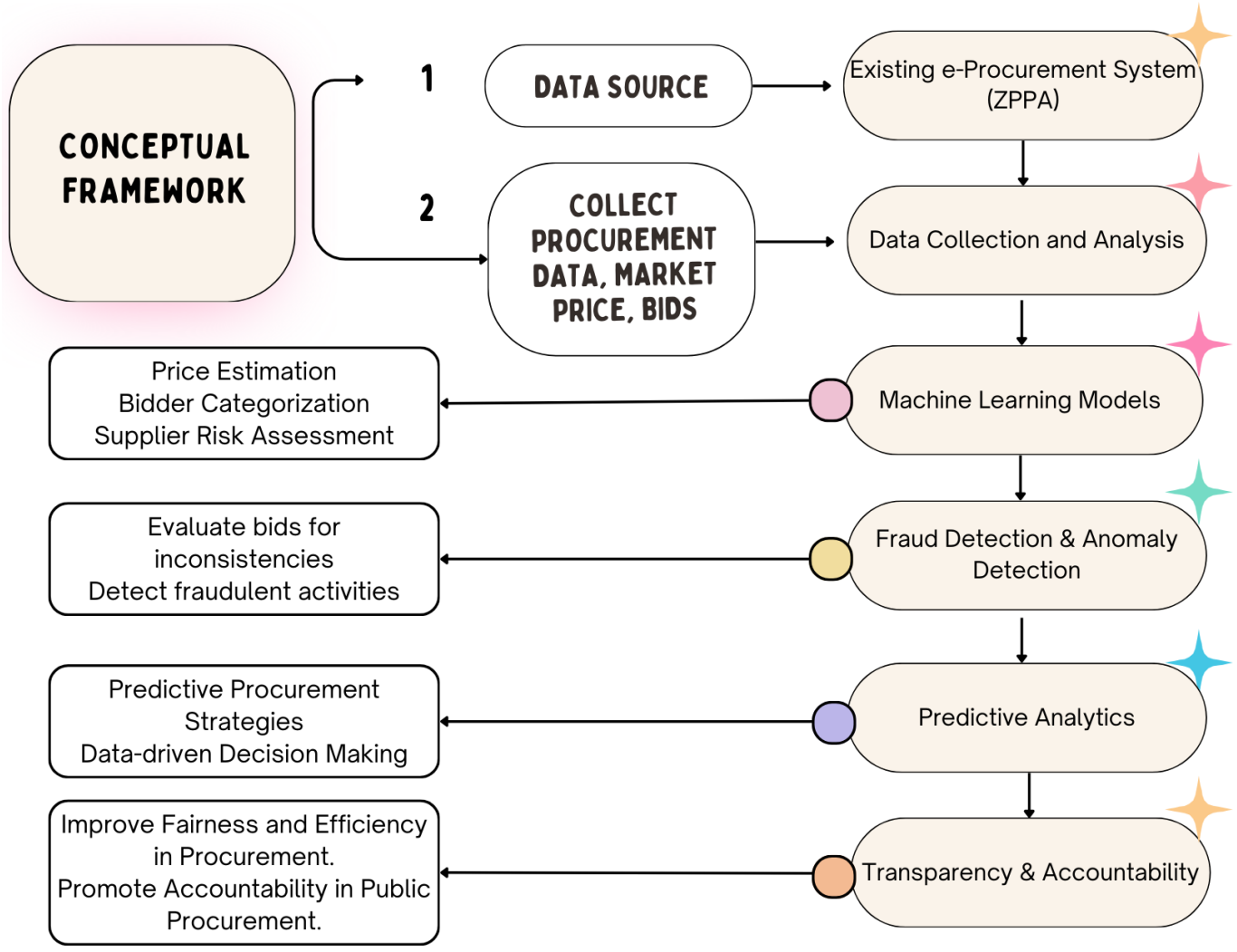


Figure 2.3.4 *Principal-Agent Problem (CFI Foundation, 2015-2024)*

2.3.5 Conceptual Framework Diagram of our Proposed System

This diagram captures the flow of how different components interact in our system concept



- **Existing e-procurement systems:** The system will investigate and gather data from other exiting systems. (ZPPA/e-GP/PACRA)
- **Data Collection and Analysis:** Collected data in terms of market prices, bids, suppliers will be analyzed to gain insight.
- **Machine Learning Models:** Provide the ML models for estimation of price, classification of bidders, and assessment of risks.
- **Anomaly and Fraud Detection:** ML algorithms will detect fraud and abnormalities based on the analysis of data.
- **Predictive Analytics:** The system will provide strategies or aid in decision-making through the extraction of insights from the data.
- **Transparency & Accountability:** The system would try to achieve is enhancing transparency, equitability, and accountability in the procurement process.

2.4 Proposed model/system

The system shall be designed as a web application since it offers a number of advantages, including being highly accessible, easy to use, and flexible. This approach nicely balances the needs of the public procurement processes, which are supposed to be highly transparent, expandable, and effective.

2.4.1 Web Application Model for Transparent e-Procurement System

- **Client-Server Architecture**

The system will follow a client-server architecture where the user (client) interacts with the system via a web browser, while the back-end server processes the data. The client-side handles user input, interface rendering, and simple validation, while the server-side manages business logic, machine learning algorithms, and data storage. This model allows for distributed usage, enabling procurement officials, suppliers, and auditors to access the system from various locations. The architecture also enhances load management and data security by separating front-end and back-end processes (Zheng & Wu, 2021).

- **Multi-Tier Architecture (3-Tier Architecture)**

The system will implement a 3-tier architecture:

- *Presentation Layer:* This layer includes the User Interface (UI) built using web technologies like HTML, CSS, and JavaScript, ensuring responsiveness and a smooth user experience.
- *Business Logic Layer:* It processes user inputs, applies machine learning models (e.g., fraud detection, bidder categorization), and manages workflows such as bid evaluation and supplier risk assessment. This layer will be built using server-side languages like Python (Django/Flask) or Java (Spring Boot).
- *Data Access Layer:* It interacts with the database, retrieving and storing procurement data. A relational database (PostgreSQL or MySQL) will ensure efficient querying and integrity of procurement records (Singh & Jadhav, 2020).

- **RESTful API**

The system will be implementing RESTful APIs, which will allow seamless communications between the components of the web application with third-party systems, such as government portals or verification systems. REST APIs enable integrations with external databases, such as supplier databases or tax records, among others, for real-time validation and thus introduce transparency in the evaluation of bidders. (Kumar & Gupta, 2022).

- **Authentication and Access Control**

The system will provide strong user authentication and role-based access control to ensure that only authorized users can access certain functionalities. For example, procurement officers, suppliers, auditors, and administrators will be assigned different roles, with sensitive operations like bid awarding or finalizing a contract provided only to authorized personnel. OAuth2 or JWT will be used for secure login of users, while SSL/TLS encryption will secure data over the web.

- **AI and Analytics Integration**

Such models will be integrated into the business logic layer through APIs or microservices, where fraud detection and categorization of bidders are performed. In real time or batch processing, they run to analyze procurement data and provide reports. A system can automatically flag irregular patterns of bidding or high-risk suppliers on the basis of real-time data inputs and thus warn procurement officials to take action in this regard.

- **Data Storage and Management**

For structured data management, such as procurement records, bidder details, bid submissions, and contracts, the system will apply a relational database management system, like MySQL and PostgreSQL. The system shall also incorporate inherent mechanisms for ensuring data integrity, managing transactions, and providing automatic backup and recovery processes to prevent data loss (Wang, 2020).

- **Reporting and Dashboard**

The system will provide an interactive dashboard and reporting tools that allow users to generate procurement reports, view performance metrics, and track anomalies. Visualization tools such as Chart.js or D3.js can be used to present data in charts and graphs. This will allow procurement officers to generate fraud detection reports, track procurement timelines, and visualize bidding trends, thereby promoting transparency and accountability (Smith & Johnson, 2018).

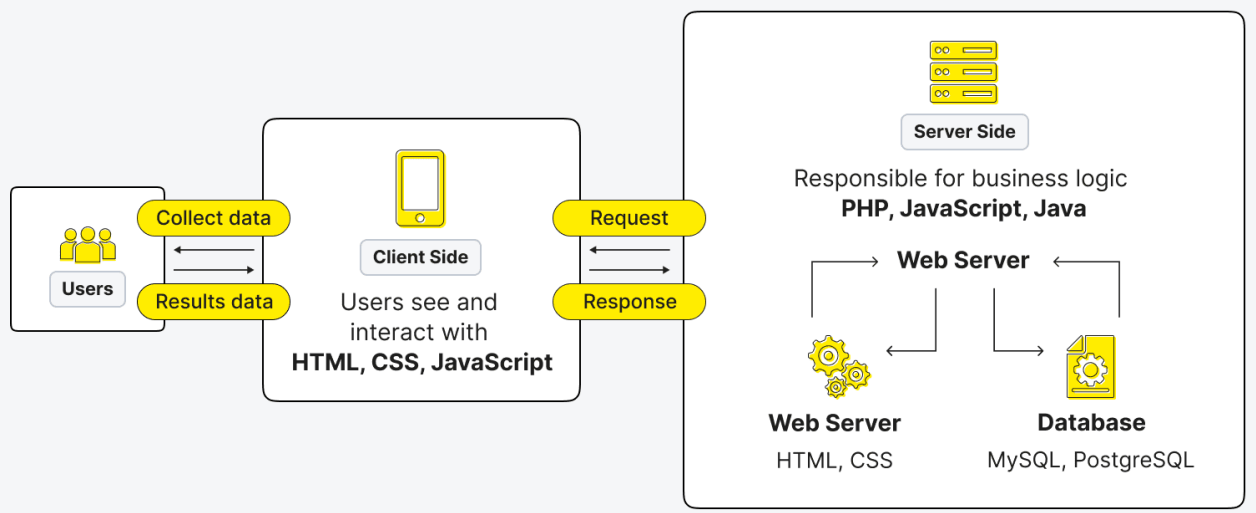


Figure 2.4.1 *Web Application Architecture (Anastasia Moroz, 2022)*

2.5 Comparison with related works

The proposed system is anticipated to have an edge in the detection of fraud, categorizing of bidders, and data-driven decision-making over other existing systems. Below, the main differences between the proposed system and other procurement platforms.

- **Machine Learning for Fraud Detection**

Existing systems such as GeBIZ (Singapore) and PEPPOL (Europe) are more compliance and regulatory transparency centric, but they do not have advanced mechanisms to countercheck fraud in real time. GeBIZ uses some rule-based controls and audits in place to identify problems after procurement, while PEPPOL addresses secure exchange of procurement documents only and does not deal with the issues of fraud prevention while carrying out the process (de Boer, Harink & Heijboer, 2021).

The proposed system will include machine learning algorithms that will permit the assessment of bidding patterns and the reporting of abnormalities as they arise. This enhances fraud detection by managing the risk of corrupt activities going undetected. (Tan & Goh, 2019).

- **Bidder Categorization using AI**

The existing systems, such as ChileCompra in Chile and Procure2Pay in India, do the classification of the bidders using some pre-defined rules or follow some manual judgment methods. These methods, according to Hoffman (2020) and Reddy (2020), are cumbersome, slower, and susceptible to human errors resulting in misplaced decisions in ranking or grading. Bidders fall into wide classes since deep data analysis is not performed, which may lead to less correct risk profiling.

Apart from improving the accuracy in bidding, AI-powered bidder categorization reduces the risk of awarding contracts to high-risk or fraudulent suppliers through real-time data inputs like past contract performance, market behavior, and financial reliability, wherein the system dynamically categorizes suppliers. This is advantageous over traditional systems that employ static categorization methods since AI keeps refining the bidder's profile continuously.

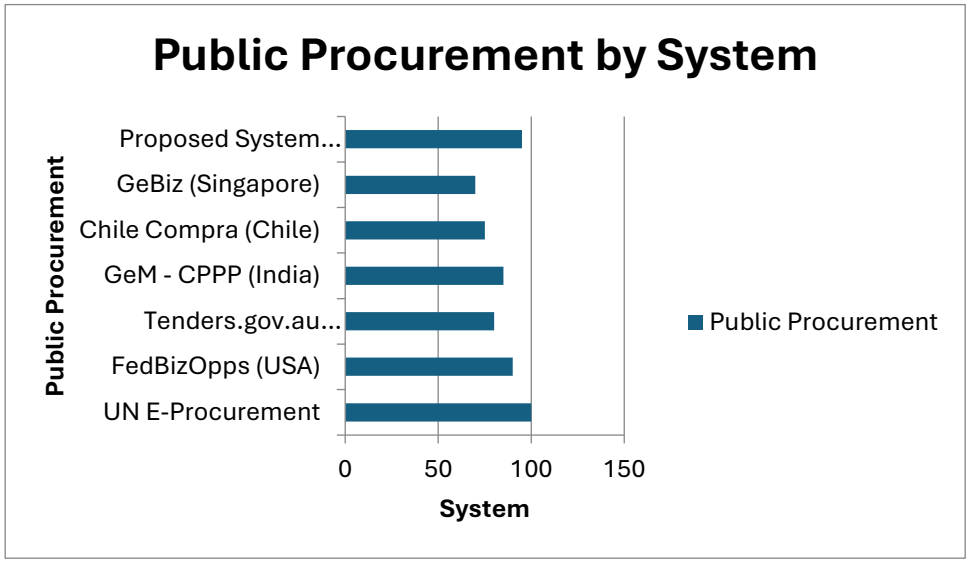
- **Data Analytics for Predictive Procurement and Decision-Making**

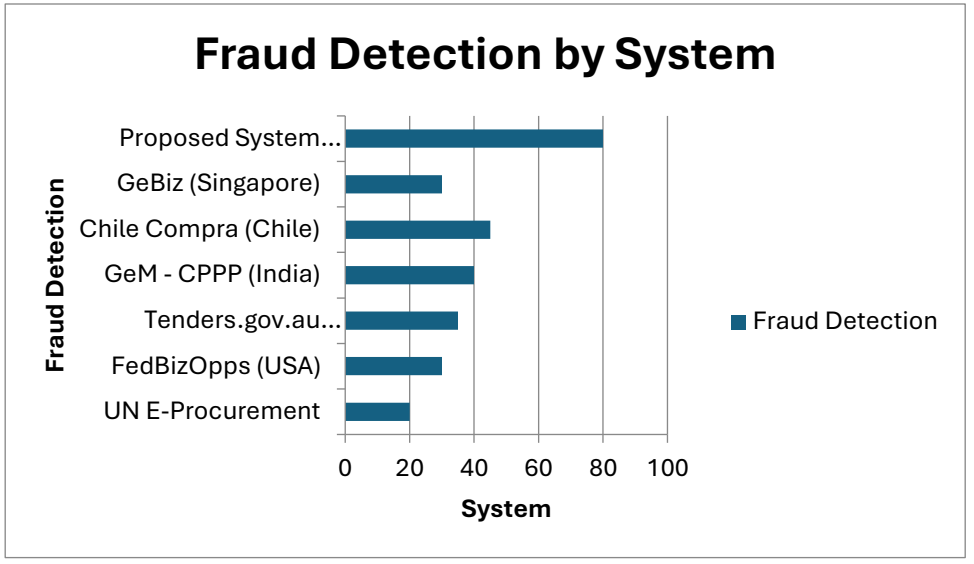
Conventional electronic procurement systems like those found in ChileCompra are oriented towards compliance, documentation, and optimizing transactions with little use of data mining for future intelligence purposes. These types of systems give analysis mainly based on retrospective review of information and basic reporting (Hoffman, 2020).

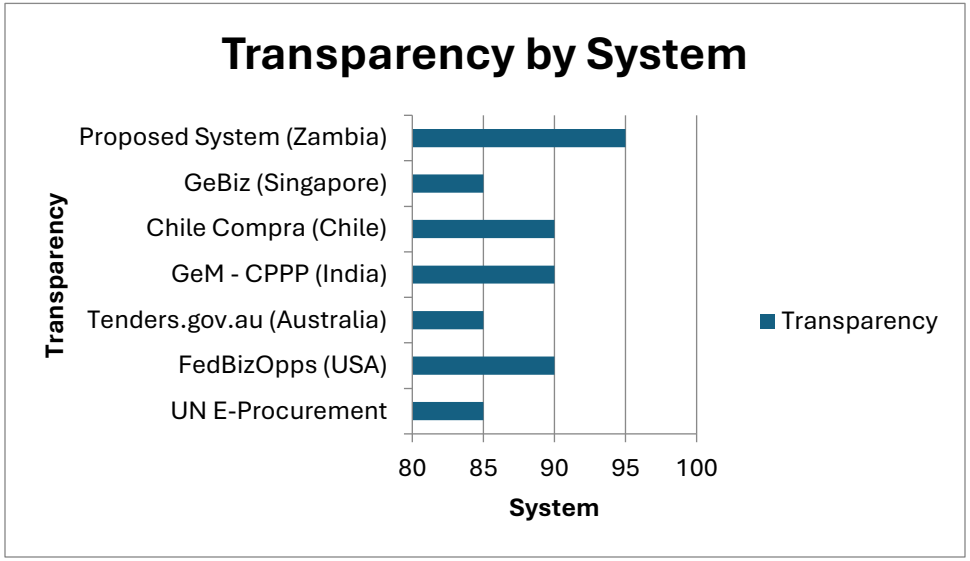
The proposed system intends to incorporate predictive analytics into the system to predict future procurement risks and market trends. Machine learning models use historical procurement data to spot patterns that give procurement officials valuable insight into what might happen in the future. This further helps predict bidder performance, qualifies supplier risk, and spots abnormalities that could suggest fraud.

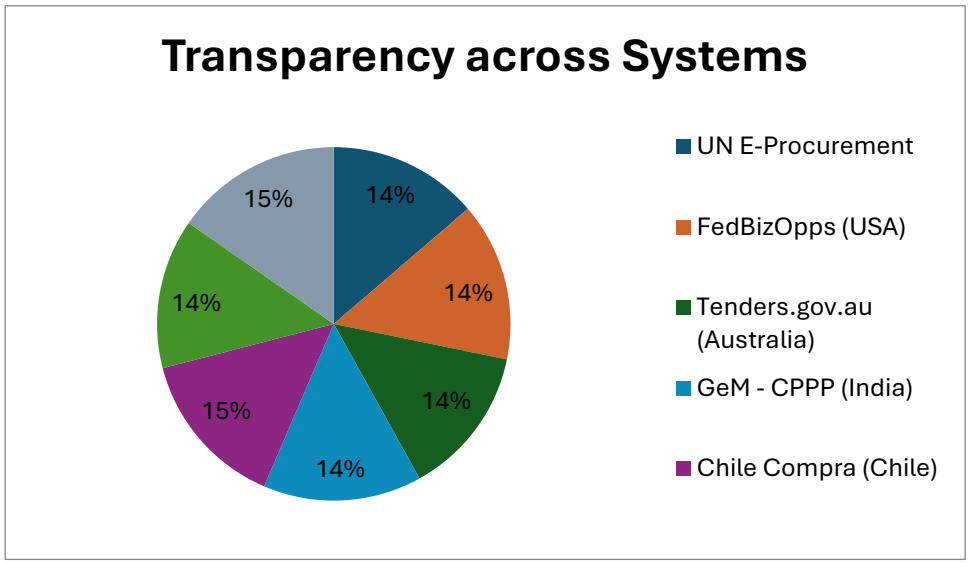
Comparison with related works

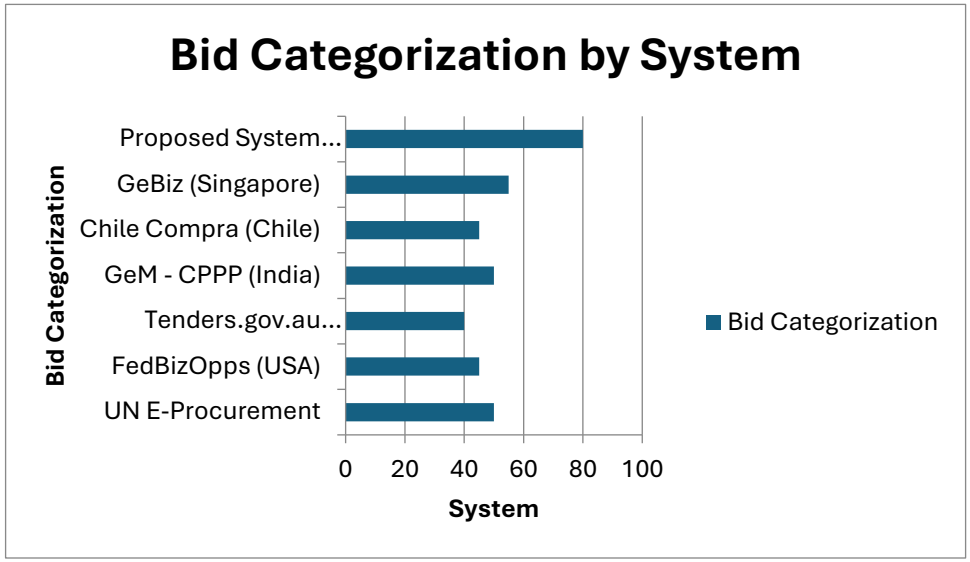
System	Country	Public Procurement	Fraud Detection	Transparency	Bid Categorization	Efficiency	Fairness	Accountability	Predictive Analytics	Data-Driven Decision Making
E-Procurement System by the United Nations	Global	High volume of international tenders; over 100,000 users globally	Limited use, primarily via audit checks	Strong, with public access to many procurement notices	Minimal, mostly categorized by procurement type and funding	Moderate, processes still often require manual input	Prioritized in fairness across nations and contracts	Transparent reporting of procurement data	Low, limited predictive analysis capabilities	Limited, mainly manual analysis
Federal Business Opportunities (FedBizOpps)	USA	Large-scale, with hundreds of thousands of contracts annually	Limited, primarily through compliance and reporting systems	High, public access to contract awards and amendments	Limited to primary contract types	High efficiency via centralized process and digital submissions	Emphasizes equal opportunity for contractors	Publicly available procurement records	Low, lacks advanced predictive analytics	Basic level, uses trends in past procurement
Tenders.gov.au	Australia	Robust, with thousands of contracts across federal and state levels	Basic fraud detection, often reviewed post-hoc	High, all procurement information publicly accessible	Limited bid categorization, primarily by industry	High efficiency due to streamlined online submission	High, equal access is a priority	Strong transparency measures with public accountability	Minimal predictive analytics	Minimal, relies on historical data
E-Procurement System (GeM - CPPP)	India	Significant volume; caters to a large base of government suppliers	Basic fraud detection, periodic audits	High, with transaction transparency	Moderate categorization by procurement type	High, with optimized processes for vendor selection	Strong emphasis on fairness for small businesses	Publicly accessible data for accountability	Low, minimal predictive analytics	Limited, primarily historical data
Chile Compra	Chile	National system handling thousands of public procurements	Basic fraud detection, periodic audits	High, with open access to procurement data	Limited categorization options by contract type	High efficiency via digital submissions	Strong focus on fair access for small vendors	High accountability with regulatory oversight	Limited predictive analytics	Limited, relies on historical purchasing patterns
GeBiz	Singapore	Comprehensive national system with thousands of procurement postings	Low-level fraud detection, mainly manual review	Strong, with open access to procurement data	Moderate, grouped by industry and contract type	High, with streamlined processing	Strong focus on equitable vendor access	Strong, with transparency and regulatory reporting	Minimal predictive analytics	Limited, primarily uses historical trends
Transparent Public Procurement Monitoring System (Proposed)	Zambia	Expected to handle significant volume of national procurement	Advanced, with ML-driven fraud detection capabilities	High, leveraging blockchain and data transparency	Advanced categorization through ML-based analysis	High efficiency with automated processes	Enhanced fairness with data-driven scoring	Strong, with real-time accountability through transparent records	High, with ML for predictive and preventive strategies	High, utilizes ML and analytics for real-time decision-making

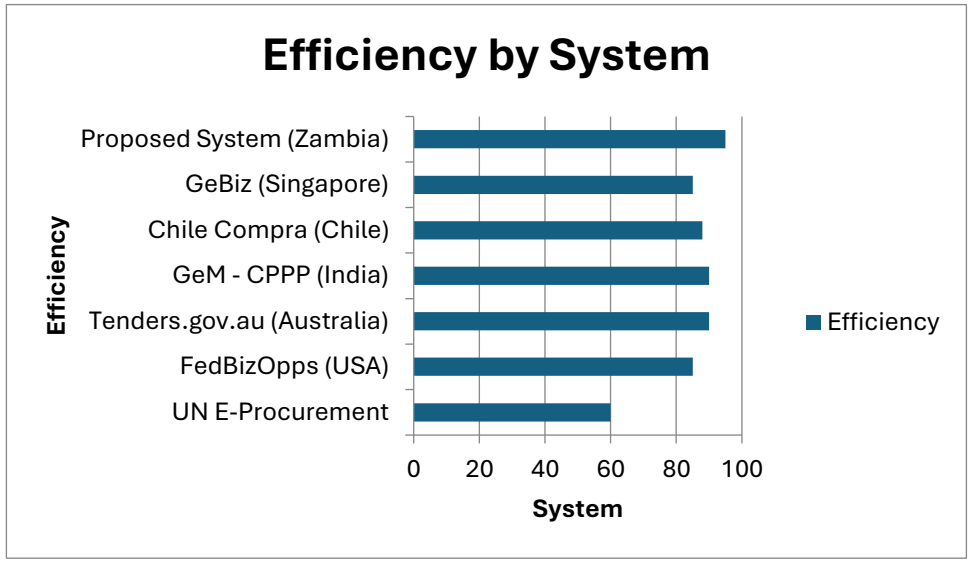

Public Procurement by System
Proposed System...
GeBiz (Singapore)
Chile Compra (Chile)
GeM - CPPP (India)
Tenders.gov.au...
FedBizOpps (USA)
UN E-Procurement
0
50
100
150
System
Public Procurement
Public Procurement

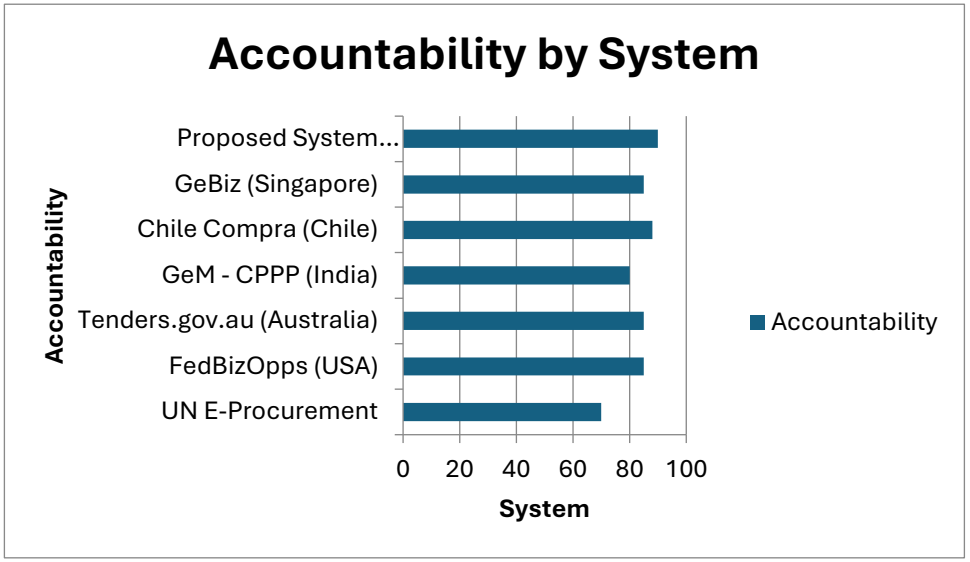

Fraud Detection by System
Proposed System...
GeBiz (Singapore)
Chile Compra (Chile)
GeM - CPPP (India)
Tenders.gov.au...
FedBizOpps (USA)
UN E-Procurement
0
20
40
60
80
100
System
Fraud Detection
Fraud Detection

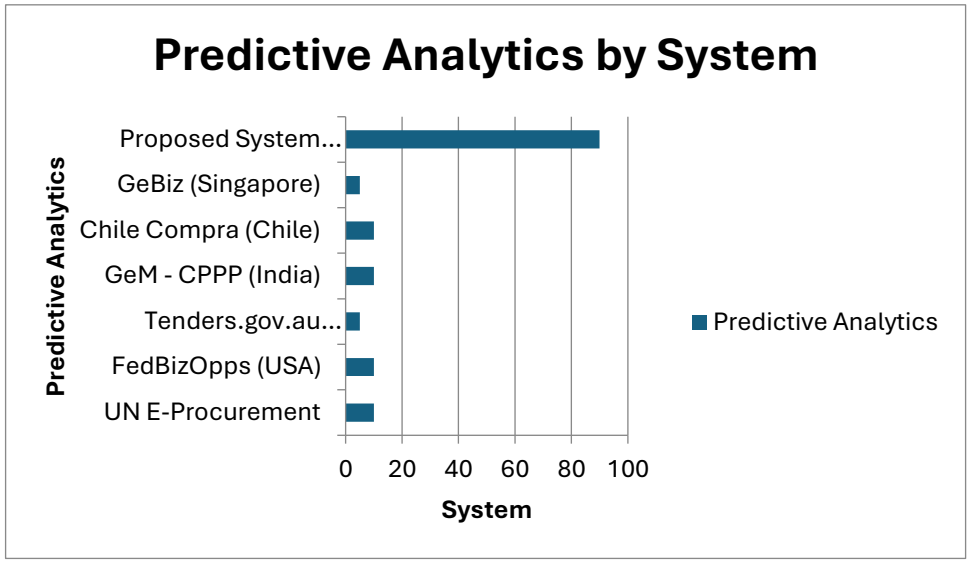

Transparency by System
Proposed System (Zambia)
GeBiz (Singapore)
Chile Compra (Chile)
GeM - CPPP (India)
Tenders.gov.au (Australia)
FedBizOpps (USA)
UN E-Procurement
80
85
90
95
100
System
Transparency
Transparency


Transparency across Systems
15%
14%
14%
14%
14%
15%
UN E-Procurement
FedBizOpps (USA)
Tenders.gov.au (Australia)
GeM - CPPP (India)
Chile Compra (Chile)


Bid Categorization by System
Proposed System...
GeBiz (Singapore)
Chile Compra (Chile)
GeM - CPPP (India)
Tenders.gov.au...
FedBizOpps (USA)
UN E-Procurement
Bid Categorization
0 20 40 60 80 100
System
Bid Categorization


Efficiency by System
Proposed System (Zambia)
GeBiz (Singapore)
Chile Compra (Chile)
GeM - CPPP (India)
Tenders.gov.au (Australia)
FedBizOpps (USA)
UN E-Procurement
Efficiency
0 20 40 60 80 100
System
Efficiency


Accountability by System
Proposed System...
GeBiz (Singapore)
Chile Compra (Chile)
GeM - CPPP (India)
Tenders.gov.au (Australia)
FedBizOpps (USA)
UN E-Procurement
Accountability
0 20 40 60 80 100
System
Accountability


Predictive Analytics by System
Proposed System...
GeBiz (Singapore)
Chile Compra (Chile)
GeM - CPPP (India)
Tenders.gov.au...
FedBizOpps (USA)
UN E-Procurement
Predictive Analytics
0 20 40 60 80 100
System
Predictive Analytics

2.6 Corruption in Zambia: Public Procurement Statistics

- **Corruption Perceptions Index (CPI)**

Rank	Country	CPI Score (100 = Very Clean, 0 = Highly Corrupt)
116	Zambia	34

Transparency International (2022)

- **Bribery Prevalence**

Question	Response Rate
How often do you pay bribes to:	
- Get public services?	24.4%
- Obtain documents or permits?	30.4%
- Secure government contracts?	43.1%

Afrobarometer (2019)

- **Procurement Irregularities**

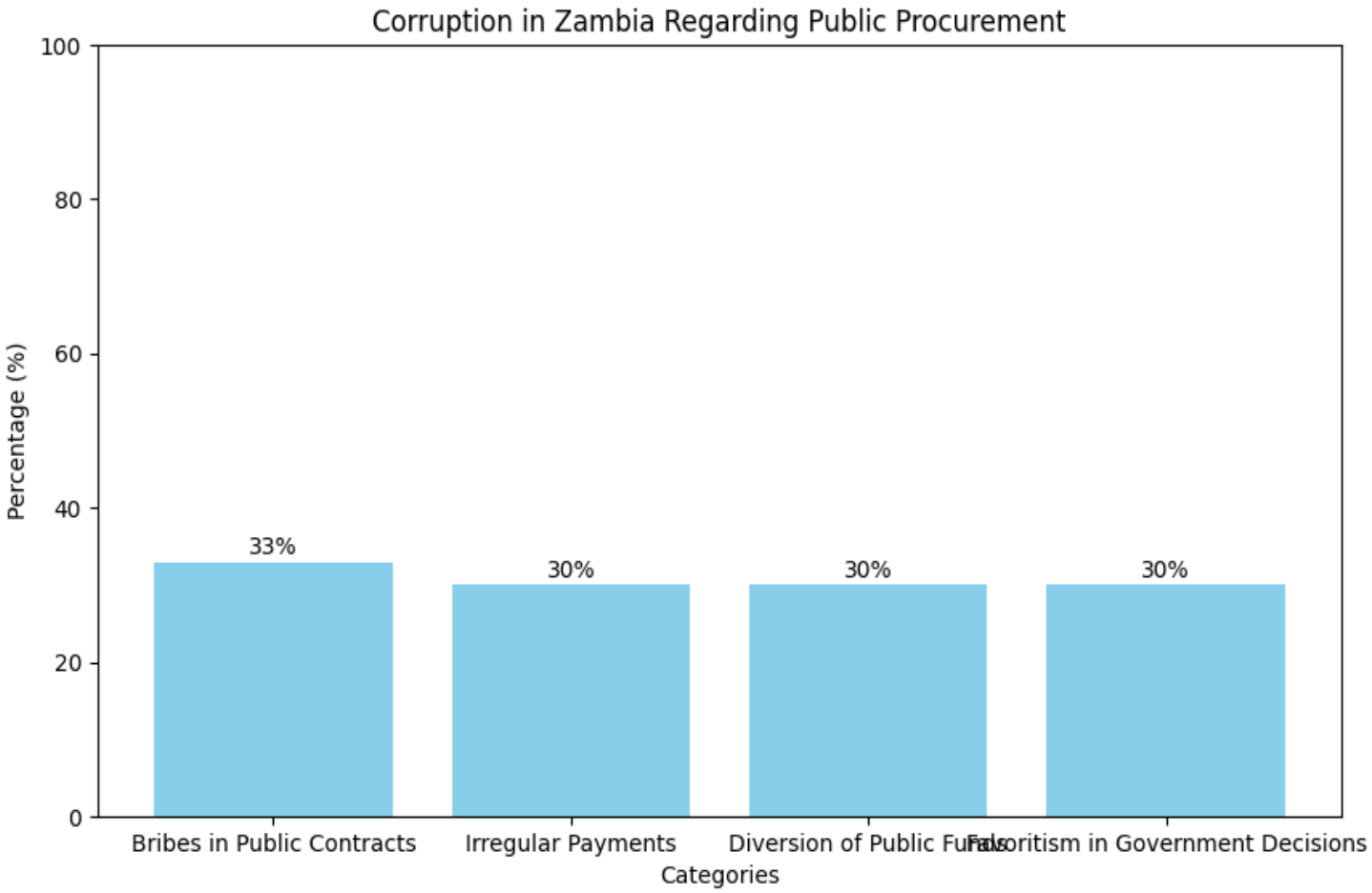
Type of Irregularity	Frequency
Unauthorized expenditures	43%
Lack of documentation	26%
Overpayments	15%
Unaccounted funds	12%
Conflict of interest	4%

Zambia's Auditor General's Report (2020)

- **Corruption Reporting**

Type of Corruption	Reports Received
Corruption in procurement	35.7%
Bribery	26.4%
Embezzlement	20.5%
Abuse of office	17.4%

Zambia's Anti-Corruption Commission (2020)



CHAPTER 3 METHODOLOGY

3.1 Research design

The research design adopted for this project is based on the *Design Science Research (DSR)* methodology. This approach is well-suited to system development projects where the primary aim is to create and evaluate IT-based solutions to solve real-world problems (Hevner et al., 2004). In this case, the system being developed integrates machine learning (ML) and data analytics to improve transparency, efficiency, and accountability in public e-procurement processes.

The DSR methodology involves iterative phases of problem identification, system development, implementation, and evaluation. This iterative approach ensures that the system's functionality evolves based on continuous feedback from stakeholders, and empirical testing using procurement data (Peffers et al., 2007).

The research will also utilize both qualitative and quantitative methods. Qualitative methods will be employed in the initial phase to analyze existing procurement systems and identify gaps where ML and analytics can enhance transparency. Quantitative methods will be used to evaluate the system's performance in terms of fraud detection, market price estimation, and supplier risk assessment.

3.2 Adopted method and justification

3.2.1 System Design and Development Tools

In this vein, the proposed system shall follow the Agile methodology, which supports flexibility and adaptability in software development. In this respect, Agile is iterative, with feedback from stakeholders leading to system evolution, so it supports continuous improvement across the development cycle, quite important while developing complex systems integrating ML models and analytics.

It will be developed using a set of technologies which will ensure the efficiency of data processing, scalability, and user experience. The components are:

Backend

- *PHP* will be utilized for server-side scripting, allowing for dynamic content management. This ensures that the system can process and deliver information in real-time, improving interaction between the server and the client.

- *Python* will be integrated for backend automation and advanced processing, particularly for machine learning tasks. Python's extensive libraries, such as Scikit-learn and TensorFlow, will facilitate the training and deployment of ML models for fraud detection, price estimation, and bidder categorization (Pedregosa et al., 2011).

- *JSON* will be used for efficient data exchange between the server and client. Its lightweight format allows for fast and seamless transmission of data, reducing latency in system responses.

Frontend:

- *HTML* will be used to structure and present content on the web pages. This will ensure that the system is accessible and easy to navigate, providing procurement officers and stakeholders with a user-friendly interface.

- Bootstrap will be employed for responsive front-end design. This framework ensures that the system is mobile-friendly and provides a consistent user interface across different devices and screen sizes (Welling & Thomson, 2009).

Database:

- *MySQL* or *PostgreSQL* databases will be used to store procurement-related data, including bid submissions, supplier information, and historical transaction records. Both databases offer robust performance and scalability, which is critical when handling large datasets in public procurement systems (Chaudhuri & Dayal, 1997). The choice of database will be determined based on the system's specific scalability and performance requirements.

3.2.3 Data Collection and Preprocessing

The objective of the data collection will be to collect information concerning procurement activities, which include but are not limited to past bidding price trends, suppliers and their performance, and records of transactions in relation to their suppliers. The data is to be primary data from open access public procurement primary data archives or secondary data where real time data is unavailable. This data will be pre-enprocessed for checking, cleaning, and readying the data for the purpose of applying machine learning techniques.

3.2.4 Machine Learning Model Development

Various machine learning models will be designed and integrated into the system:

- **Supervised Learning:** The estimation of market price and detection of fraud can be done using algorithms like linear regression and support vector machines. These are models that work well with structured data and can predict an outcome based on historical data about procurements.
- **Unsupervised Learning:** Clustering algorithms such as k-means will be used in classifying suppliers and outlier detection. Such models will enable automatic detection of patterns in the data, hence making meaningful insights on bidder behavior and associated risks possible

Each model will be trained on 70% of the data and tested on the remaining 30% to ensure generalizability. Performance metrics such as accuracy, precision, recall, and F1-score will be used to evaluate the models.

3.3 Systems Analysis

According to (NCC 2008, P. 5-61) "*If the analysts fail to understand the requirements of the system, the chances of it turning into an effective solution are slim*". This section therefore focuses on the systems analysis of the system to be designed and implemented with the core knowledge of the current ZPPA/e-GP system.

3.3.1 Current System and Its Disadvantages

The current ZPPA/e-GP system faces significant limitations due to its lack of integration with key institutions like PACRA and ZRA. Additionally, procurement officers are unable to access crucial information on suppliers, such as registration details or tax compliance, without manual verification. Further, the system does not flag anomalies or potential corruption risks, leading to inefficiencies in detecting irregularities. Its lack of transparency further hampers accountability, as it does not provide stakeholders with real-time data or visibility into procurement activities. These issues collectively undermine the effectiveness and integrity of the procurement process.

Example of Current ZPPA/e-GP Three Tier Distributed System

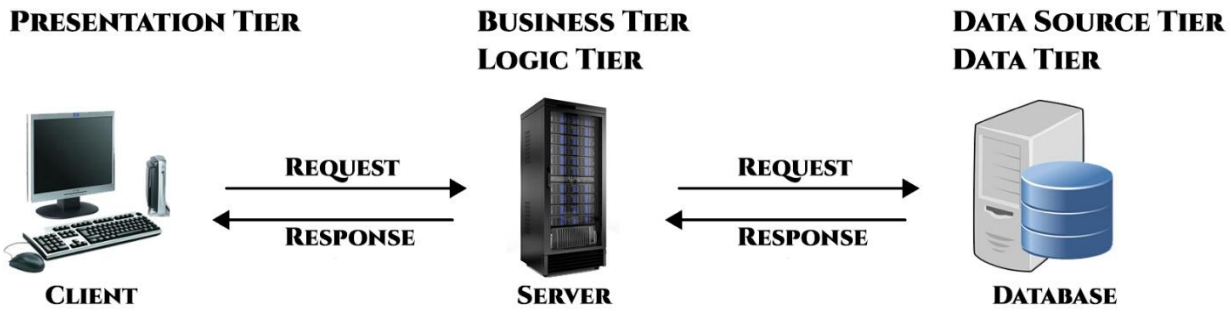


Figure 3.3.1 *ZPPA/e-GP Multi-Tier Architecture example (Shafi Shaik, 2022)*

The architecture described for the ZPPA/e-GP system can be categorized as Multi-Tier Architecture (or N-Tier Architecture). This architectural style separates the system into distinct layers, each responsible for specific aspects of the system's functionality.

- *Presentation Layer (User Interface Layer):* This is where users interact with the system, typically through a web portal or application interface.
- *Application Layer (Business Logic Layer):* This layer contains the core functionality and business logic of the application, managing processes like procurement management, bid evaluation, and supplier registration.
- *Data Layer:* This tier handles data storage and management, interacting with databases to retrieve and store information.
- *Security Layer:* Although not always categorized as a separate layer, security measures are critical throughout the architecture to protect data and manage user access. (ZPPA, 2024)

3.3.2 Requirements Specifications

In order to achieve goals for the proposed system, functional and non-functional requirements must clearly be defined. This is done so as to have an enlightened requirements specification for the new system being developed.

- **Functional requirements Table**

"These are requirements that concentrate on what the proposed system should do in order to meet user requests". Bennett, McRobb and Farmer (2005, p.131)

Module	Functional Requirement	Description
Analysis	Analyze existing e-Procurement systems	Identify areas for ML/Analytics improvement
Analysis	Identify data sources	Integrate relevant data feeds
Market Price Estimation	Develop ML model for market price estimation	Provide estimated market price reports
Bidder Analysis	Develop ML model for bidder analysis	Assess bidder credibility/risk
Bidder Analysis	Provide bidder classification reports	
Fraud Detection	Develop ML models for fraud detection	Alert potential fraud/high-risk suppliers
Fraud Detection	Assess supplier risk	Evaluate bid evaluation models
Predictive Procurement	Apply data analytics for predictive strategies	Develop data-driven decision-making tools
Predictive Procurement	Provide procurement forecasting reports	
Efficiency, Fairness, Accountability	Assess ML/Analytics impact on efficiency	Evaluate fairness in procurement
Efficiency, Fairness, Accountability	Monitor accountability	Provide performance reports
User Interface	Develop intuitive user interface	Provide real-time analytics dashboards
Integration	Integrate with existing procurement systems	Ensure seamless data exchange
Security	Implement robust security measures	Ensure data encryption/access controls

- **Non - Functional requirements Table**

Non-Functional Requirement	Description
Scalability	Adapt to increasing data/usage
Reliability	Ensure consistent system performance
Performance	Optimize processing speed/efficiency
Usability	Ensure user-friendly interface
Security	Protect data/access with encryption/access controls

- **Assumptions and Dependencies Table**

Assumption/Dependency	Description
Existing system data availability	Access to historical procurement data
Market data feeds	Access to relevant market data
Procurement expert collaboration	Input from procurement experts

- **Acceptance Criteria Table**

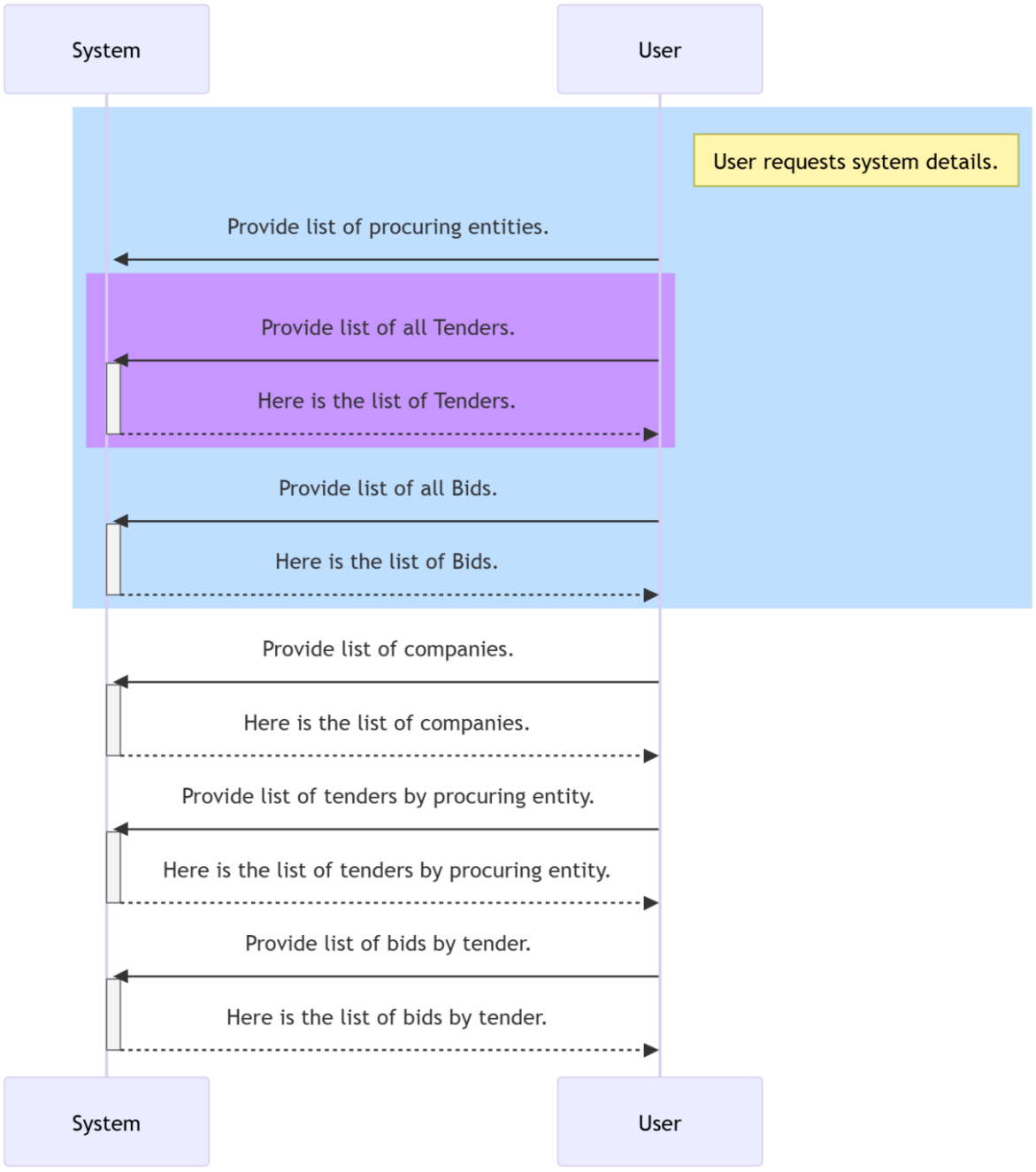
Acceptance Criterion	Description
System integration	Successful integration with existing systems
Market price estimation accuracy	Accurate market price estimation
Fraud detection effectiveness	Effective fraud detection/risk assessment
Procurement efficiency improvement	Improved procurement efficiency/fairness

3.4 Systems Design

Works of Xiaoping Jia (2007 P. 4) describe system design as: "*the primary concern of decomposing complex problems into manageable components*". Simply put, bringing to light the analyzed requirements discussed in the previous chapter in terms of procedures and processes needed to develop the proposed system.

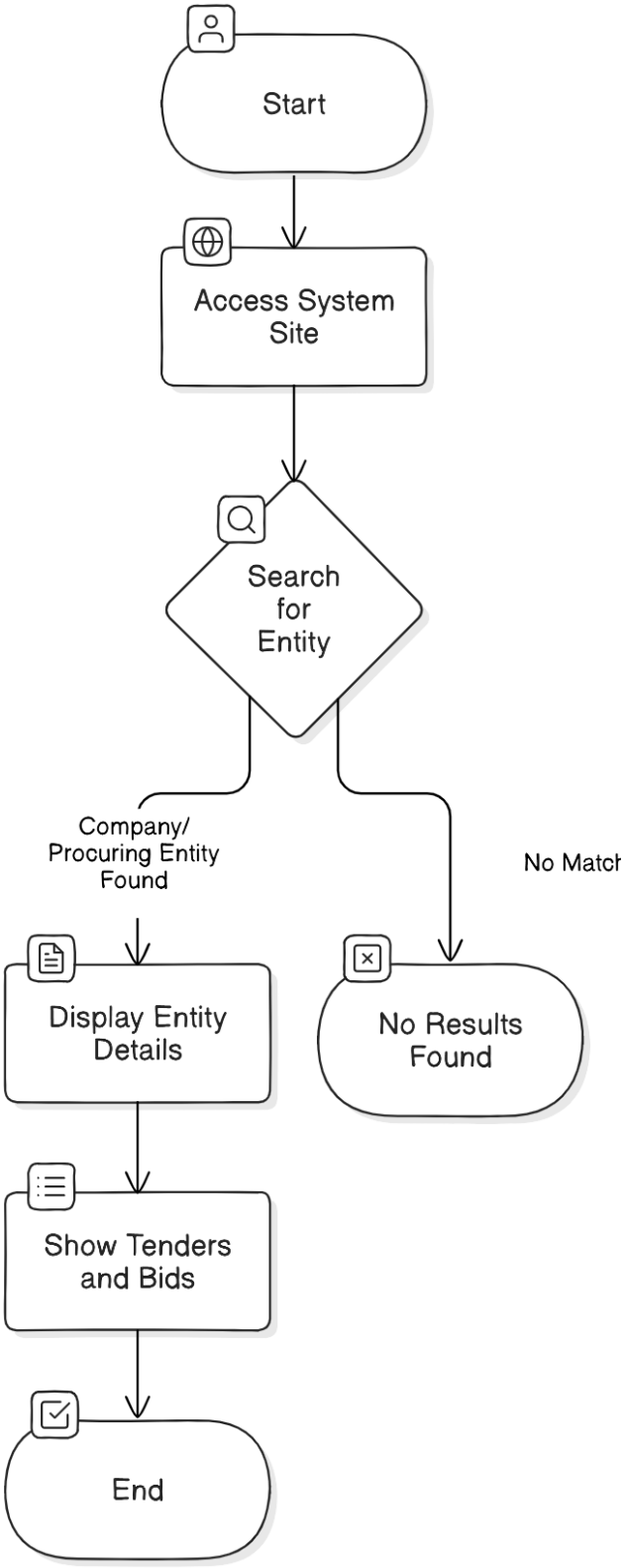
3.4.1 Systems Interface Scenario

The system interface should give the following details:

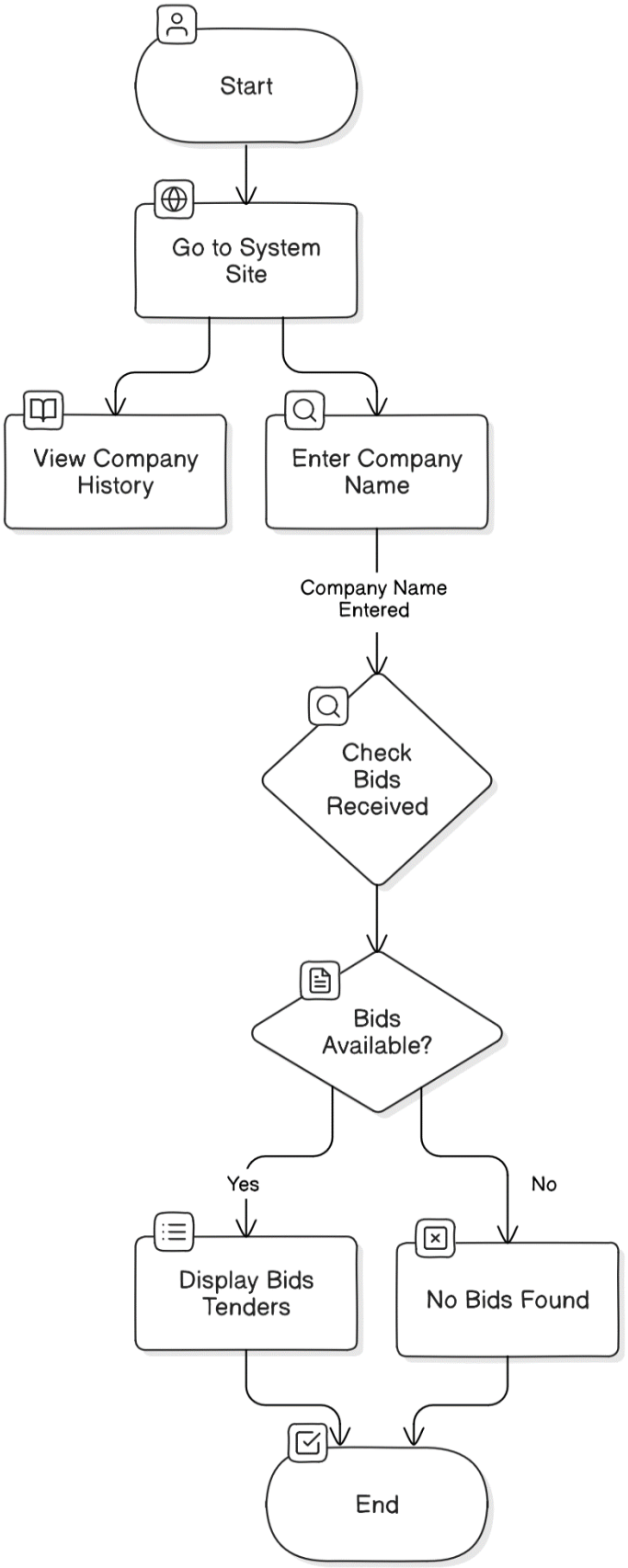


3.4.2 Activity Diagrams

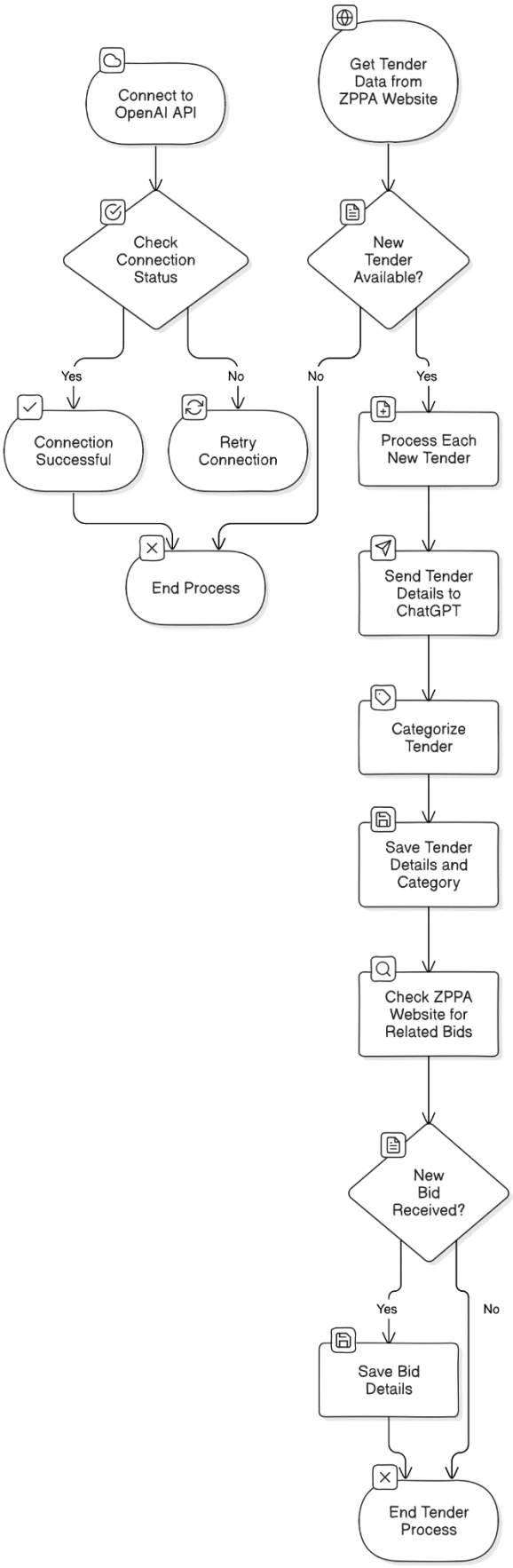
- **Scenario One (1)**



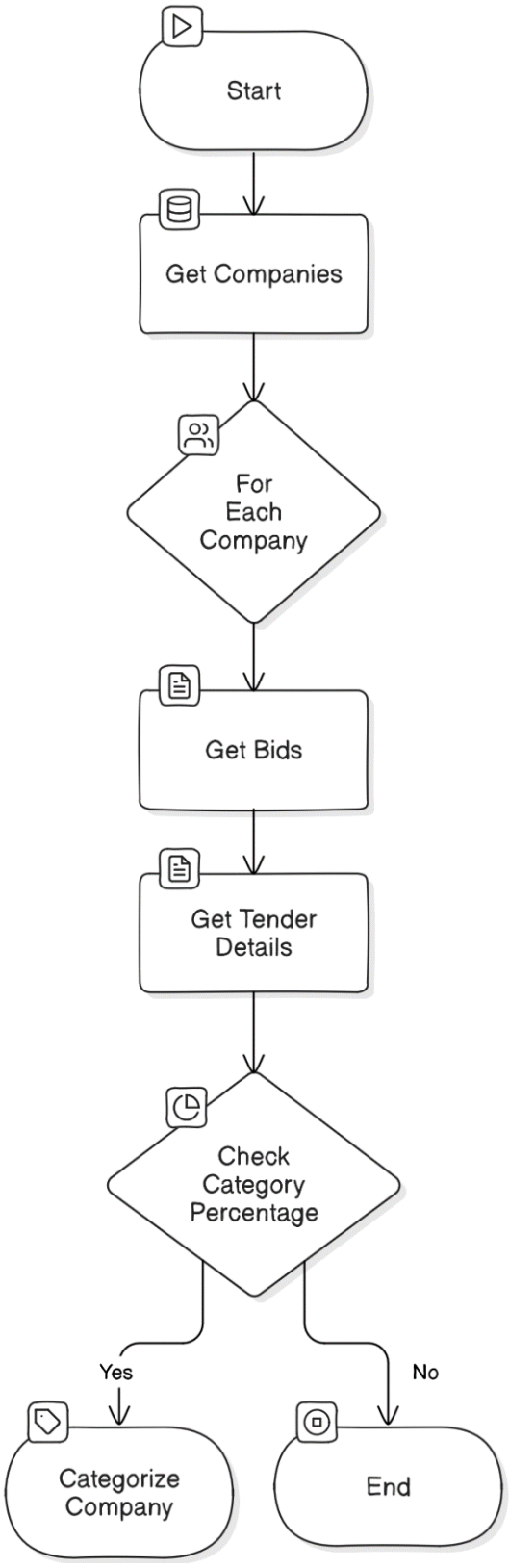
- **Scenario Two (2)**



- **Collection and Categorization of Tenders**



- **Categorization of Companies**



- **System Flow (Implemented Scenario)**

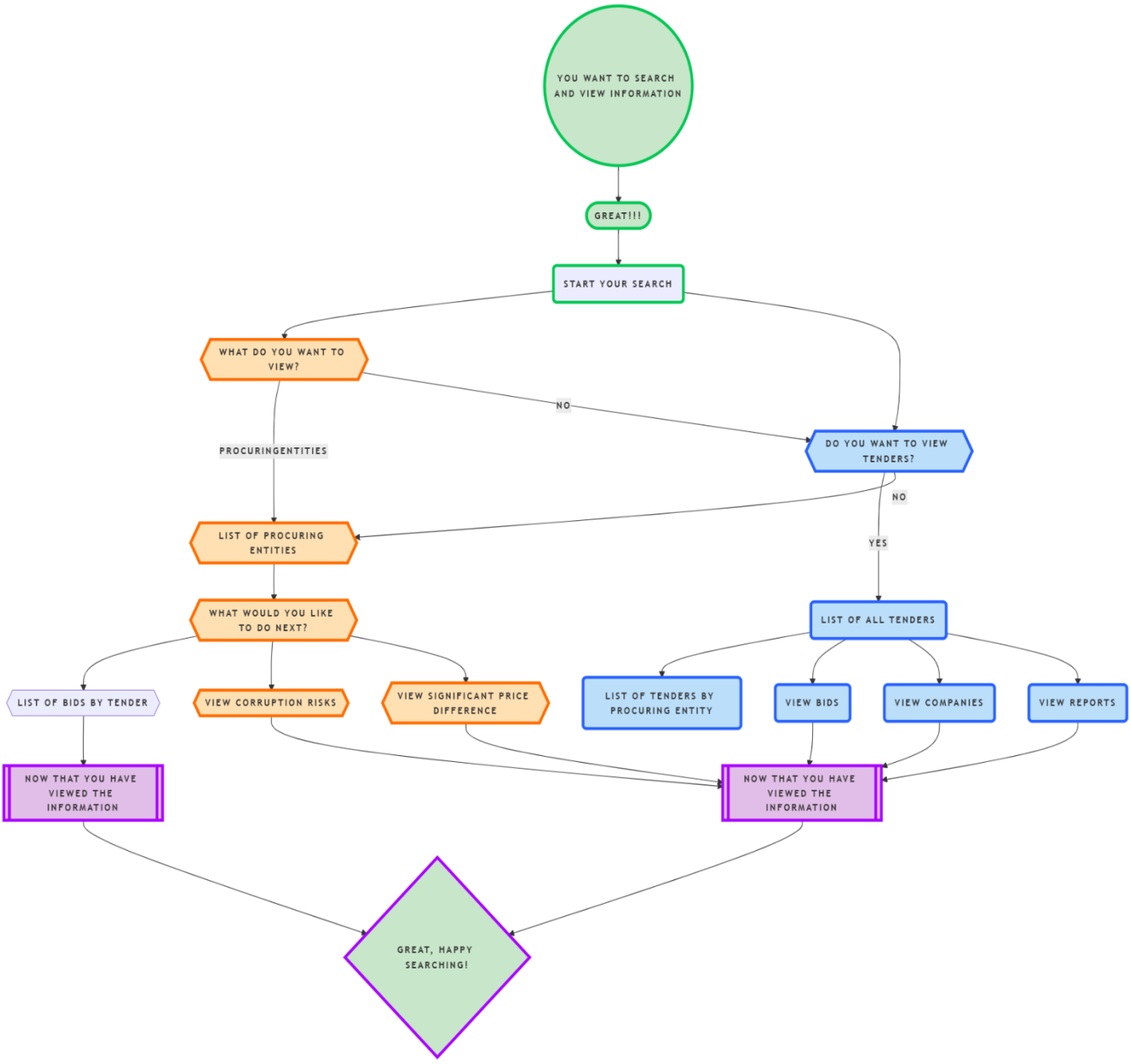


Figure 3.4.2 *Proposed System Outlook*

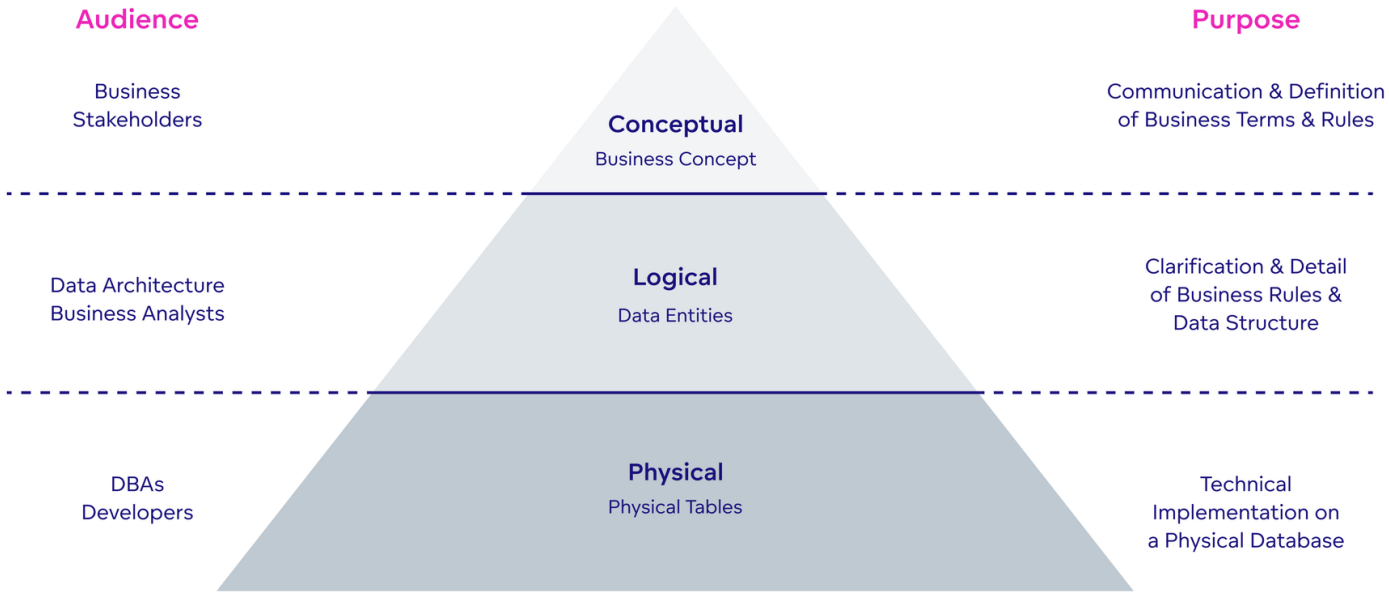
3.5 Database Design

3.5.1 Conceptual Data Model Diagram

"*The first step in conceptual database is to build one or more data models of the data requirements of the enterprise*". A conceptual data model contains: "Entity Types, Relationship Types, Attributes and Attributes Domains, Primary Key and Alternate Key and lastly Integrity Constraints". States Connolly and Begg (2010, P. 420)

Why Conceptual data Model?

"*Provides a visual representation of theoretical constructs (and variables) of interest*". "*Theories present a systematic view of phenomena by specifying relations among variables using a set of interrelated constructs/variables, definitions and propositions* (Kerlinger, 1979)."



Business Rules:

- *A Tender Notice must reference a Procuring Entity*

Every tender notice is linked to the procurement entity that issues it, and each tender relates to a specific procurement plan. The reference number of the tender should match the procuring entity's active procurement plans.

- *A Tender can be associated with one or more categories.*

Each tender must specify the industry or industries to which it belongs. For example, a tender can be relevant to the "Construction" and "IT Services" industries.

- *A Bid must be associated with a Tender Notice.*

Each bid submitted by a company must reference a specific tender notice. A bid cannot exist without a tender to bid for.

- *A Company may submit multiple bids for different tenders, but each bid must be unique to the tender.*

Companies can participate in multiple tenders, but a company cannot submit multiple bids for the same tender (unless the tender procedure explicitly allows it).

- *Each Bid must include a financial value.*

Bids must specify the financial value (the amount the company is bidding) and whether any bid security has been provided (i.e., a guarantee that the bidder can meet the tender requirements).

- *A Company can belong to multiple industries.*

Companies may operate across multiple industries. This is represented by the many-to-many relationship between the Company and the Industry.

- *A Tender Notice must specify at least one of the following: works, supplies, or services.*

Tenders can be for different types of procurement (e.g., construction works, supply of goods, or services), and every tender notice must indicate what it covers.

- *All tenders must include an expected delivery date and submission deadline.*

A tender notice must define an expected delivery date for the goods/services and a deadline for the submission of bids.

- *A tender with only one bidder should be flagged as a corruption risk.*

- *When there are only two bidders and the percentage difference between them is greater than 50% both bids and the tender and the bids should be flagged as suspicious.*

- *When there are more than two bidders, extremely high and extremely low bids (and the tender) should be marked as suspicious.*

3.5.2 Entity Relationship Diagram

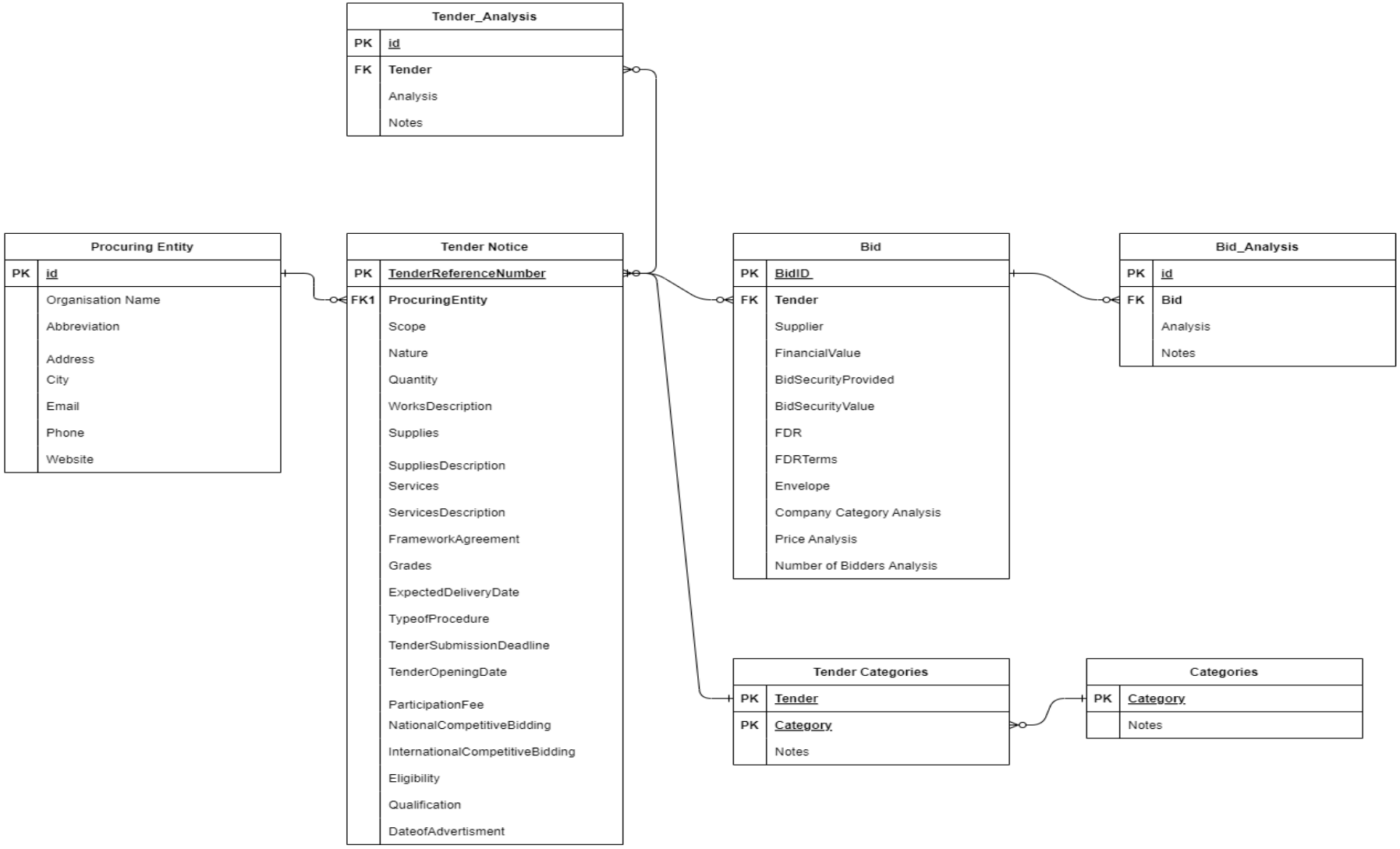


Figure 3.5.2 *Entity Relationship Diagram*

3.5.3 Relationships between Tables

- Tender Analysis → Tender

Relationship: One-to-One
TenderAnalysis.TenderID references Tender.TenderID

- Tender Notice → Procuring Entity

Relationship: Many-to-One
TenderNotice.ProcuringEntityID references ProcuringEntity.ProcuringEntityID

- Bid → Tender

Relationship: Many-to-One
Bid.TenderID references Tender.TenderID

- Bid Analysis → Bid

Relationship: One-to-One
BidAnalysis.BidID references Bid.BidID

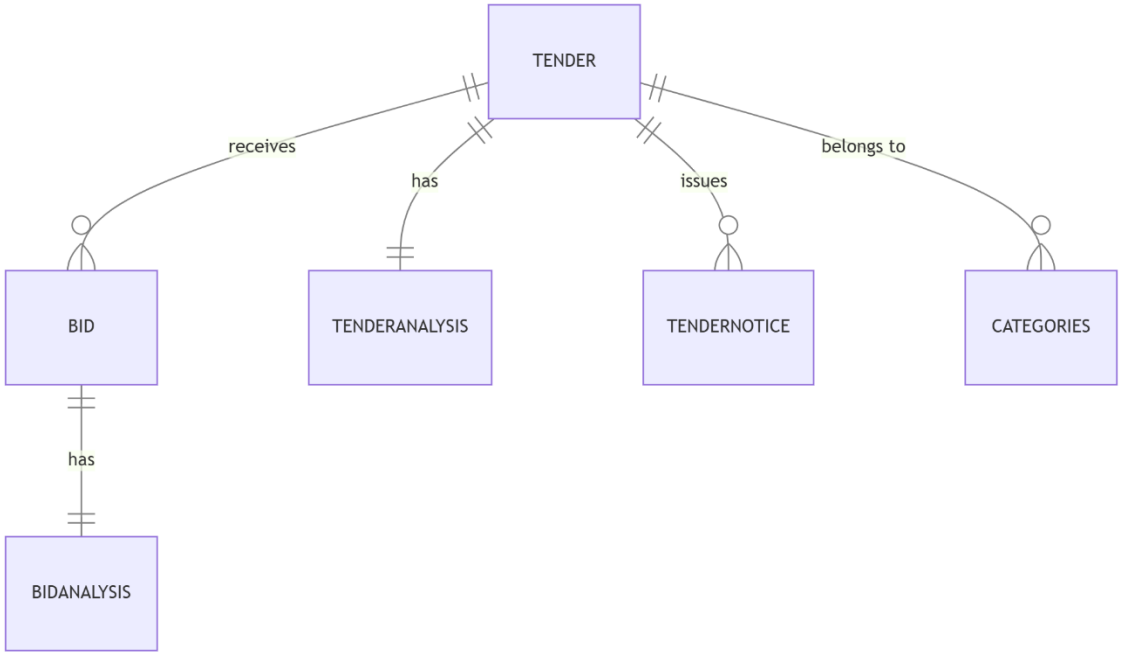
- Tender → Categories

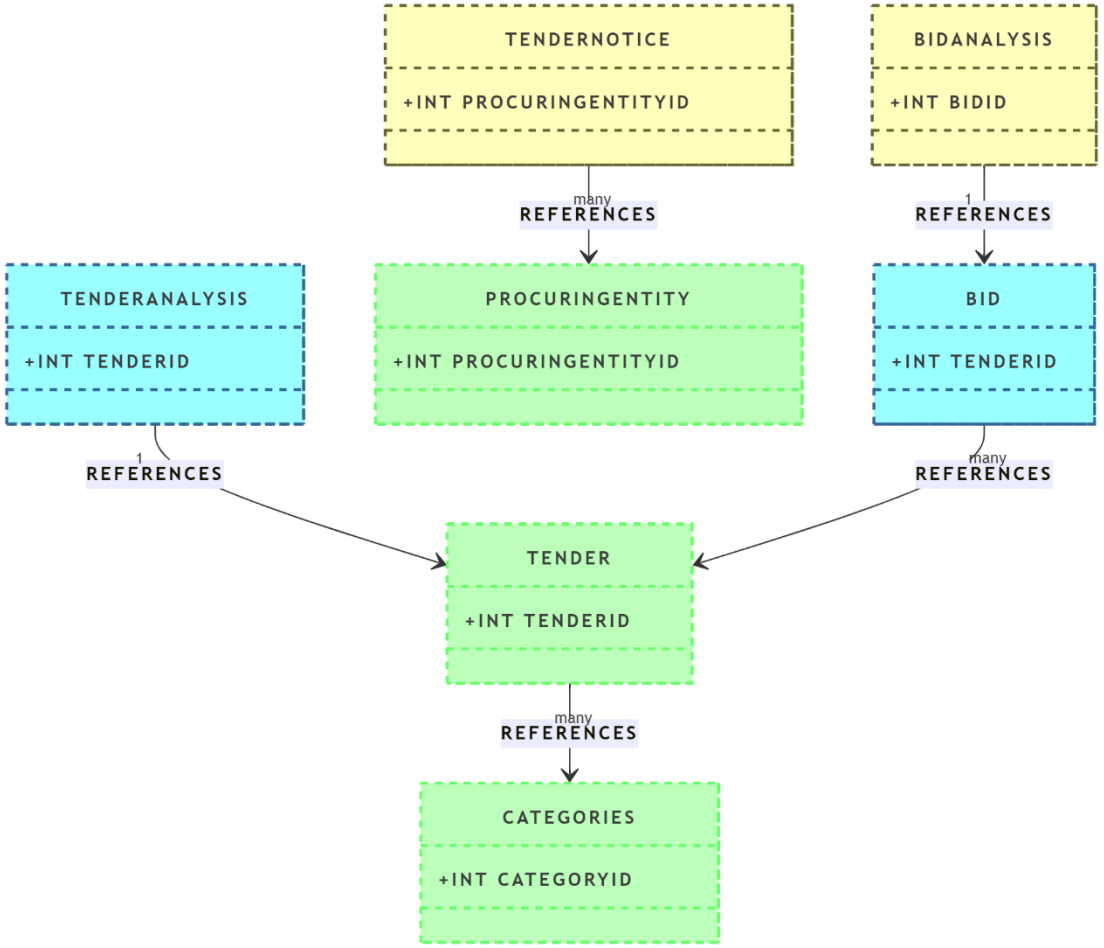
Relationship: Many-to-One
Tender.CategoryID references Categories.CategoryID

Relationships

Entity	Primary Key (PK)	Foreign Key (FK)	Relationship
Tender Analysis	TenderID	TenderID → Tender.TenderID	One-to-One
Tender Notice	TenderReferenceNumber	ProcuringEntityID → ProcuringEntity.ProcuringEntityID	Many-to-One
Bid	BidID	TenderID → Tender.TenderID	Many-to-One
Bid Analysis	BidAnalysisID	BidID → Bid.BidID	One-to-One
Tender	TenderID	CategoryID → Categories.CategoryID	Many-to-One
Procuring Entity	ProcuringEntityID	N/A	N/A
Categories	CategoryID	N/A	N/A

Other Related Diagrams





3.5.4 AI Automation in Tender Categorization

AI will play a critical role in automating the categorization process, ensuring that tenders are accurately classified according to their respective industries. This process involves analyzing tender details such as scope, nature, and specific requirements, and determining which industry they fall under. The AI system will leverage machine learning algorithms to continuously improve its accuracy in identifying relevant industries for each tender.

- *Categorization of Tender*

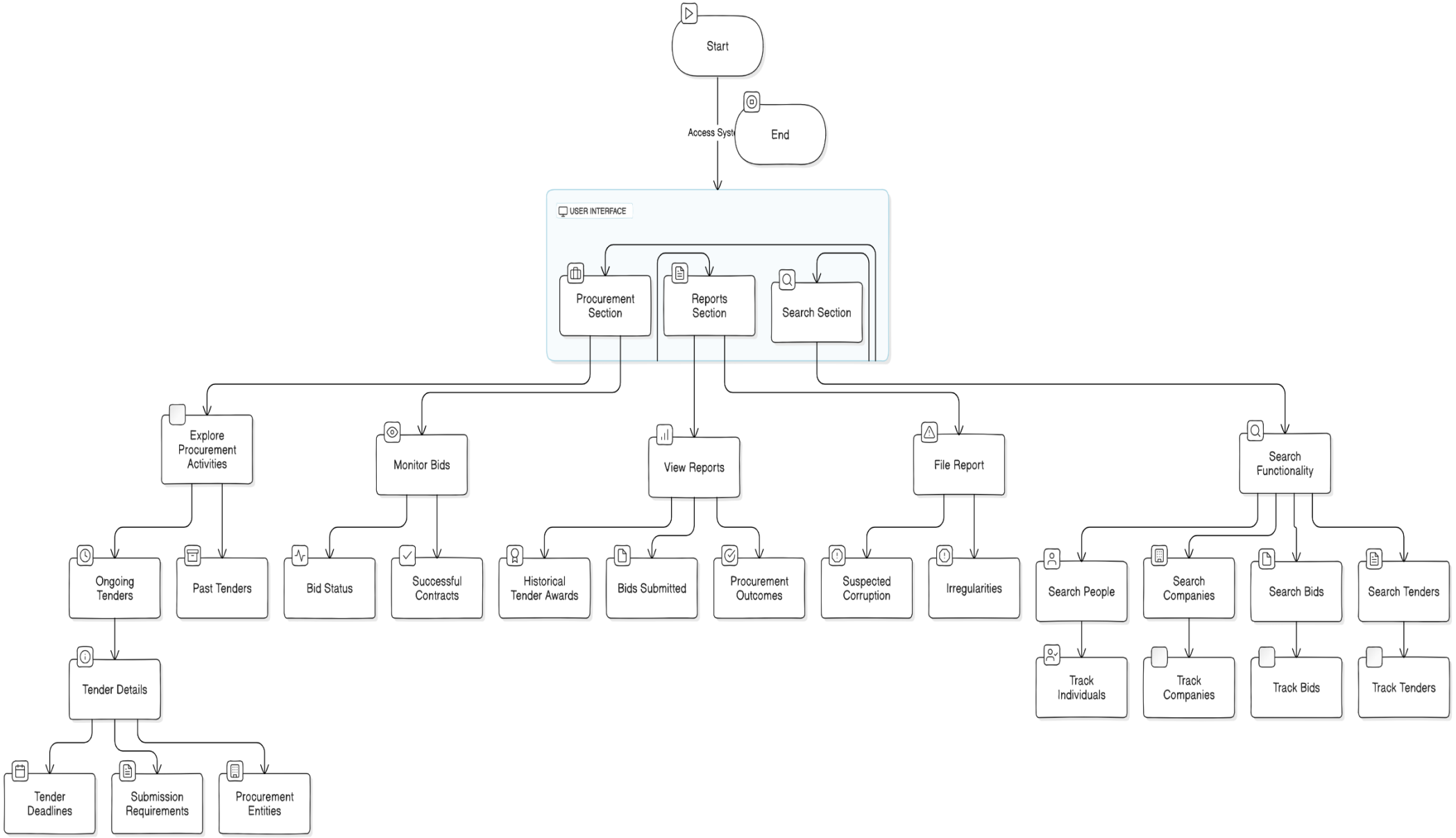
The AI system will assess and categorize tenders based on their content, including the scope of work, services, supplies, or products involved. By doing so, it will determine the most appropriate industry classification for each tender, ensuring that tenders are accurately grouped within relevant sectors. This automated process will allow for greater consistency and accuracy compared to manual categorization.

- *Fraud Detection and Flagging Suspicious Activity*

The AI system will also be equipped to flag suspicious tenders and detect potential fraud or corrupt activities. It will analyze patterns, pricing anomalies, and irregularities in tender submissions. For example, tenders that appear significantly overpriced or show patterns of consistent mis categorization may be flagged for further investigation. The AI will also help identify potential conflicts of interest or repeated instances of certain companies winning tenders under questionable circumstances.

By leveraging advanced analytics and machine learning, the AI system will contribute to improving transparency and accountability in the tendering process, ultimately safeguarding against corruption and inefficiencies.

3.5.6 Proposed System Flowchart



3.6 Association of research method to project

3.6.1 Analyze and Estimate Market Price

The statistical methods Percentage difference and Interquartile ranges were used to identify bid values that were outliers and therefore suspicious.

3.6.2 Analyze and Categorize Tenders

Large language models will be used to analyze the titles of tenders and compare them to preexisting categories to see which ones they fit into.

3.6.3 Impact of ML and Analytics on Efficiency, Fairness, and Accountability

In order to measure its impact, both qualitative and quantitative methods of research should be used. This would involve carrying out questionnaires and interviews for procurement officials, suppliers, and other stakeholders to establish qualitative data on the perceived impacts of the system. Further, it would involve a statistical evaluation of some quantitative metrics, such as a fraud reduction rate and process efficiency pre and post implementation. The mixed-method approach adopted herein helps to comprehensively examine the changes induced in both transparency and accountability on account of the system.

3.7 Research data and datasets

3.7.1 Types of Data Collected

The project will make use of both primary and secondary data. Primary data shall be drawn from the ZPPA website for identification of current procurement processes, problems encountered, and user needs. Secondary data, on the other hand, will be obtained from existing systems like PACRA, interviews, surveys, and other datasets that are publicly available.

- **Primary Data:**

Data from ZPPA focused on ascertaining the current procurement practices and challenges. As noted by Creswell (2014), the main data needed to bring out realistic facts from the users and implementers of the procurement systems.

- **Secondary Data:**

Data used will include procurement transaction data, historical bidding data, supplier profiles, contract awards, and financial records. It shall be retrieved from publicly available sources such as the e-GP system, PACRA, interviews, surveys, ZPPA itself and other public procurement audits of Zambia. As it has been evidential with Smith et al., 2017, that secondary data has proved to be a rich source in the training of machine learning models in pattern analysis and anomaly detection.

3.7.2 Data Privacy and Security

Because procurement data is very sensitive, immense attention is being paid to data privacy and security. Any personal or confidential information within the data sets may be anonymized using local data protection regulations, such as the Data Protection Act of 2021. It will include secure storage of all data and access to the data by authorized personnel only in line with government data privacy policies and ICT security guidelines.

3.8 Ethical concerns related to the research

3.8.1 Fairness and Non-Discrimination

The system inherently aims at increasing transparency in the procurement process by use of machine learning. In addition, there has been an increase in the ethical concern of algorithmic bias: any data used to train machine learning models will reflect existing or historical inequities in procurement decisions; it would, therefore, not be in the interest of the system to foster such algorithms that might further unfair treatment of suppliers or bidders. The datasets should be representative and diverse, so that the risk of biased evaluation of suppliers is at a minimum.

3.8.2 Accountability and Transparency

Even though the system seeks to enhance openness, it is equally important to ensure that there is control over the processes of making decisions developed. For instance, there have been machine learning algorithms developed to evaluate tenders and check for fraudulent practices. However, there is the possibility of decision-making being done without any clarity where the explanation for the result given by the algorithm is not readily available (Rahwan et al., 2019). This concern about the so-called 'black box' approach in order systems is even more heightened when there are publicly funded projects and contracts involved.

In order to avoid this, the quality management system will be built in such a way that all processes of making decisions will be clear, traceable and above all open for examination. In addition to that, there will be human intervention at key stages of the process in order to ensure responsibility and provide reasoning for systems-based decisions.

Ethical Considerations in Research (Enago Academy, 2023)

3.8.3 Preventing Misuse and Corruption

The system would be subject to abuse in terms of fraud and anomalies detection in procurement processes if it were not safeguarded. A chance still exists that some of the individuals who are involved in corrupt activities would try to manipulate or bypass the system. This is also a very important ethical issue, looking at the fact that public procurement is one of the areas the Zambian government largely spends its money on.
Strong security protocols backed by periodic auditing and oversight will be instituted as preventive measures against unauthorized access and manipulation. The system should also have anti-corruption principles in conformance with the Anti-Corruption Commission Act of 2012, aimed at reducing unethical practices within public procurement.

3.8.4 Informed Consent and Stakeholder Engagement

The other pivotal ethical issue will be to make sure that all stakeholders-suppliers, government agencies, and citizens-are well informed about the functioning of the system and how it is going to use their data. It must also be the case that stakeholders may give consent to use their data within the system, in obedience to the ethics guidelines on using data.

There will also be regular consultations with key stakeholders, including government officials, suppliers, and civil society, for transparency and to instill confidence in the people regarding the deployment and operation of the system.

3.9 Chapter Summary

In summary, this project uses the DSR methodology in order to develop an information technology-based solution to enhance transparency, efficiency, and accountability in public e-procurement. DSR is iterative; this provides scope for continual feedback to improve the solution.

Qualitative and quantitative approaches both shall be adopted. The Agile methodology informs system development. With respect to the technologies used, the system relies on PHP, Python, JSON, HTML, Bootstrap, and MySQL/PostgreSQL. Machine learning models apply in the forecast of market prices, analysis of bidders, and fraud detection.

The methodology involves data collection from primary sources through user feedback and secondary sources, which are procurement datasets. Key ethical considerations regarding data privacy, data equity, and data security guarantee that the system operates within the legal and ethical parameters set by the constitution.

CHAPTER 4 DATA, EXPERIMENTS, AND IMPLEMENTATION

4.1 Appropriate modelling in relation to project

4.1.2 Waterfall Model

I chose the Waterfall design methodology for my system because it is structured and successive, meaning that it moves from one phase to another only when the first is completed. This methodology presents clarity and a reduction of ambiguity in the well-set requirements, design of the system architecture, and an orderly implementation of the features. The scope of changes within a well-structured waterfall model is very minimal during its development, hence better control and predictability.

The Waterfall methodology is quite suitable for this project because the requirements are well understood and stable. Its emphasis on thorough documentation ensures that all processes, decisions, and designs are recorded for easy reference in the future, as well as audit purposes.

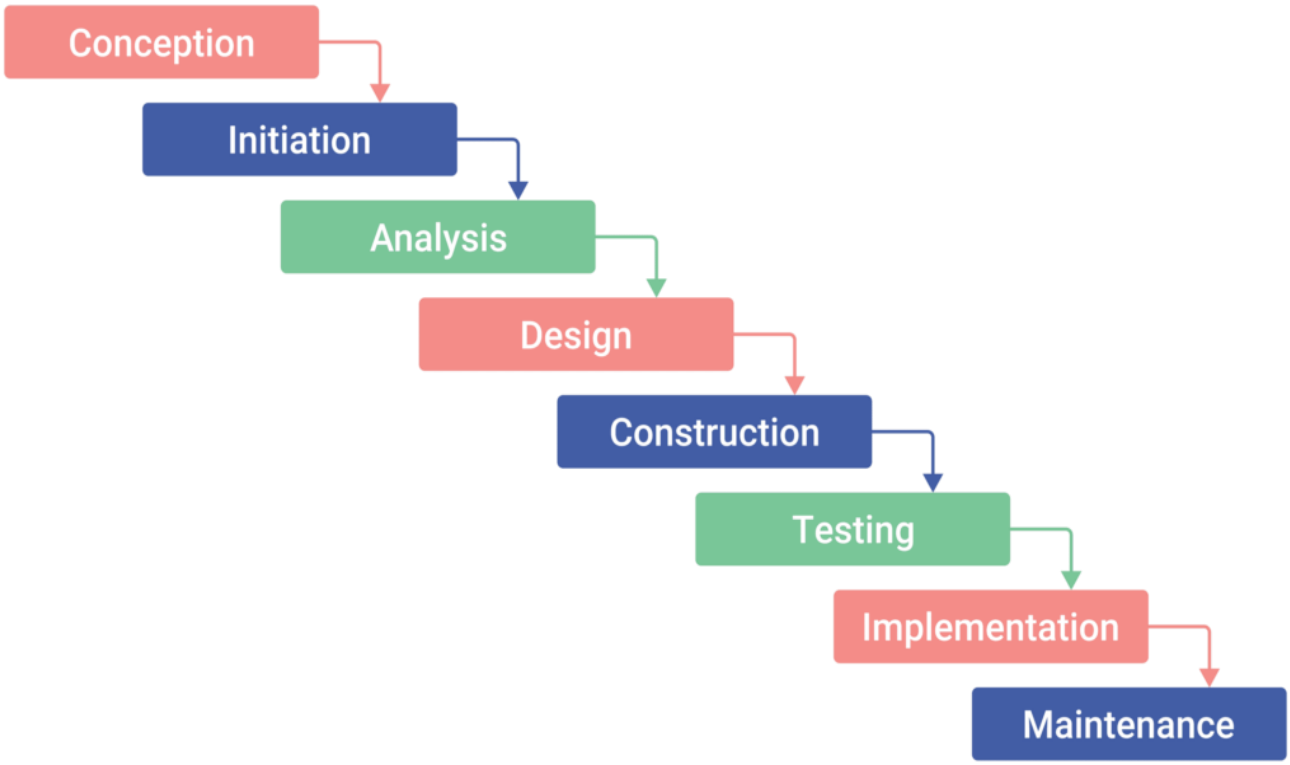


Figure 4.1.2 *Waterfall Model - Software Planet Group 2024*

4.2 Techniques, algorithms, mechanisms

This section explains the main methods, strategies, and processes designed to build the system. Each of the components addresses the specific task that is necessary for analyzing, storing, and displaying procurement information, especially for the implementation of machine learning and analytics for tender monitoring and evaluation in real time.

4.2.1 Large Language Models for Tender Classification

Generative Pre-trained Transformer (GPT) models, which are also classified as Large Language Models (LLMs), possess advanced features that enhance their usability in the Natural Language Processing Arena. It is possible to use these models to ease the processing and the analysis of very large amounts of textual data such as procurement textbooks. In this project, LLMs will be used to work with tender documents in order to extract information needed for their classification. This classification aims at organizing tenders according to their types, purposes, estimating values, and other relevant parameters. Due to the model's ability of discerning patterns and understanding their context, tenders will be further interrogated for red flags such as anomalies or signs of frauds in utilization which will allow for risk predictive analytics. As a result, text analysis is one of the major activities of this project, and the data insights will assist all the relevant stakeholders in making their decisions.

4.2.2 Web Applications for Data Display

To facilitate the visual representation of the analyzed data, a web-based application interface will be created. Therefore, stakeholders will have an easy platform to access and make sense of the procurement data. The web application aims to present suitable tender data alongside its analytics and notification of any concerning anomalies that may arise. Frontend application design using technologies like HTML, Bootstrap and JavaScript among others, will provide an interactive and responsive design that will allow easy access to the web application on any type of device. This interface will be important in promoting transparency and simplicity in accessing procurement data for both the technical and non-technical users.

4.2.3 Web Scraping for Data Collection

The project shall apply web scraping in the collection of relevant procurement data from targeted sources, such as official procurement portals. Python-based web scraping scripts for the automation of data retrieval are meant to ensure consistency and efficiency in the collections. It will be capturing information relating to newly published tenders, historical procurement data, and related documents necessary for machine learning analysis. The extracted data from web scraping will be formatted to be able to be ingested by the database and add both breadth and depth of data available for analytics.

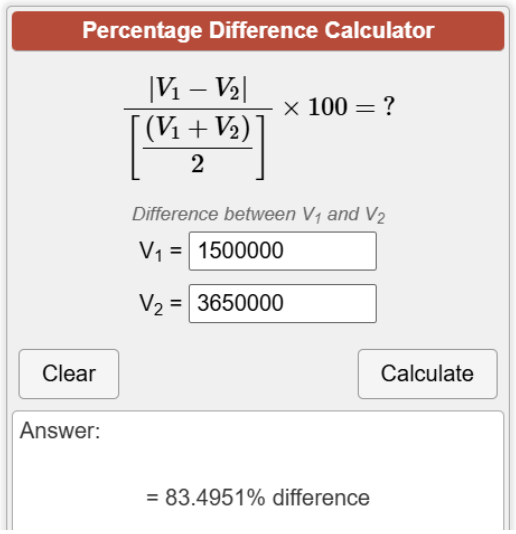
4.2.4 Database Management for Data Storage

This system is designed to securely store large volumes of procurement data and permit access through active database management. The RDBMS, which could be MySQL or PostgreSQL, would be used for structured and scalable storage to manage raw data and processed outputs. This structure effectively allows for the querying of data, real-time and historical data integration-elementary components that provide predictive insights and analyze procurement trends.

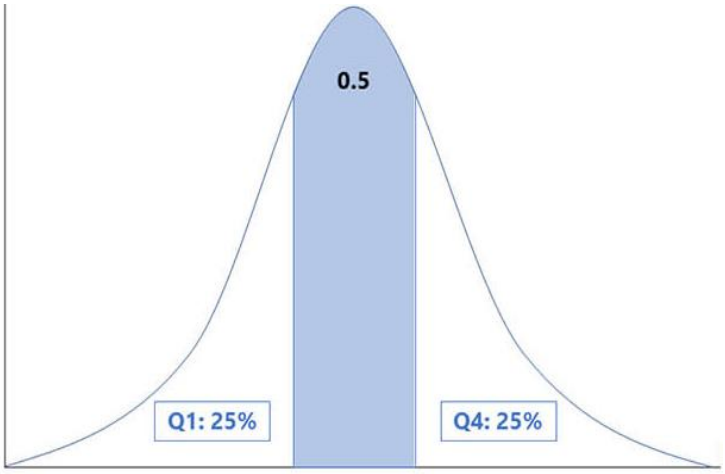
4.2.5 Statistics for corruption risk analysis

In order to ascertain fairness and detect irregularities in our system, it is imperative to analyze bid data very carefully. A comparison of bid prices is one such method for determining significant deviations, an indicator of a possible anomaly or risk of corruption. Analytical mechanisms will be different, contingent on whether a single or multiple tenders were received for a specific calling, as the number of bids significantly inhibits the robustness of insights gleaned from these analyses.

When a tender receives only two bids, the **Percentage Difference method** is employed in order to assess the relative variation of the prices. This method has a simple and effective means of disclosing substantive differences in the bid amounts. In the case where a bid is either much higher or much lower than the other, it will bring about further investigation into the causes of the difference. This approach is simple and provides transparency in the cases of low competition but does not complicate analysis in such a situation.



Percentage Difference Calculator



Interquartile Range Calculations & Graphs

When a number of bids are received, sophisticated statistical methods like **Interquartile Range (IQR)** are used to detect outliers in bids as opposed to other methods. The inter-quartile method measures the distance to the middle 50 % of the bid prices so that one can find whether any bids are highly deviating from the average. The method can expose highly divergent bids either above the contract price or below the contract price, which is considered one of the malpractices in the procurement process. Addressing such anomalies creates a much fairer bidding environment and fosters trust in public procurement processes. The logic underlying application of those methods creates a more rigorous data-driven approach for bid evaluation, which relates directly to the broader agenda of transparency and accountability in procurement.

4.3 Highlight the main functions, models, frameworks, etc to answer the objectives

- **Web scraping**

1. *Procuring entities data*

The process began with web scraping to download a list of all procuring entities from the ZPPA website: ZPPA Procuring Entities.
https://eprocure.zppa.org.zm/epps/app/viewPublication.do?selectedItem=app/viewPublication.do

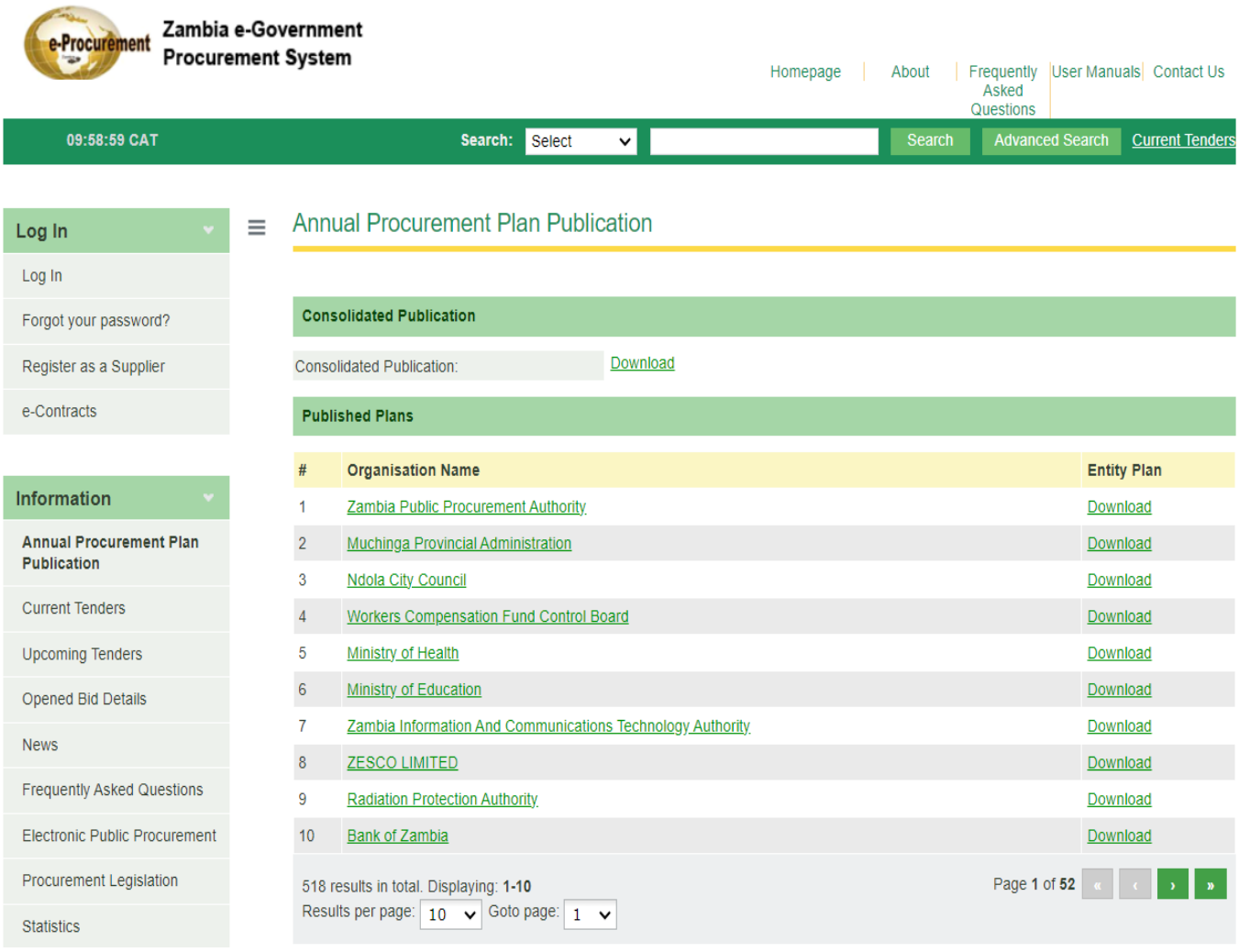


Figure 4.3a *ZPPA website (Annual Procurement Plan, 2024)*

Once the Excel file was downloaded, Python code was executed to connect the Excel document to the database, named Procurement. The script was then run, and the procuring entities in the Excel file from ZPPA were exported directly into the database.

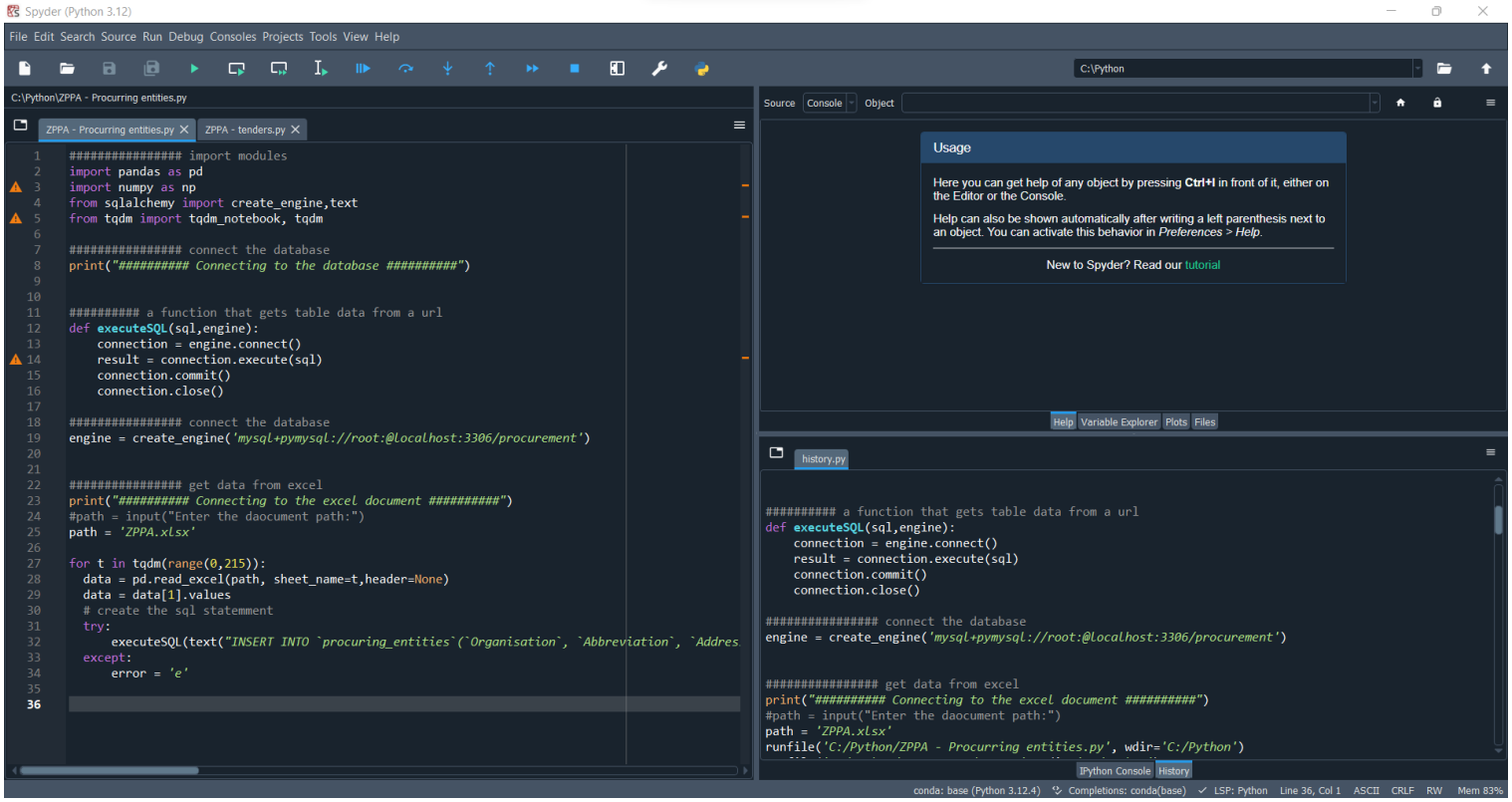


Figure 4.3b *Python execution - connecting the excel document to the database (Spyder v5)*

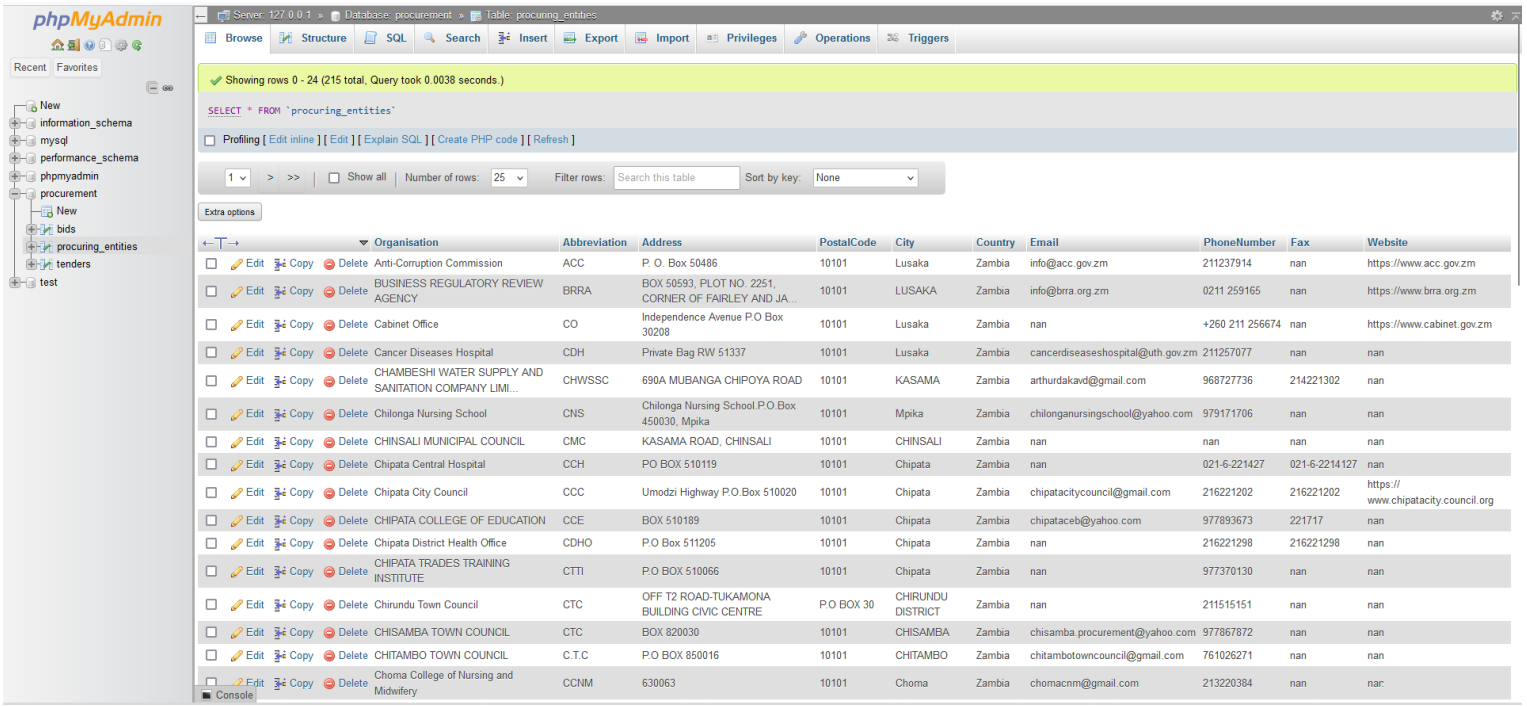


Figure 4.3c *Procuring entities exported directly into the database (MySQL)*

2. *Acquiring tenders and bids*

Alongside this critical function in the process comes the entering of all tender data into the database, using a set of pre-defined steps. This is achieved through importing the necessary modules and loading specific functions that automate the tender management process.

The following workflow describes the process for:

- **SQL Statements Run**: A function executes, running the appropriate SQL statement necessary to prepare the database for new entries.
- **Extract Table Data from URL**: This is another function that retrieves table data directly from the given URL, in this case it's the ZPPA website URL, to automatically gather tender information.
- **Extract List DataTable (dt) from URL**: Now, it uses a special function to get the list of DataTables from that URL, ensuring all the data is collected in an organized manner.
- **Database Connection**: fetching tenders and initiating the processing therein. In doing so, it searches the tender entered and saved in the respective columns in the database.
- **Update Tender Details:** If a tender does exist, the function will ensure that information concerning the same is updated.
- **Bid Opening Date and Bid Saving:** The function fetches the bid opening date and saves the bid accordingly, so that not a single critical detail related to the tender is missed.
- **Final Tender Update:** The information of the tender is updated in the last stage with newly entered data for the same so that the database receives the most updated and complete information.
- **Tender categorisation with LLMs**
- **Corruption risk analysis**

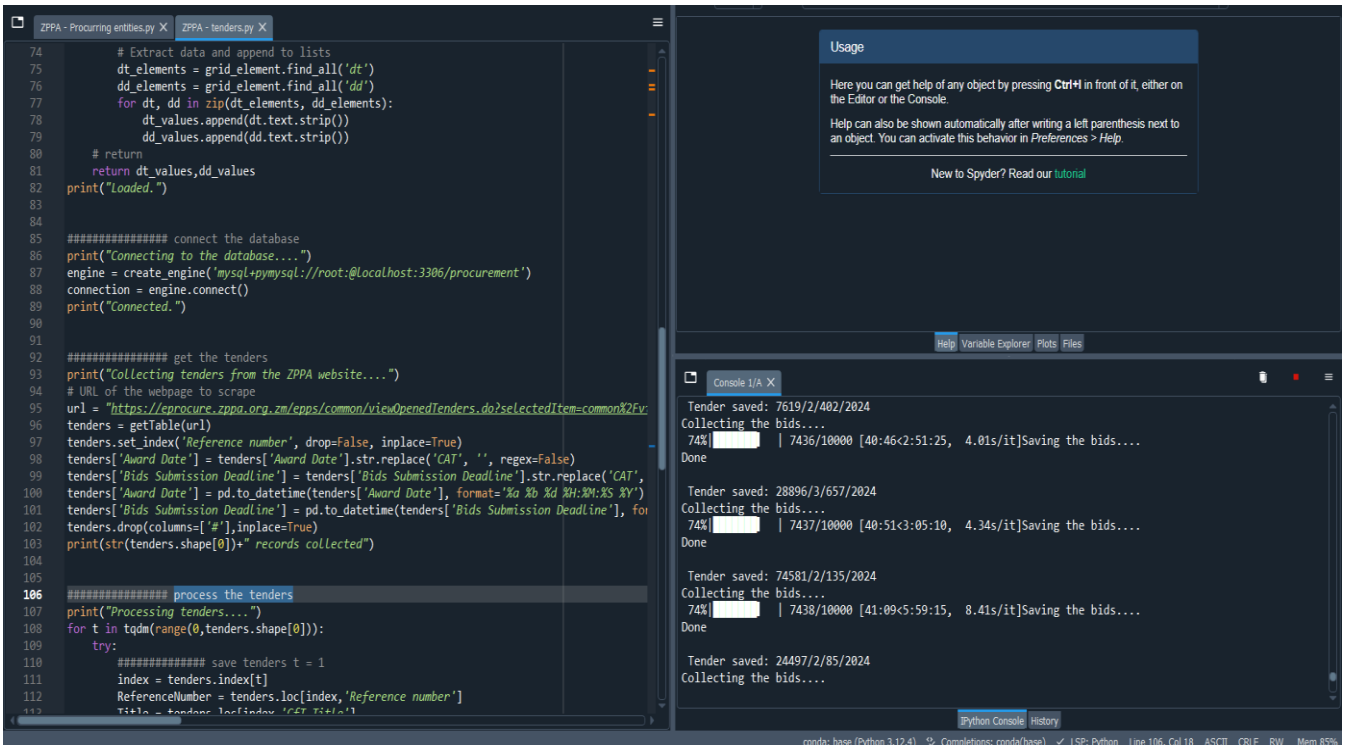


Figure 4.3d *Entering of all tender data into the database (Anaconda, Spyder v5)*

The tender information processing is designed to be fully automated for efficiency and accuracy, enhancing the management of procurement information. This process should be done periodically since ZPPA normally updates tender information quite often. The process should be regularly run to ensure that correct and most current tender data has been captured.

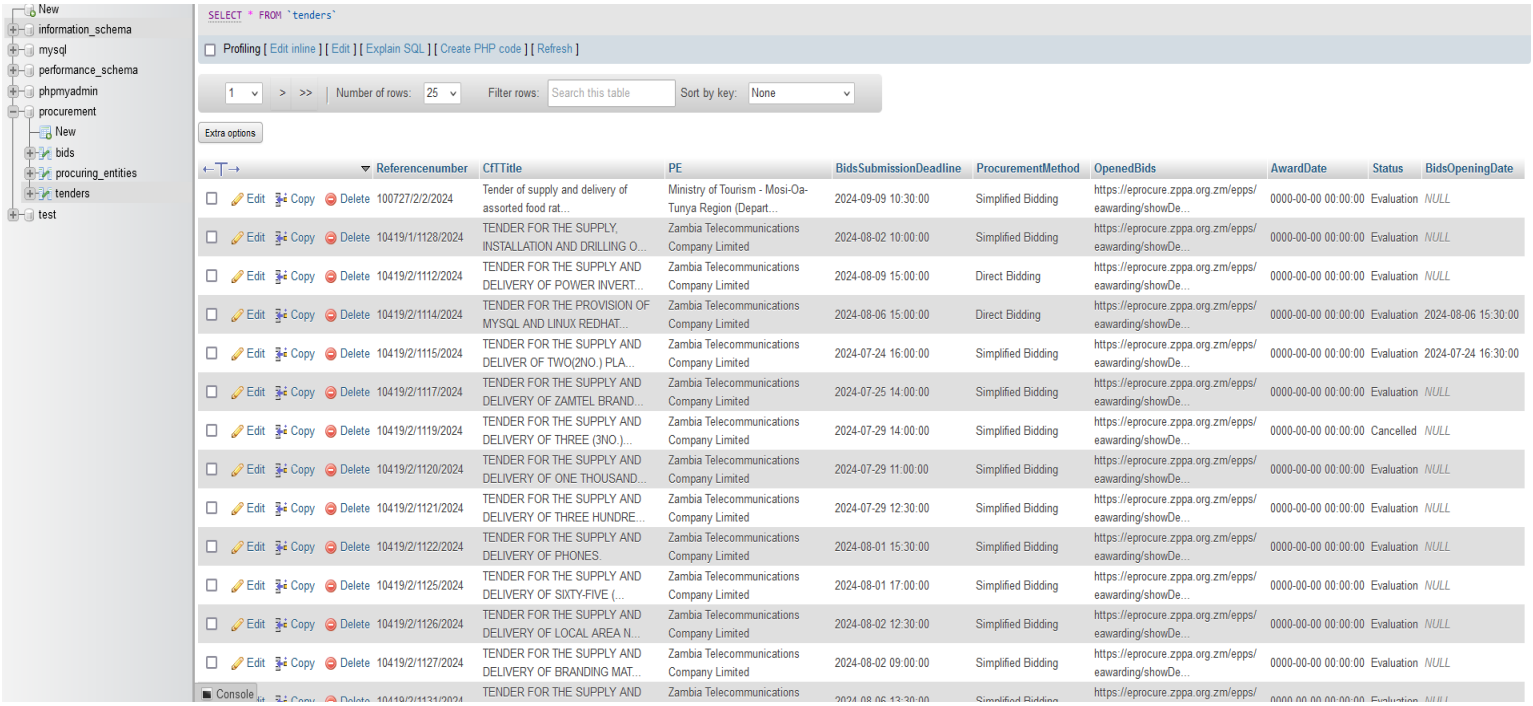
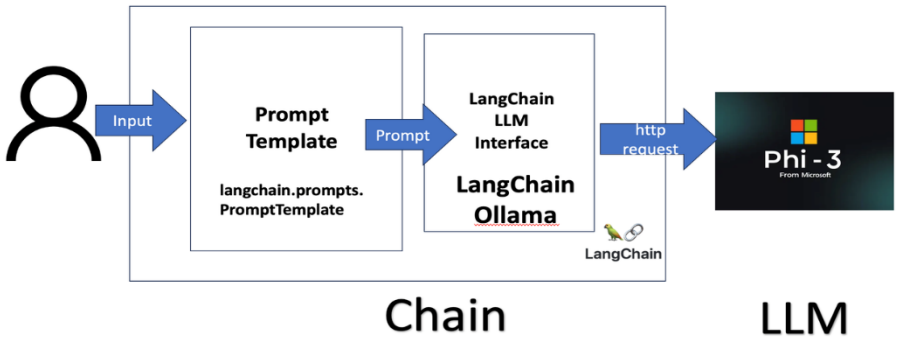


Figure 4.3e shows successful database entry from an automated tender information process, designed for periodic updates to capture the latest ZPPA data accurately and efficiently.

4.4 Ollama

A framework for a Large Language Model for natural language processing and generation; in our system, it will be used, among other things, for analysing procurement data, tender and bid classification as well as to generate evidence needed to strengthen decision-making processes. Above all, its advanced language processing capabilities will provide transparency, efficiency, and accountability for public e-Procurement.



CHAPTER 5 RESULTS AND DISCUSSIONS

5.1 Results Presentation

- Importing Tenders and Bids

```
97  print("Collecting tenders from the ZPPA website....")
98  # URL of the webpage to scrape
99  url = "https://eprocure.zppa.org.zm/epps/common/viewOpenedTenders.do?selectedItem=common%2FviewOpenedTenders.do&T01_ps=100"
100 tenders = getTable(url)
101 tenders.set_index('Reference number', drop=False, inplace=True)
102 tenders['Award Date'] = tenders['Award Date'].str.replace('CAT', '', regex=False)
103 tenders['Bids Submission Deadline'] = tenders['Bids Submission Deadline'].str.replace('CAT', '', regex=False)
104 tenders['Award Date'] = pd.to_datetime(tenders['Award Date'], format='%a %b %d %H:%M:%S %Y')
105 tenders['Bids Submission Deadline'] = pd.to_datetime(tenders['Bids Submission Deadline'], format='%a %b %d %H:%M:%S %Y')
106 tenders.drop(columns=['#'],inplace=True)
107 print(str(tenders.shape[0])+" records collected")
108
109
110 ################ process the tenders
111 print("Processing tenders....")
112 for t in tqdm(range(700,tenders.shape[0])):
113     ############## save tenders
114     index = tenders.index[t]
115     ReferenceNumber = tenders['Reference number'].values[t]
116     Title = tenders['CfT Title'].values[t]
117     ProcuringEntity = tenders['PE'].values[t]
```

Figure 5.1a *Importing Tenders and Bids into our database from ZPPA website*

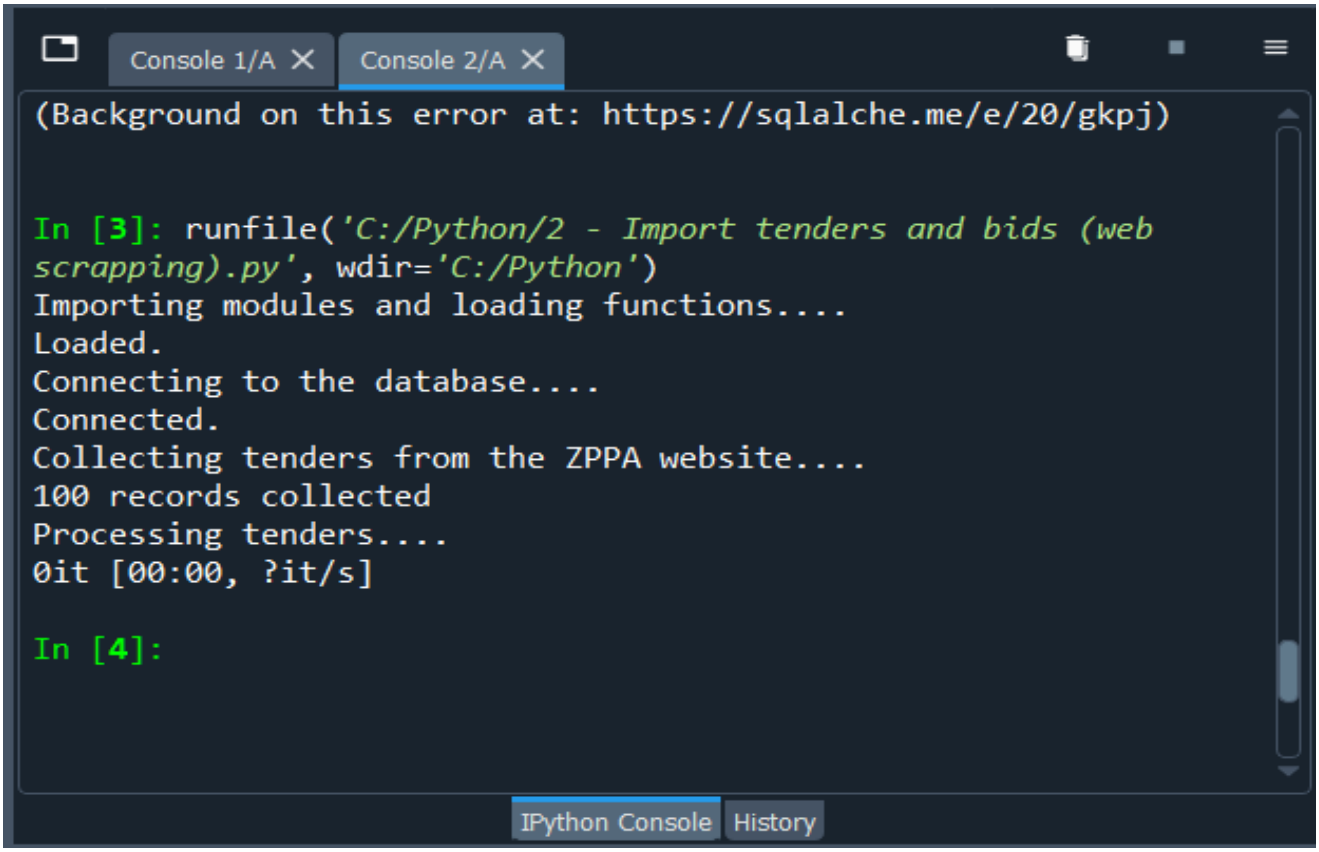


Figure 5.1b *Shows successful results of bids and tenders collected*

- Categorization of Tenders

```
engine = create_engine("mysql+pymysql://root:@localhost:3306/procurement")


####### Retrieve data from the database
categories = pd.read_sql("SELECT * FROM `procure_categories`", engine)
tenders = pd.read_sql("SELECT * FROM `procure_tenders` WHERE `Referencenumber` NOT IN (SELECT DISTINCT `tender` FROM `procure_tende


####### process the tenders t=0
for t in tqdm(range(0,tenders.shape[0])):
    Referencenumber = tenders['Referencenumber'].values[t]
    CfTTitle = tenders['CfTTitle'].values[t]
    # check which category it falls under (loop through all the categories)
    for i in range(0,categories.shape[0]):
        category = categories['Category'].values[i]
        ## create the question
        question = "The following are the details of a tender that was advertised: "+CfTTitle+". Does it fall into the category: "+
        ## get an anwer from ollama
        answer = questionOllama(question)
        answer = answer.lower()
        # Check if 'yes' appears in the text
        if 'yes' in answer:
            from sqlalchemy import create_engine
            sql = "INSERT INTO `procure_tendercategories`(`Tender`, `Category`) VALUES ('"+Referencenumber+"','"+category+"')"
            # save it to the database
            engine = create_engine("mysql+pymysql://root:@localhost:3306/procurement")
            with engine.connect() as connection:
                result = connection.execute(text(sql))
                connection.commit()
                connection.close()
```

Figure 5.1c *Bid Categorization*

The code above seeks to automate categorizing tenders based on textual analysis via the decision-making from the chatbot and would go a long way toward easing manual categorization. Yet, its use of dynamic SQL in INSERT operations and repeated database connections may not be optimized for scalability or security.

```
In [3]: runfile('C:/Python/2 - Import tenders and bids (web
scrapping).py', wdir='C:/Python')
Importing modules and loading functions....
Loaded.
Connecting to the database....
Connected.
Collecting tenders from the ZPPA website....
100 records collected
Processing tenders....
0it [00:00, ?it/s]

In [4]: runfile('C:/Python/3 - Categorise tenders.py', wdir='C:/
Python')
 21%|          | 54/255 [3:34:05<46:47:51, 838.17s/it]
```

Figure 5.1d *Showing Progress of Categorization*

- Categorization Results

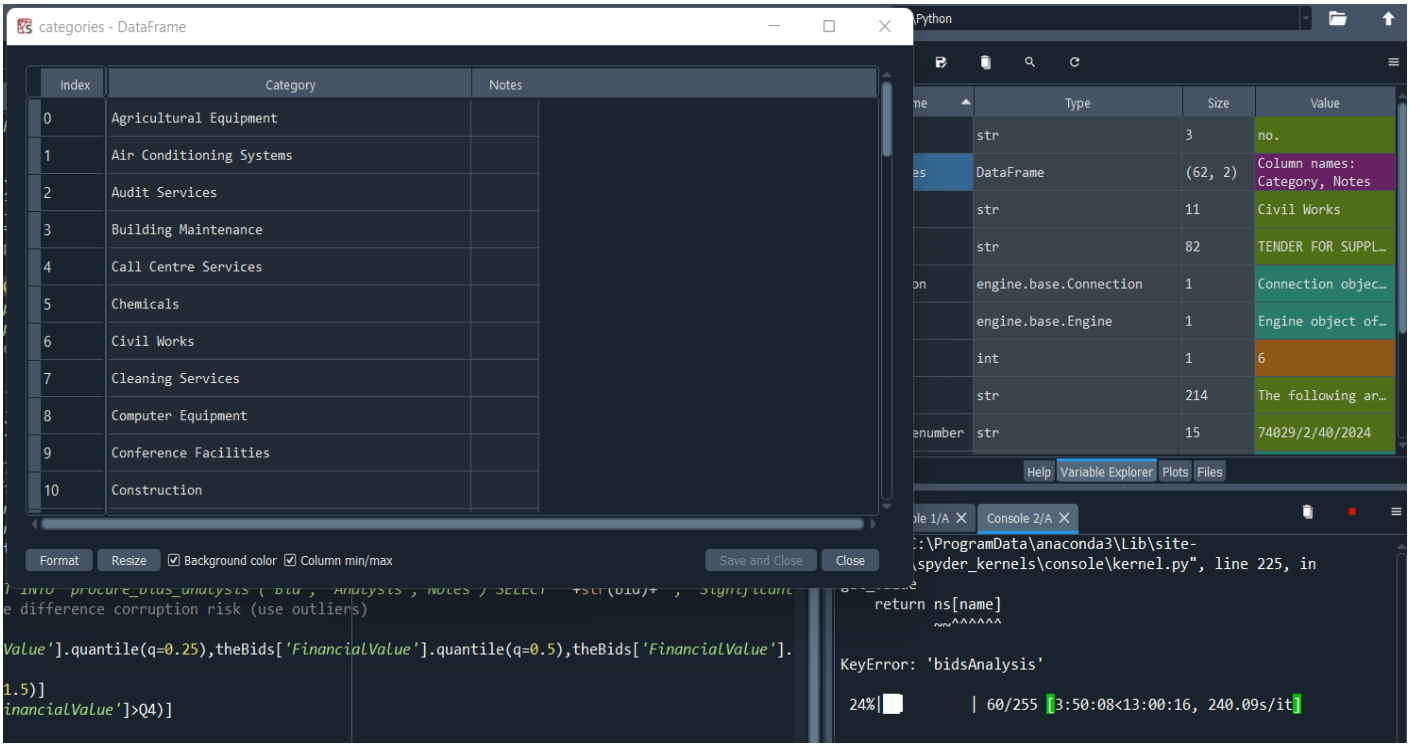


Figure 5.1e *Successful Categorization*

- Corruption Risk Results

Index	BidID	Tender	Supplier	FinancialValue	ecurityProv	lSecurityVal	FDR
0	120882	85752/2/28/2024	Simon Mwanza General Dealers	390.5	N/A	0	0.0
1	120909	20820/3/482/2024	New Horizon Printing Press	27561.6	N/A	0	0.0
2	120910	85692/2/22/2024	MACHIMAX GENERAL DEALERS	2118	N/A	0	0.0
3	120948	28896/3/632/2024	Krishna Travels Limited	65800	N/A	0	0.0
4	120957	66307/2/85/2024	Kom Trade	6690	N/A	0	0.0
5	120960	20820/3/482/2024	COMBINED PRINT MEDIA AND ADVERTISING LIMITED	67500	N/A	0	0.0
6	120964	81376/2/144/2024	Claving Enterprises Limited	24050	N/A	0	0.0
7	120968	39303/2/205/2024	EXTRIM PROJECTS ZAMBIA LIMITED	2900	N/A	0	0.0
8	120974	57250/3/193/2024	Protea Hotel by Marriott Lusaka	19750	N/A	0	0.0
9	120975	20820/2/483/2024	HANDYMAN'S PARADISE LTD	43154.1	N/A	0	0.0
10	120992	25204/2/1088/2024	PRENAM GENERAL DEALERS	0	N/A	0	0.0
11	120993	20759/2/59/2024	CLOTHES CLUB	46425.6	N/A	0	0.0
12	120998	44583/3/180/2024	Zambia Daily Mail	16837.6	N/A	0	0.0
13	121003	75975/3/638/2024	ZAMTEL	2298	N/A	0	0.0
14	121006	27007/3/216/2024	Relaxin Executive lodge	21425	N/A	0	0.0
15	121007	81376/2/142/2024	PATENDE INVESTMENTS LIMITED	7500	N/A	0	0.0
16	121013	74581/2/126/2024	ELIWAMA ENTERPRISES LIMITED	2.3064e+06	N/A	0	0.0
17	121015	81376/2/141/2024	Dosig Genaral Dealers	7580	N/A	0	0.0
18	121027	74581/2/125/2024	ELIWAMA ENTERPRISES LIMITED	832600	N/A	0	0.0
19	121032	5907/2/1539/2024	TOP ENERGY ZAMBIA LIMITED	990000	N/A	0	0.0
20	121046	44583/2/179/2024	ZAMTEL	45000	N/A	0	0.0
21	121050	67685/2/11/2024	THE OFFICE STORE LIMITED	18828	N/A	0	0.0
22	121054	39303/2/205/2024	Mays Logistics Limited	4000	N/A	0	0.0

Index	id	Bid	Analysis	Notes
0	1	147878	Single Bidder	
1	2	147880	Single Bidder	
2	3	147428	Single Bidder	
3	4	147584	Single Bidder	
4	5	148938	Single Bidder	
5	6	122403	Single Bidder	
6	7	123722	Single Bidder	
7	8	124171	Single Bidder	
8	9	124890	Single Bidder	
9	10	125705	Single Bidder	
10	11	126444	Significant Price Difference (Percentage Difference)	Significantly Lower Price

Figure 5.1f *Corruption Risk Results*

- List of Procuring Entities

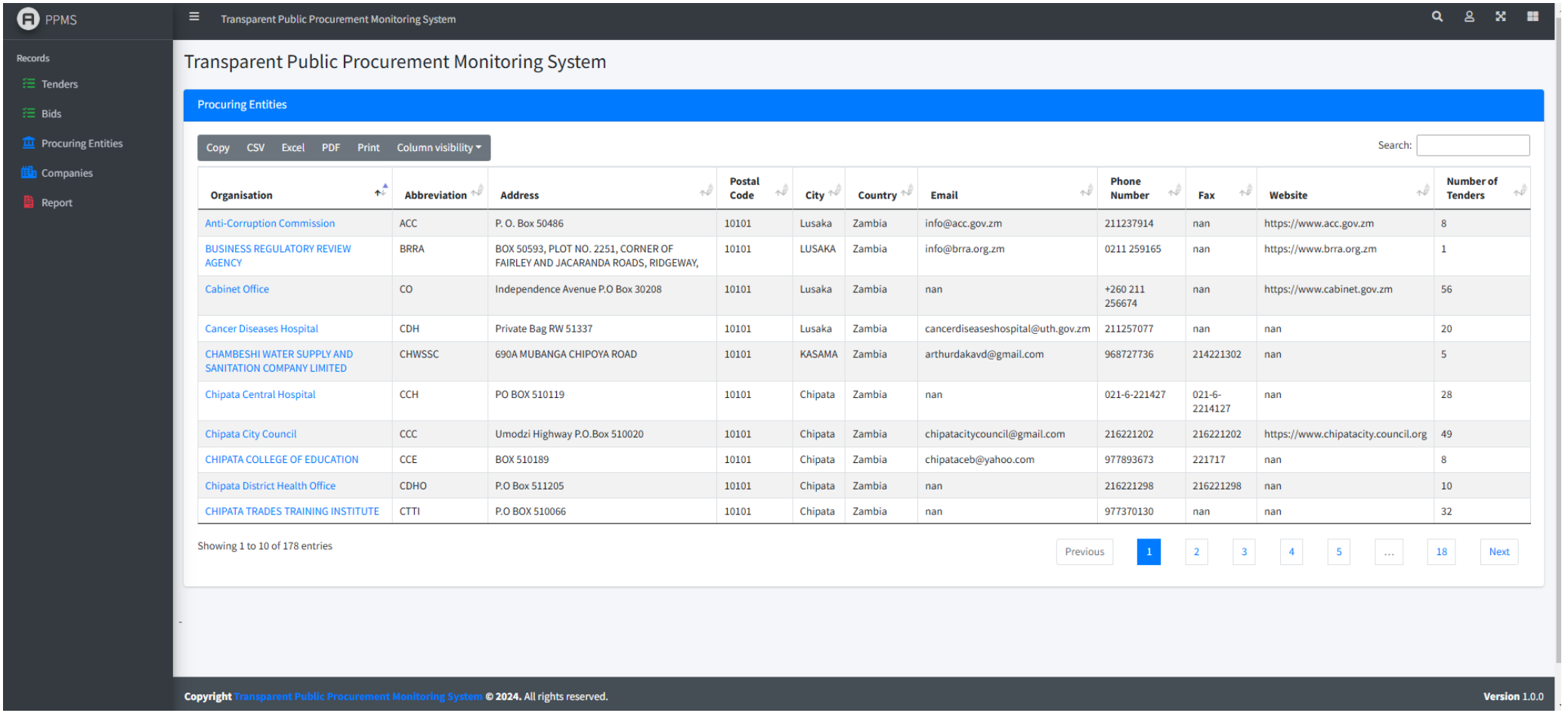


Figure 5.1g *List of Procuring Entities*

The image above, extracted from the system, displays a list of all procuring entities retrieved from ZPPA, seamlessly connected to our procurement database and presented within our system interface.

- List of Tenders

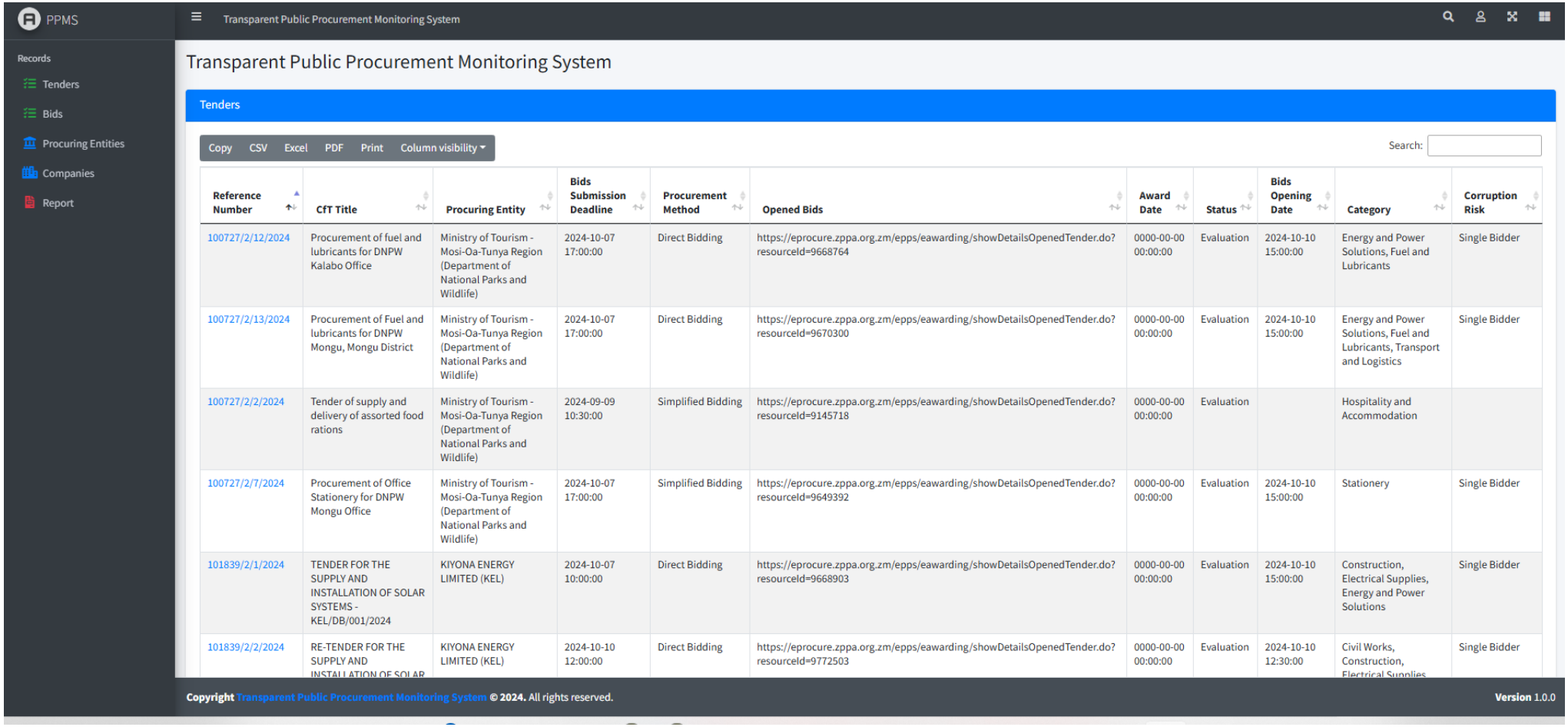


Figure 5.1h *List of Tenders*

The above picture shows a list of all Floated Tenders imported from ZPPA into our database and presented on our system, along with their associated Corruption Risk.

- List of Bids

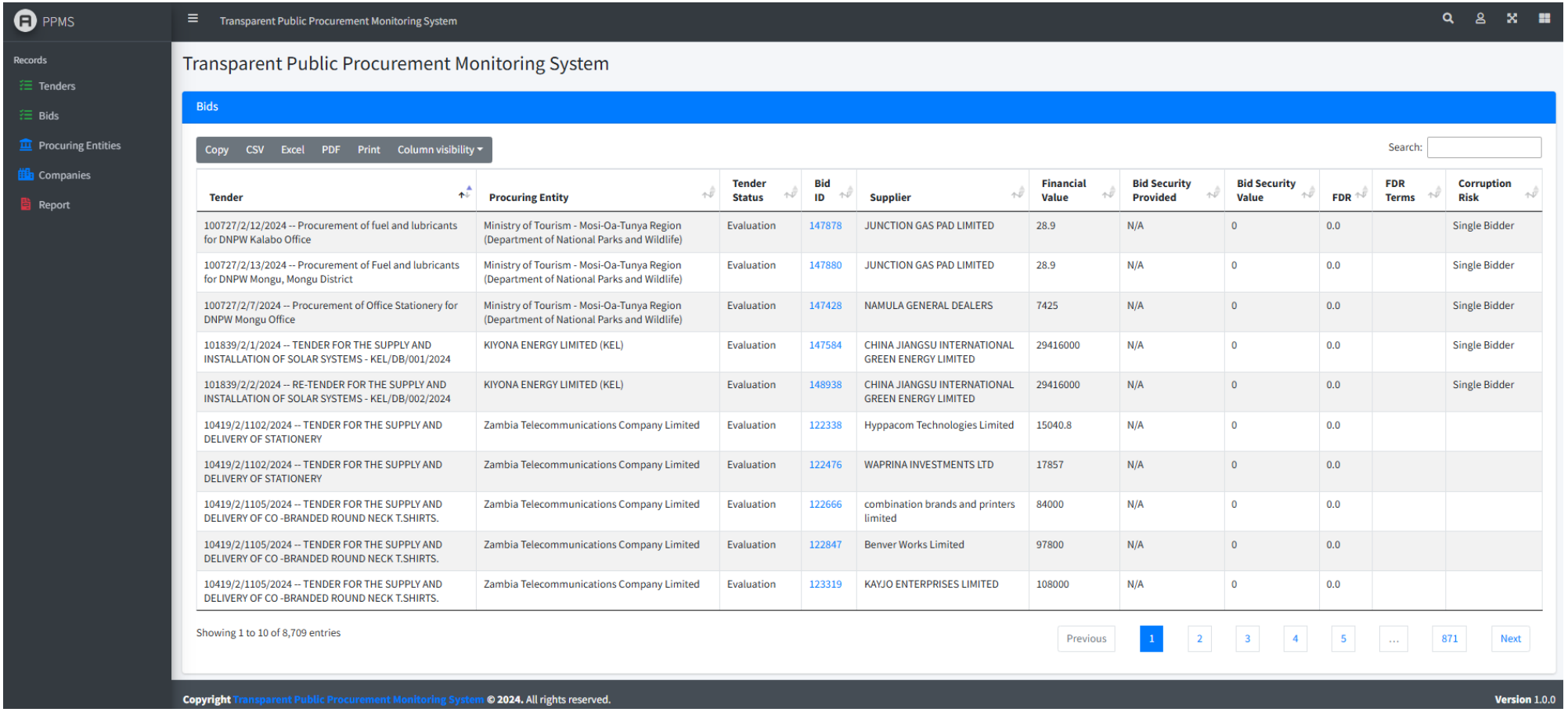


Figure 5.1i *List of Tenders*

The image above displays all the bids in our system, along with their associated risk components.

- List of Companies

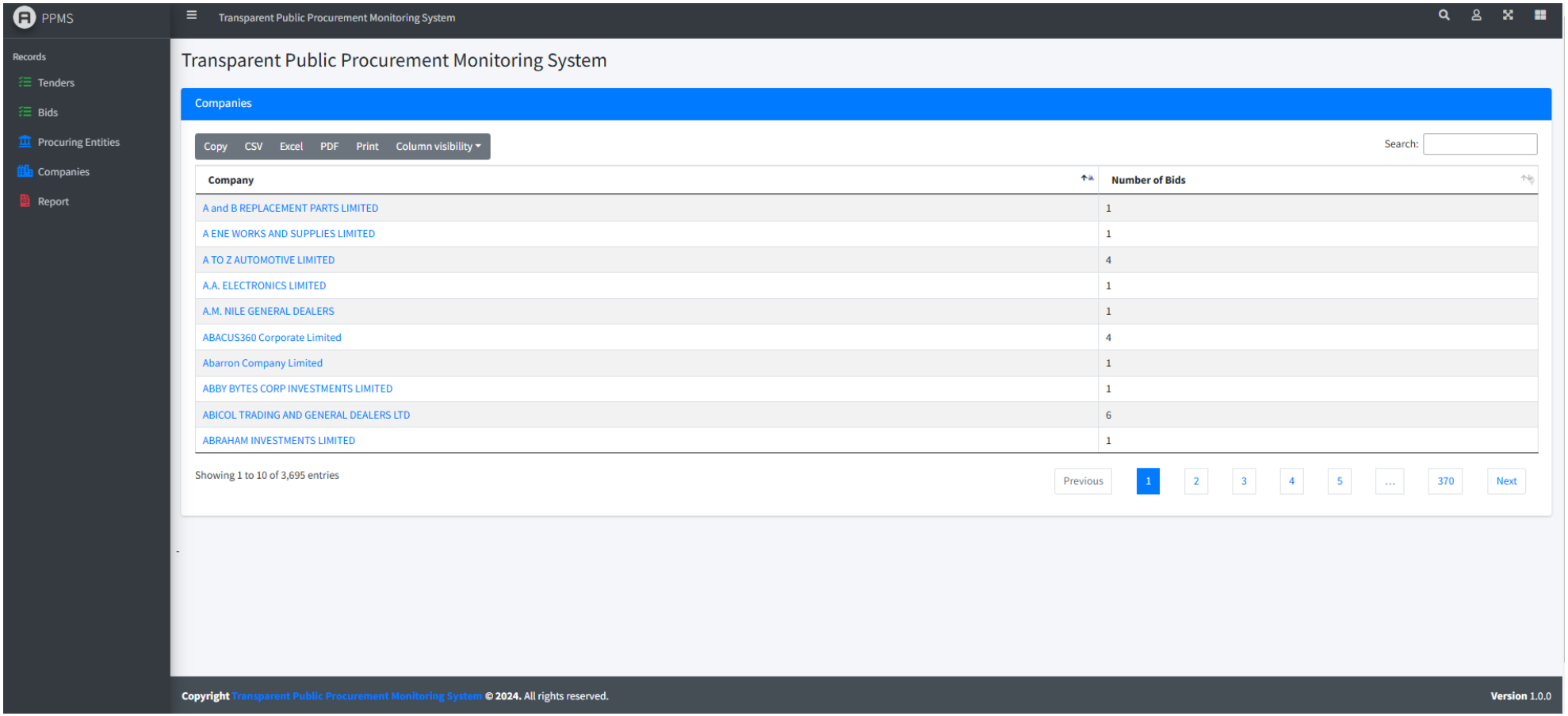


Figure 5.1j *List of Companies*

Above is a list of all companies and the number of bids each company has participated in.

- Procuring entity and its Tenders

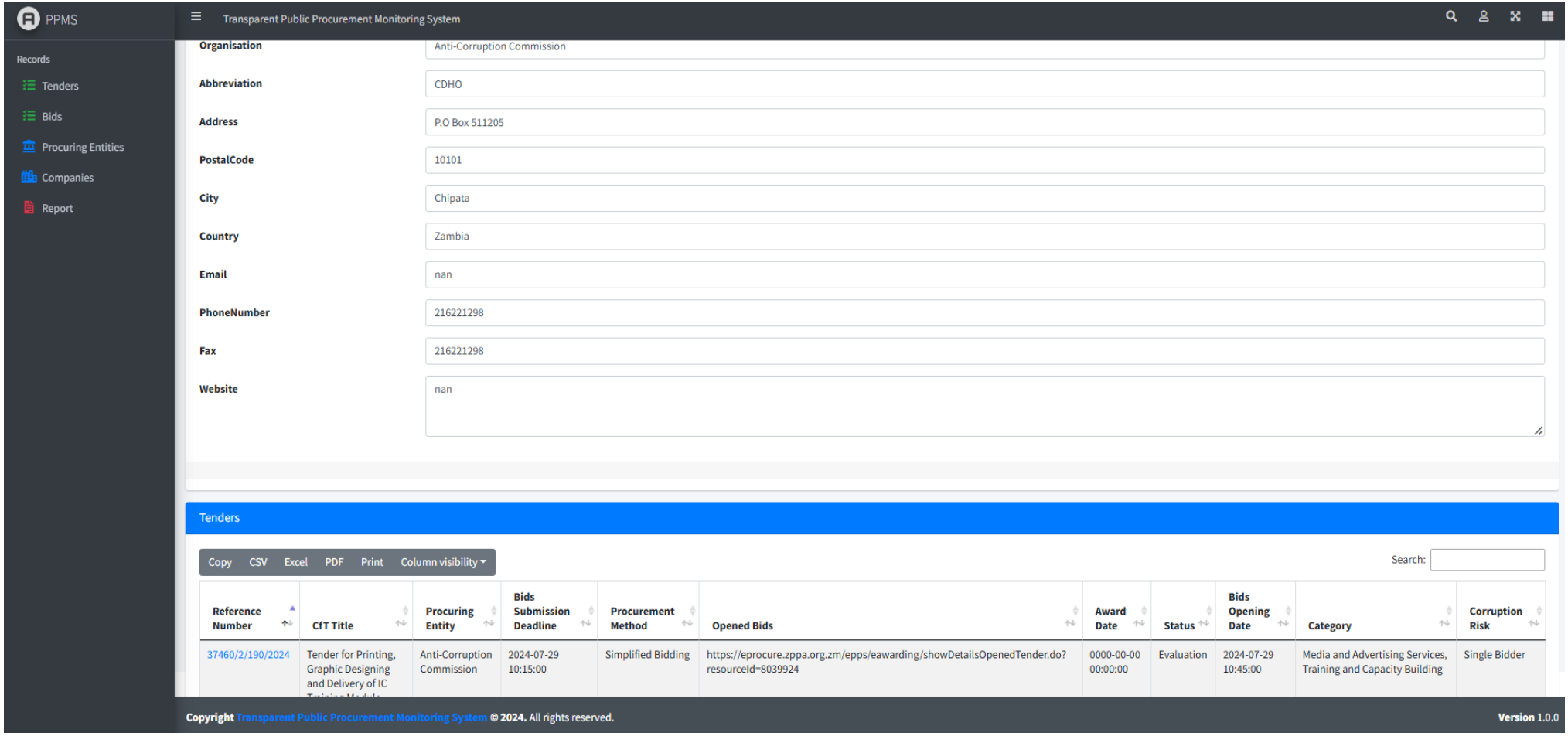


Figure 5.1k *Procuring Entity and Tenders its participated in*

The above picture shows a Procuring Entity and the tenders in which it has participated since its inception.

- Tender and its Bids

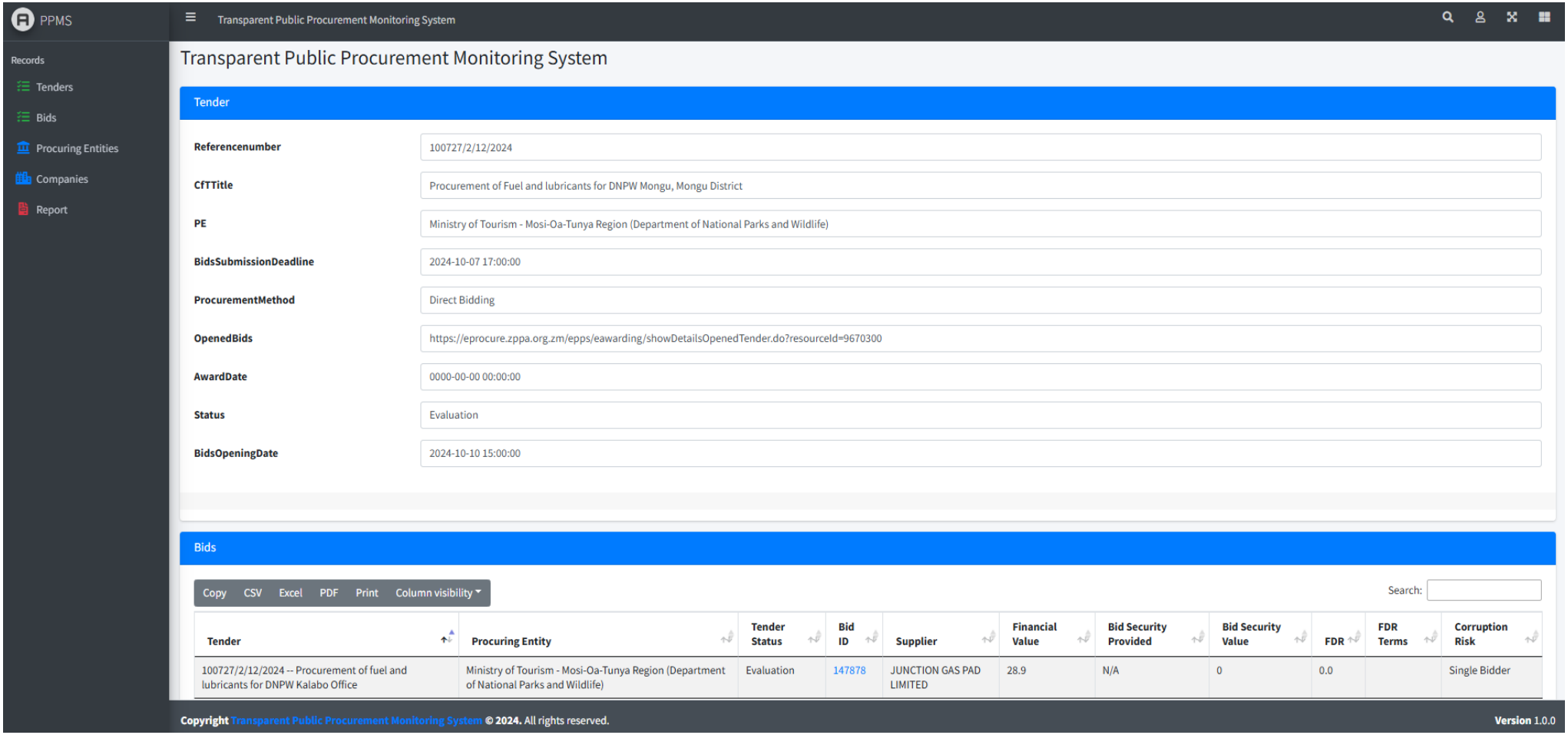


Figure 5.1l *Tender and Bidders.*

The above picture shows one tender with a list of bidders underneath it, showing also the corruption risk by the side.

5.2 Analysis of Results

Well, I successfully collected data from **178** Procuring Entities, which served as a strong basis for the procurement activity analysis. This system also aggregated data on **5,356** Tenders and **8,709** Bids that could be reviewed comprehensively regarding the trends in the bidding process.

The preliminary analysis indicates deep concern areas around transparency and corruption in public procurement as most Zambian companies do not use the ZPPA e-GP system. For instance, about **35.68%** of the tenders ended with a single bid, which points toward a lack of competition. Another fact is that **6.05%** of the tenders, which were available to two bids only, also showed much higher differences in prices between the bids, which indicates there might be some irregularities. Furthermore, **3.58%** of tenders that could hire three or more bids included bids with price outliers deviating clearly from what would be expected.

This brings to fore the contended need for a machine-learning and analytics-enabled approach to address these very issues of transparency, improve fairness, and accountability in public procurement.

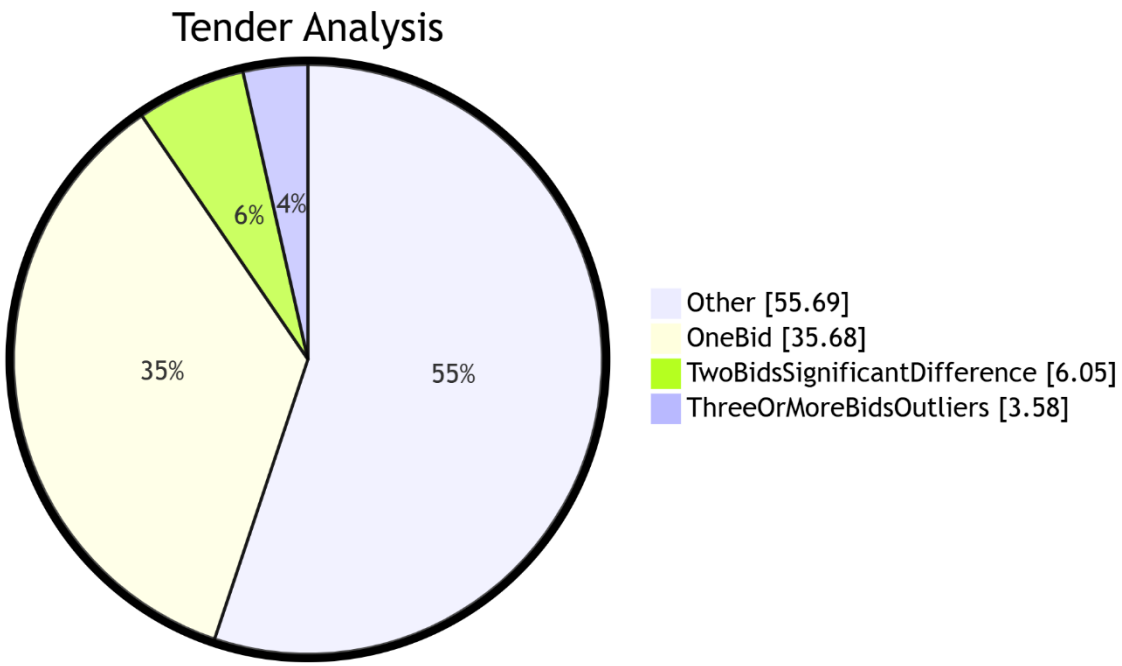


Figure 5.2a *Tender Analysis Percentages*

- Tender and Bid Corruption Analysis Percentages

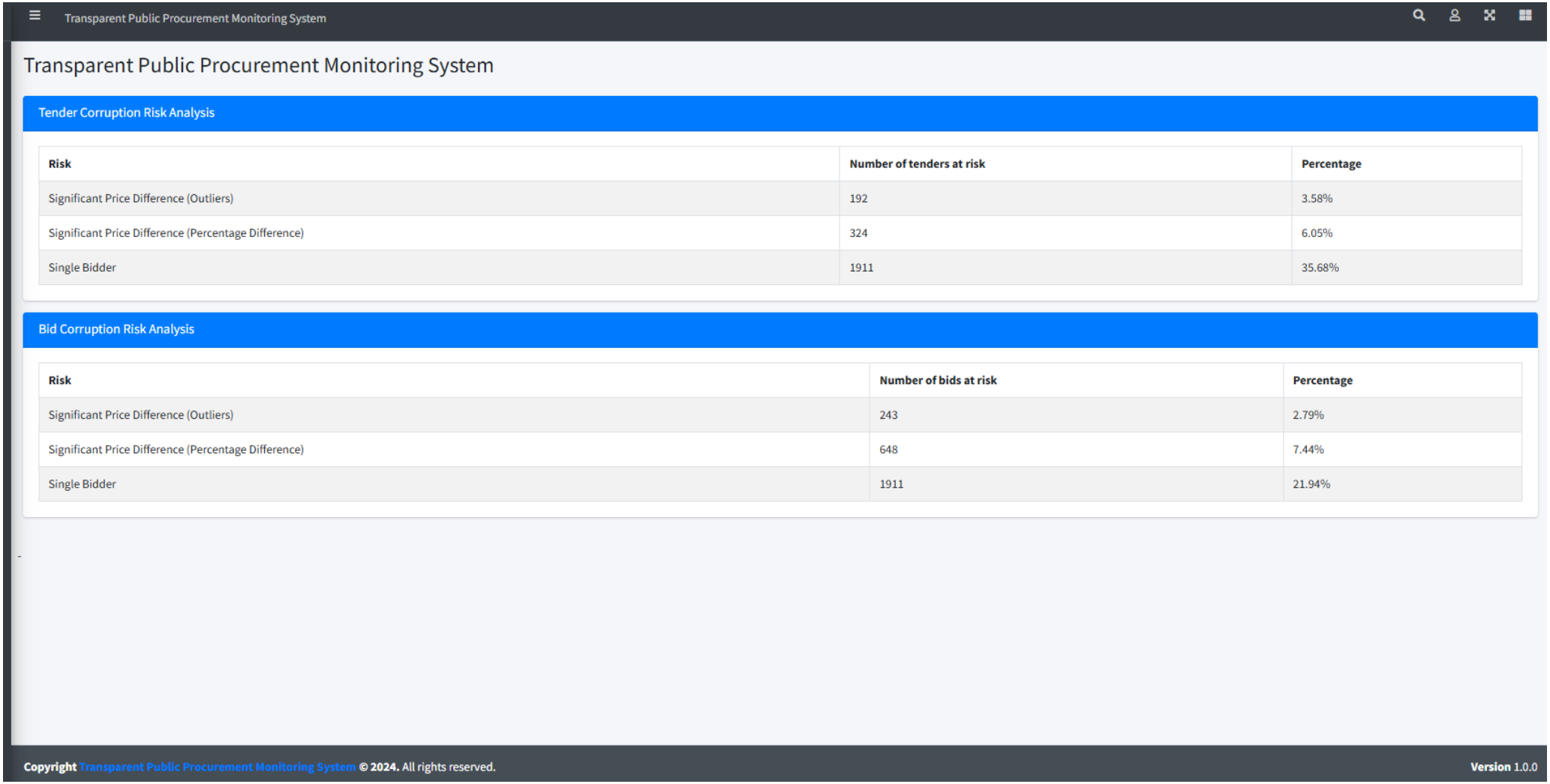


Figure 5.2b *Tender and Bid Corruption Analysis Percentages*

5.3 Comparison to Related Work

There are a lot of articles and reports on government procurement systems and their associated challenges but, unfortunately, there has not been a sufficiently solid, data-driven analysis of public procurement processes in Zambia. Of course, much literature revolves around policy recommendations and regulatory frameworks or anecdotal evidence. Very few studies take the plunge and use data in exposing inefficiencies or risk of corruption lurking within all those processes. My system has therefore gone an entirely different way by probing deep into procurement data to obtain patterns, discrepancies, and risks overlooked by conventional methods. It analyzes thousands of tenders and bids and delivers valuable hints of what measures need to be put in place to close transparency gaps and irregularities that compromise the fairness and accountability of public procurement processes.

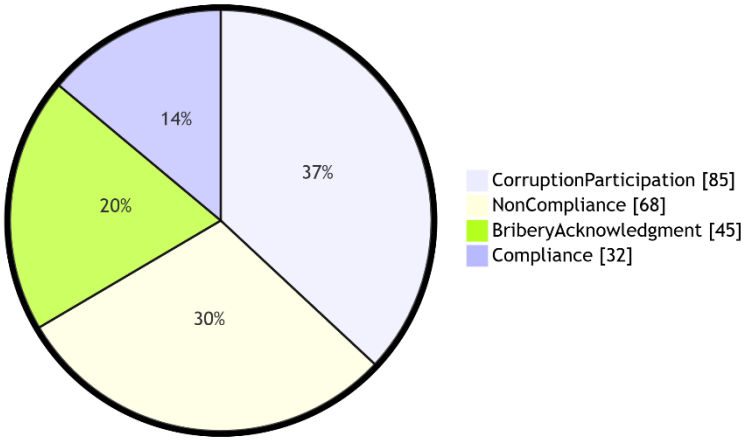
What is new in this work is the focus on data as the main resource toward improved public procurement. An approach different from all those that typically begin by making generalizations is that which my system employs to identify risks of corruption by means of indicators such as single bidders in tenders or significant price movements compared with a controlled average. This quality of scrutiny has not been brought on Zambia before, which makes my approach groundbreaking and ideal for entering improvements of transparency and fairness in the procurement space. It demonstrates how technology can bring such differences between policy and practice in tangible solutions that are supported by data.

- **Madagascar**

A report on the public procurement system in Madagascar, for instance, reports that **68%** of all public contracts awarded during 2013-2014 did not meet established procurement standards; this therefore hints at several flaws within the system. This could question transparency because non-compliance invites a potential corruption risk: in the non-standard process of awarding a contract, this is much easier to manipulate and operate under unfair procedures.

The report further states that even though **85%** of firms had been engaged in procurement processes that were designed to facilitate corruption, only **45%** of them recognized bribery in the process. Herein lies a vital contradiction between the general existence of corruption and its actual reporting, thus making the whole issue endemic and difficult to untangle within the current public procurement system.

(Reported Corruption vs. Experience of Corruption in Public Procurement Contracts Bernard Gauthier | Frédéric Lesné)

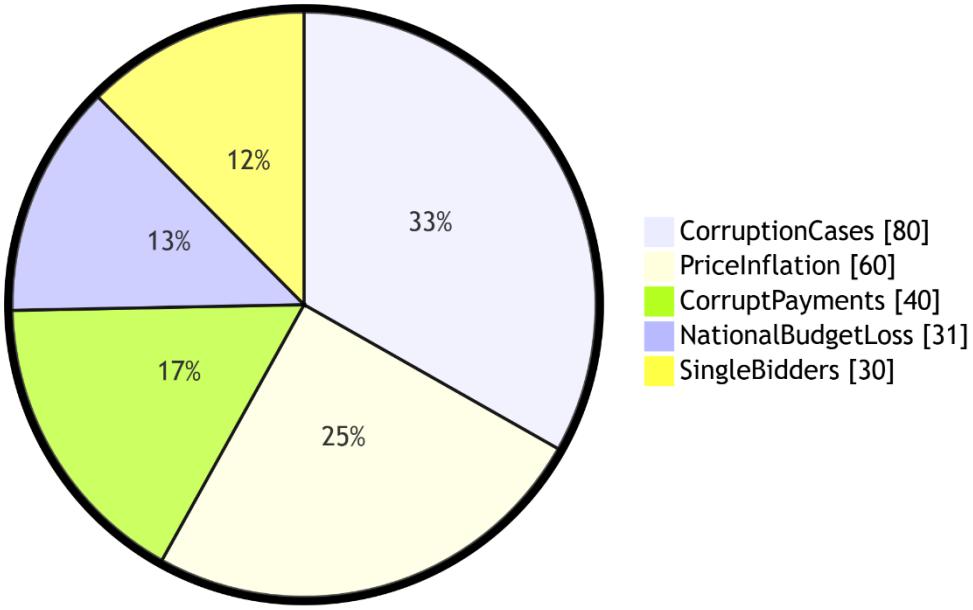


- **Kenya**

Public procurement remains one of the most susceptible sectors to corruption in Kenya, as research shows, contributing about 80% of the entire country's reported cases of corruption. The figure reveals the necessity of procurement to the realm of public finance and governance. Another disturbing aspect is that many goods and services are procured at prices inflated by over **60%** against the regular market rates instead of **15-20%.** Such kinds of activities do adversely affect the standards of public service delivery while they cause public discontent and substantial financial losses. Further, around **25%-30%** of tenders face only one bidder for the acquisition making opportunities to thrive in improper practice even higher. (EACC, 2018) (University of Nairobi, 2014).

Corruption ranges from one end of public procurement to the other with businesses spending **30%** to **50%** of project costs on provision of kickbacks for contracts. A symptom of a more sizable malignancy: costs balloon and fair competition shrivels. According to estimates, procurement corruption swindles about **30%-33%** of the Kenyan national budget each year, thereby punishing development and economic progress heavily. Findings that advocate for immediate comprehensive reform including installation of monitoring systems, rigorous supervision, and advanced technologies such as machine learning with the aim of detection and prevention against such malpractices. The base of such efforts must also include transparency, equity and accountability in Kenya's fight against such challenges in the restoration of integrity in public procurement. (Stratford Journal, 2020) (AFRICOG, 2015).

SEALING CORRUPTION LOOPHOLES IN PUBLIC PROCUREMENT



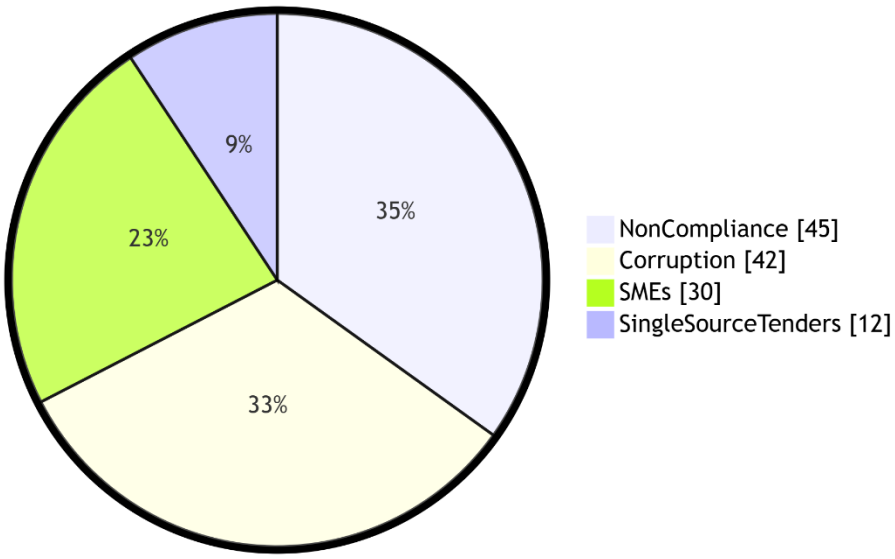
(Evaluation-of-corruption-in-the-public-procurement – A Kenyan experience)

➢ South Africa

There exists a very strong legislative framework upon which the South African public procurement system is built: The Public Finance Management Act (PFMA) and the Municipal Finance Management Act (MFMA), among others, impose stringent restrictions over public expenditure. They also show that this framework is the Preferential Procurement Policy Framework Act (PPPFA), which is supposed to really advance Black Economic Empowerment (BEE) by favoring historically disadvantaged groups in procurement processes-in fact, approximately **30%** of contracts are supposed to be given to small and medium-sized enterprises (SMEs) so that local businesses would benefit from such processes and become inclusive. It also adds that the much-expected Public Procurement Bill would smoothen the legal framework, improve compliance, and boost accountability for all processes procurement. (Transparency International, 2022)

On the downside, both internal and external challenges take away the efficacy of that system amid this foundation. Luckily, however, corruption has remained, up to this time, one of the seasoned former issues, with **R30 billion** worth being surrendered each year on account of irregular procurement practices. Substantially showing gaps in compliance against standards, the Auditor-General had recently established a **45%** non-compliance rate in regard to some essential tenders. Most often single-source tender deserving but about **10-15%** of the cases, thus avoiding competitive bidding, raises questions over transparency. Besides, there are tremendous price differences in the order of **100-200%** as compared to the market average. Agreements under public-private partnerships add more complexity, as misaligned incentives often lead to such higher costs without a similar number of advantages. It is important to address these challenges through enhanced supervision, technological advancement, and better enforcement of the existing regulations. (Auditor-General of South Africa, 2022) (Public Sector Supply Chain Research, 2021)

(Fraudulent award of tenders and South African government By Phuti Frans Manaka)



(Corruption-in-the-public-procurement – South Africa)

- **Public Procurement Comparison Across Countries**

Indicator	**Zambia (Our System Data)**	**Kenya**	**South Africa**	**Madagascar**
Single-Bid Tenders (%)	35.68%	10-15%	10-15%	Significantly higher non-compliance (68% tenders not complying with Procurement Code)
Tenders with Price Variations (Significant Price Differences) (%)	6.05%	Up to 100%-200% for some cases	Price inflation reported, though varies per sector	Price inflation significant, with observed inflation rates as high as 200%
Tenders with Price Outliers (%)	3.58%	45%	Price variations (reported as high as 100%-200% inflation for similar tenders)	Large discrepancy in pricing of tenders (not explicitly mentioned, but inferred as a key issue)
Corruption or Procurement Irregularities	Potential corruption and fraud risks (as observed in pricing issues)	Corruption costs government billions (approx. 30 billion KSh annually)	R30 billion loss annually due to procurement irregularities	Irregularities in up to 68% of tenders awarded outside public procurement codes
Price Inflation and Competitive Bidding Compliance	Price variations as high as 200% in some cases (comparable to market rates)	Significant price inflation across certain tenders	Price variations and single-source tenders bypass competitive bidding	Much higher risk of significant inflation with contracts compared to market
Regulatory Compliance	Less regulatory compliance monitoring (data from Auditor General suggests improvement needed)	Reported non-compliance rate of 45% in key tenders (study by Auditor-General)	45% non-compliance rate according to the Auditor-General	Widespread failure to comply with proper procurement codes, highly irregular award methods
Preferential Procurement Focus (e.g., BEE or Similar Initiatives)	No explicit preferential procurement policy similar to South Africa's BEE framework	Public-private partnerships under scrutiny for inefficient allocation of resources	Preferential Procurement Policy to promote BEE initiatives (about 30% earmarked for SMEs, particularly those owned by previously disadvantaged groups)	Limited explicit preferential procurement measures reported

5.4 Implications of Results.

- **Loss of Value for Money:** Tenders with price outliers or inflated bids may lead to excessive government spending, depriving the public of much-needed services and infrastructure development.

- **Missed Opportunities for SMEs and Local Economic Growth:** Low competition means that small and medium-sized enterprises cannot participate in public procurement processes, let alone benefit from them, and, as such, thwart localized economic empowerment.

- **Necessity for Policy Reform:** The need to review our Zambian regulatory frameworks, drawing lessons from other countries.

- **Underutilization of the Procurement System**: With a very high proportion (**35.68%**) of tenders having a single bid, there are signs that competitive tendering is not followed in many other cases. This absence of competition undermines principles of transparency and accountability in public procurement.

- **Risk of Corruption**: The very big price differences on **6.05%** tenders that have been only responded to by two bidders, as well outliers on **3.58%** of tenders that have received bids by more than two bidders, indicate a possibility for bid rigging or collusion-strong evidence of corrupt practices. These symptoms require stricter monitoring and investigation.

- **Low Stakeholder Participation**: Low numbers of bidders in a tender may indicate little promotion of opportunities, besides having possible barriers to taking part, including convoluted qualification processes and lack of confidence in the system.

Some disturbing features come out of this system where serious flaws can be observed in Zambia public procurement, with a large portion having characteristics of suspiciousness inviting necessary intervention.

With such red flags raised, one would have thought of a deep investigation into such anomalies, with law enforcement and oversight agencies like the ACC, the Zambia Police, and the Auditor General's office involved. Their involvement should be to establish whether these irregularities emanate from systemic weaknesses, deliberate collusion, or outright corruption. The strengthening of the investigation frameworks, with stringent standards for accountability, ensures the shut of these procurement cases once and for all so people may gain confidence in the system. Indeed, sentences send out strong notice that no fraudulent deal shall pass muster to enforce equal opportunities and ethical dealings in the Zambian Procurement Process.

CHAPTER 6 SUMMARY AND CONCLUSION

6.1 Summary of main Findings

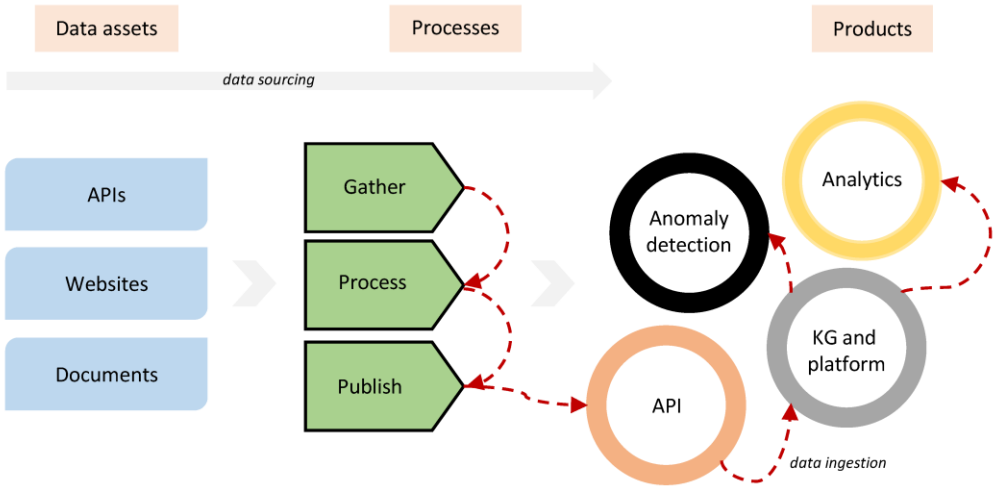
- **Machine Learning**

Machine learning and large language models have completely flipped the script with regard to how complex datasets may be analyzed, especially along qualitative dimensions hitherto unreachable by classical algorithms. These models can process and provide meaning from huge volumes of unstructured data in order to come up with patterns and insights that go beyond numerical analysis.

For example, while analyzing public procurement entities, machine learning has allowed us to discover unusual transactions, unreliable pricing patterns that seem not to conform to any simple quantitative scheme, and to identify fraud that usually depends on a complicated combination of the two. Such new horizons grant access to better illumination into an increasingly intricate system, revealing within it the essential but often silent inefficiencies and irregularities that thwart transparency and accountability.

Machine learning has changed the entire working space of computer analysts, as they can now treat activities that were impossible in the past using computation without yet defining their quantitative reach. Moving beyond numbers, such data invite analysts to consider the entire text of narratives and behavior of bidders, along with nuanced contextual factors. That leap is qualitative, so it allows identifying risks and corruption patterns today, not just by counting numbers but through disambiguation of relationships, languages, and contexts in that data. Such depth of analysis is the last mile connecting raw data to intelligence-to-action.

This shift ensures that decisions are made from holistic and, indeed, data driven, quantitative measures when necessary and qualitative insights when needed. Machine learning will have freed the analyst from the limitations of numbers only as it becomes possible to analyze broader systemic problems and solutions. One technology advance into the other because of the robustness derived from the merger between advanced methods.



6.2 Contribution to the body of knowledge

6.2.1 Limitations of the Research

While the system has shown promise in analysing bid data and detecting irregularities in public procurement, its capacity is limited to the data scope available during the study. For example, we could not analyse ownership and shareholder data of companies that won successful tenders. This is because it was not possible to get comprehensive data from the Patents and Companies Registration Agency (PACRA), which maintains the registry of company ownership and shareholder details in Zambia. Without this critical level of information, the system could not go into potentially very significant relationships, such as cross-ownership among bidders or conflicts of interest linked to shareholder ties.

Such missing insights create a void in the overall efficacy of the procurement monitoring system. Identifying elements of and relationships established from ownership is crucial for the detection of more sophisticated types of corruption such as bid collusion or undue influence. Bid collusion or undue influence do not require take place between two companies but more often than not will involve the interconnectedness of many other companies and/or hold-up entities with hidden affiliations that can only emerge after an extensive analysis of shareholders and ownership data. By incorporating this aspect of analysis, the system would significantly enhance its ability to detect not only pricing or bidding irregularities but also some systemic risks related to corporate governance and violations of ethics.

In addition to the inability to access PACRA's data on company ownership and shareholder information, several other limitations hinder the system's comprehensive monitoring capabilities. Firstly, the analysis relies heavily on data completeness and accuracy from Procuring Entities. However, data inconsistencies, such as missing tender documentation, incomplete bid records, and inadequate metadata about procurement processes, limit the system's ability to produce an all-encompassing evaluation. This challenge underscores the need for standardized reporting practices and mechanisms to ensure data integrity at every stage of the procurement lifecycle.

Secondly, while the system excels at identifying quantitative anomalies, such as significant price variations and single-bid submissions, it is limited in its ability to capture qualitative aspects of procurement processes. Factors such as bidder reputation, project execution quality, and past performance, which often contribute to procurement outcomes, are beyond the system's analytical scope due to a lack of structured and accessible data. Expanding the system to incorporate qualitative metrics could yield a more holistic assessment of procurement processes and outcomes.

6.2.2 Future works

Going forward, to maximize the potential of the Transparent Electronic Public Procurement Monitoring System's effectiveness and impact, it is necessary to work with relevant key stakeholders within Zambia's governance and oversight landscape. This system has been shown to be quite good at identifying suspicious patterns in public procurement; however, it will need partnerships with institutions such as the Zambia Public Procurement Authority (ZPPA), Patents and Companies Registration Agency (PACRA), and law enforcement agencies like the Financial Intelligence Centre (FIC) and the Anti-Corruption Commission (ACC) to realize that full potential. This will be integrated through one single database where the entities will share resources and information on the integrated platform for tracking procurement data, ownership of companies, financial transactions, and past records for corruption and misconduct.

This would, therefore, enhance proactive detection and investigation of procurement irregularities. For instance, suspicious tenders highlighted through the system may be further checked against data by PACRA to show links indicating a conflict of interest or fraud through hidden ownership structures. In the same way, money laundering or kickbacks for certain tenders identified through financial patterns by FIC could complement procurement data. These would be value-added joint ventures that increase accountability, further enhance transparency and compliance, and provide timely and actionable insight to the decision-maker/investigator. Positioning the system as a shared resource for these agencies will not only engender inter-agency collaboration but also position the system at the core of modernizing Zambia's public procurement oversight mechanisms.

REFERENCES

Ahmed, E., Mahmood, A.N. and Hu, J., 2016. A survey of network anomaly detection techniques. Journal of Network and Computer Applications, 60, pp.19-31.

Arrowsmith, S., 2010. Public procurement: Basic concepts and the coverage of procurement rules. In: S. Arrowsmith, ed. public procurement regulation: An introduction. Nottingham: University of Nottingham Press, pp.1-24.

Bai, C. and Sarkis, J., 2022. The role of procurement transparency in promoting sustainability. Journal of Cleaner Production, 315, 128100.

Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., et al., 2001. Manifesto for Agile Software Development. Available at: https://agilemanifesto.org (Accessed: 8 October 2024).

Benedict, D., Barry, S. and Sadiq, M., 2018. A survey of machine learning techniques in public procurement fraud detection. International Journal of Advanced Research in Artificial Intelligence, 7(4), pp.58-63.

Bertin, J., 2010. Semiology of graphics: diagrams, networks, maps. Redlands, CA: ESRI Press.

Bose, R., 2019. Advanced analytics and machine learning in e-procurement. Journal of Business Research, 101, pp.211-218. Available at: https://doi.org/10.1016/j.jbusres.2019.04.020

Brockett, P.L., Xia, X. and Derrig, R.A., 2002. Using Kohonen's self-organizing feature map to uncover automobile bodily injury claims fraud. Journal of Risk and Insurance, 69(3), pp.245-274.

Bryman, A. and Bell, E., 2015. Business Research Methods. 4th ed. Oxford University Press.

Chandola, V., Banerjee, A. and Kumar, V., 2009. Anomaly detection: A survey. ACM Computing Surveys, 41(3), pp.1-58.

Chaudhuri, S. and Dayal, U., 1997. An overview of data warehousing and OLAP technology. ACM SIGMOD Record, 26(1), pp.65-74.

Chen, L. and Liu, F., 2023. The role of machine learning and analytics in promoting transparency in procurement. International Journal of Supply Chain Management, 18(3), pp.87-99. Available at: https://doi.org/10.1080/ijscm.2023.11.002 (Accessed: 5 October 2024).

Costa, A.A. and Tavares, L.V., 2018. Advanced data analytics for public procurement: New opportunities to harness big data. International Journal of Information Management, 40, pp.70-79. Available at: https://doi.org/10.1016/j.ijinfomgt.2018.01.004

CUTS Lusaka, 2021. Enhanced oversight and monitoring of public procurement in Zambia's education sector. Available at: https://cuts-lusaka.org/enhanced-oversight-and-monitoring-of-public-procurement-in-zambias-education-sector/

Davila, A., Gupta, M. and Palmer, R., 2003. Moving procurement systems to the internet: The adoption and use of e-procurement technology models. European Management Journal, 21(1), pp.11-23.

Dada, O.A., Fashola, F.A. and Abubakar, M.S., 2022. Enhancing transparency in public procurement through e-procurement systems: A study of the Nigerian experience. International Journal of Public Sector Management, 35(3), pp.335-351.

European Commission, 2015. Public procurement - a study on corruption in the public procurement process. Publications Office of the European Union. Available at: https://doi.org/10.2799/488782

European Parliament, 2021. Artificial Intelligence in public procurement. Available at: www.europarl.europa.eu/RegData/etudes/STUD/2021/691722/IPOL_STU(2021)691722_EN.pdf

Ferraz, C., Finan, F. and Moreira, D.B., 2021. Corruption learning: Evidence from driving tests in Brazil. Journal of Public Economics, 204, p.104553.

Flynn, A. and Davis, P., 2017. Investigating the implementation of e-procurement and the impact on public procurement in UK local government. Journal of Purchasing and Supply Management, 23(3), pp.207-219. Available at: https://doi.org/10.1016/j.pursup.2017.07.001

Gershon, M., 2005. Public procurement: efficiency and transparency. Journal of Public Policy and Administration, 15(3), pp.231-249.

Government of Zambia, 2012. Anti-Corruption Commission Act 2012. Lusaka: Government Printers.

Han, J., Pei, J. and Kamber, M., 2011. Data Mining: Concepts and Techniques. 3rd ed. Morgan Kaufmann.

Hevner, A.R., March, S.T., Park, J. and Ram, S., 2004. Design Science in Information Systems Research. MIS Quarterly, 28(1), pp.75-105.

Hertie School, n.d. How public organizations can use AI in anti-corruption. Available at: https://www.hertie-school.org/en/digital-governance/research/blog/detail/content/how-public-organisations-can-use-ai-in-anti-corruption-what-we-know-so-far-and-why-we-need-to-learn-more-about-it

Kühn, M. and Weingarten, P., 2018. The role of big data and analytics in the public procurement process: A framework and research agenda. Public Administration Review, 78(6), pp.905-913. Available at: https://doi.org/10.1111/puar.12912
Li, X., Wang, Y. and Sun, Z., 2023. Model evaluation metrics for fraud detection and risk assessment in e-procurement. Journal of Computational Intelligence Systems, 34(2), pp.125-138. Available at: https://doi.org/10.1016/j.jcis.2023.01.005 (Accessed: 5 October 2024).

Mishra, P. and Kumar, R., 2021. e-Procurement Systems in Developing Countries: Challenges and Benefits. International Journal of Public Administration, 44(5), pp.343-354.

Nguyen, T.H. and Tran, P., 2023. The role of big data and analytics in decision-making processes in procurement. International Journal of Production Economics, 242, p.108184. Available at: https://doi.org/10.1016/j.ijpe.2021.108184 (Accessed: 5 October 2024).

Nguyen, T.H. and Schniederjans, D.G., 2022. The role of machine learning in public procurement: bidder evaluation and risk assessment. Journal of Public Procurement, 22(3), pp.215-235. Available at: https://doi.org/10.1108/JOPP-03-2022-0013 (Accessed: 5 October 2024).

OECD, 2016. Preventing corruption in public procurement. OECD Publishing. Available at: https://doi.org/10.1787/9789264265820-en

OECD, 2016. Corruption in public procurement. Available at: https://www.oecd.org/gov/ethics/Corruption-Public-Procurement-Brochure.pdf

Open Contracting Partnership, 2022. Let's get real about AI's potential to end corruption in procurement. Available at: https://www.open-contracting.org/2022/01/20/lets-get-real-about-ais-potential-to-end-corruption-in-procurement/

Reddy, A. and Singh, V., 2022. Digital Procurement in India: Enhancing Transparency and Reducing Corruption. Journal of Public Procurement, 22(1), pp.56-71.

Saini, S., 2022. Leveraging AI to fight corruption: Lessons from Brazil and India. Available at: https://www.linkedin.com/pulse/leveraging-ai-fight-corruption-lessons-from-brazil-indias-saini-0lrde/

Schniederjans, D.G. and Hales, D.N., 2016. Cloud computing and its impact on e-commerce in the procurement process: Realities and myths. Journal of Business Logistics, 37(1), pp.20-30.

Transparency International, 2020. Corruption in Public Procurement: A Global Survey of the Public Procurement System. Available at: https://www.transparency.org/en/publications/corruption-in-public-procurement-a-global-survey-of-the-public-procurement-system

Uyar, A., 2021. Machine learning in public procurement: Techniques and applications. Expert Systems with Applications, 167, p.114176. Available at: https://doi.org/10.1016/j.eswa.2020.114176 (Accessed: 5 October 2024).

Yahya, S., 2019. Public procurement fraud: The role of big data analytics. Journal of Public Procurement, 19(2), pp.105-123. Available at: https://doi.org/10.1108/JOPP-05-2018-0042 (Accessed: 5 October 2024).

Government of Zambia. (2021). National Anti-Corruption Policy. Lusaka: Government Printer.

Peters, G. W., & Panayi, E. (2016). Understanding modern banking ledgers through blockchain technologies: Future of transaction processing and smart contracts on the internet of money. Banking Beyond Banks and Money, 239–278.

Tapscott, D., & Tapscott, A. (2016). Blockchain revolution: How the technology behind bitcoin is changing money, business, and the world. Penguin.

Government E Procurement: Digital Transformation in Public Procurement: Lessons for Startups Available: https://fastercapital.com/content/Government-E-Procurement--Digital-Transformation-in-Public-Procurement--Lessons-for-Startups.html Accessed 29 October 2024

Blockchain in Procurement — The Ultimate Guide of 2024 Available: https://procurementtactics.com/blockchain-in-procurement/ Accessed 29 October 2024

Cheng, Y., Tu, Z., Meng, F., & Liu, Y. (2020). Document analysis using natural language processing: A comprehensive review. Information Processing & Management, 57(4), 102-116.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998-6008.

https://www.uonbi.ac.ke/sites/default/files/public_procurement_loopholes.pdf

https://eacc.go.ke/default/wp-content/uploads/2018/09/Evaluation-of-corruption-in-the-public-procurement.pdf

https://uir.unisa.ac.za/bitstream/handle/10500/28404/dissertation_manaka_pf.pdf?sequence=1&isAllowed=y

https://ferdi.fr/dl/df-CDSUJyyvS8yHE6r8fRaykECJ/ferdi-p242-reported-corruption-vs-experience-of-corruption-in-public.pdf

APPENDICES

APPENDIX A – Schedule of work

Task No.	Duration (Days)	Phase	Task Description
1	1	**START**	Project Kick-off Meeting
2	3	**INITIATION**	Identify Stakeholders
3	1		Set Up Initial Meetings
4	1		Assess Current Public Procurement System
5	2		Review Interactions with Related Agencies
6	4	**PLANNING**	Conduct Literature Review on e-Procurement & ML
7	6		Analyze Similar Existing Systems in Public Procurement
8	1		Gather Data Requirements for Procurement
9	3		Research ML Methodologies and Tools
10	2		Develop Gantt Chart and Project Timeline
11	12	**EXECUTION**	Conduct System Analysis for Procurement Monitoring
12	16		Design System Architecture and Machine Learning Framework
13	3		Select Technology Stack (Frontend & Backend)
14	1		Choose Development Methodology (e.g., Agile)
15	1		Finalize Programming Languages and Tools
16	30		Develop and Code Core System Features
17	5		Implement Initial System Modules
18	5		Integrate ML Model with Data Sources
19	5		System Testing (Unit & Integration Testing)
20	2		User Acceptance Testing (UAT)
21	2		Prepare Deployment Plan
22	1		Deploy System
23	30	**MONITORING**	Monitor System Performance and Collect Feedback
24	4	**CLOSING**	Evaluate System Effectiveness
25	2		Prepare and Present Final Project Report
26	1		Closing Meeting and Project Handover
27	0	**END**	Project Closure

APPENDIX B – Project Charter

Project Title: Design and Implementation of a transparent electronic public procurement Monitoring System through the use of machine learning and analytics.

Project start Date: 09/09/2024 **Projected Finish Date:** 27/12/2024

Project Manager: Jack Bwaaza

Project Objectives:
This project examines the existing e-Procurement system for points where the application of machine learning and analytics can facilitate greater transparency. The project then applies ML models in two stages: first, to get market price estimates, and secondly for categorizing bidders. The different models studied in this work include those for fraud detection, supplier risk assessment, and bid evaluation. This project will also apply data analytics to predictive procurement strategies and analytics-informed decision-making, eventually assessing the impact of ML and analytics on efficiency, equity, and accountability within public e-Procurement processes.

Success Criteria: Successfully design and implementation of the system.

Approach:

- Investigate the current system
- Design and Implementation of the new system and ensure the system is able to evaluate ML models for fraud detection, supplier risk assessment, and bid evaluation in e-Procurement
- Test the system

Roles and Responsibilities

Name	Role	Position	Contact Information
Clive Choongo	Sponsor	CEO	CliveC@yahoo.com
Jack Bwaaza	Project Manager	Manager	jbwaaza@gmail.com
Christopher Chavunma	Team member	Database Administrator	chavuma2013@yahoo.com
Ethel Nunkwe	Team member	Consultant - ZPPA	Ethel.nunkwe@zppa.org.zm
Bonaventure Moonga	Team member	Systems Analyst	bonaventure@bgsgroup.co.zm
Jose S. Simuyi	Team member	Project Consultant	simuyi@yahoo.com
Kelly Ikowa	Team member	Procurement officer	IkowaK@gmail.com

Comments:

APPENDIX C – Project Budget

1 Systems Development and Design				
Unit	**Input type**	**Quantity**	**Unit cost**	**Cost (ZMW)**
System Architecture Design	Lump sum	1	50,000	50,000
Software Development (Frontend & Backend)	Lump sum	1	200,000.00	200,000.00
Database Design and Setup	Lump sum	1	40,000	40,000
Machine Learning Model Development	Lump sum	1	100,000.00	100,000.00
User Interface/User Experience (UI/UX) Design	Lump sum	1	30,000	30,000
Subtotal 1				**420,000.00**
2 Technology and Infrastructure				
Cloud Hosting and Infrastructure	Annual	1	100,000.00	100,000.00
Development and Collaboration Tools	Unit	10	3,000.00	30,000.00
Security and Compliance Tools	Lump sum	1	50,000	50,000.00
Subtotal 2				**180,000.00**
3 Personnel Costs				
Salaries for Development Team	Unit	5	12,000.00	60,000.00
Training and Capacity Building	Unit	10	5,000.00	50,000.00
Subtotal 3				**110,000.00**
4 Testing and Deployment				
System Testing (User Acceptance Testing)	Lump sum	1	20,000.00	20,000.00
Deployment Costs	Lumpsum	1	30,000.00	30,000.00
Subtotal 4				**50,000.00**
5 Ongoing Maintenance and Support				
Annual Maintenance and Updates	Annual	1	50,000.00	50,000.00
Technical Support and Helpdesk Services	Annual	1	30,000.00	30,000.00
Subtotal 1				**80,000.00**
TOTAL =				**840,000.00**

APPENDIX D – Data Samples

- **Procurement Details**

Tender ID	Project Name	Project Type (Goods, Services, Works)	Estimated Budget	Contract Value	Start Date

- **Bidder Information**

Bidder ID	Company Name	Registration Number (PACRA)	Country of Origin	Bid Amount	Bid Status (Awarded, Not Awarded)

- **Fraud Indicators**

Number of Projects Previously Won	Average Bid Deviation	Supplier Risk Rating

- **Market Analysis**

Average Market Price	Price Variance	Competitor Prices

- **Bid Evaluation Criteria**

Technical Score	Financial Score	Total Score

https://eprocure.zppa.org.zm/epps/app/viewPublication.do?selectedItem=app/viewPublication.do

1	Organisation Name :	ZCCM INVESTMENTS HOLDINGS PLC
2	Organisation Abbreviation:	ZCCM-IH
3	Address:	30048
4	Postal Code:	10101
5	City:	Lusaka
6	Country:	Zambia
7	Email:	kabwekd@zccm-ih.com.zm
8	Phone Number:	26021122035
9	Fax:	
10	Website:	http://www.zccm-ih.com.zm

APPENDIX E – Code Snippets

➢ Procuring Import (Excel Import Data)

```
import pandas as pd
import numpy as np
from sqlalchemy import create_engine,text
from tqdm import tqdm_notebook, tqdm


########## a function that gets table data from a url
def executeSQL(sql,engine):
   connection = engine.connect()
   result = connection.execute(sql)
   connection.commit()
   connection.close()

################ connect the database
engine = create_engine('mysql+pymysql://root:@localhost:3306/procurement')
connection = engine.connect()

################ connect to the excel document
print("########## Connecting to the excel document ##########")
#path = input("Enter the daocument path:")
path = 'C:\Python\ZPPA.xlsx'

################ save excel data in the database
for t in tqdm(range(0,212)):
  data = pd.read_excel(path, sheet_name=t,header=None)
  data = data[1].values
  # create the sql statemment
  try:
     executeSQL(text("INSERT INTO `procuring_entities`(`Organisation`, `Abbreviation`, `Address`, `PostalCode`, `City`, `Country`, `Email`,
`PhoneNumber`, `Fax`, `Website`) VALUES
('"+str(data[0])+"','"+str(data[1])+"','"+str(data[2])+"','"+str(data[3])+"','"+str(data[4])+"','"+str(data[5])+"','"+str(data[6])+"','"+str(data[7])+"','"+
str(data[8])+"','"+str(data[9])+"')"),engine)
  except:
     error = 'e'
```

➢ Importing Tenders and Bids (Web Scraping)

```
print("Importing modules and loading functions....")
import pandas as pd
import numpy as np
from sqlalchemy import create_engine,text
import requests
from bs4 import BeautifulSoup
from tqdm import tqdm_notebook, tqdm
from datetime import datetime, timedelta


########## a function that gets table data from a url
def executeSQL(sql,engine):
   try:
      connection = engine.connect()
      connection.execute(sql)
      connection.commit()
      connection.close()
      return True
   except:
      return False
```

```
########## a function that gets table data from a url
def getTable(url):
    # Send an HTTP GET request to the URL
    response = requests.get(url)
    # Check if the request was successful
    if response.status_code == 200:
        # Parse the HTML content of the page using BeautifulSoup
        soup = BeautifulSoup(response.text, 'html.parser')
        # Find the table with the ID "T0"
        table = soup.find('table')
        if table:
            # Initialize empty lists to store the data
            headers = []
            data = []
            # Extract table headers
            header_row = table.find('thead').find('tr')
            for header_cell in header_row.find_all('th'):
                headers.append(header_cell.text.strip())
            # Extract table rows
            rows = table.find('tbody').find_all('tr')
            for row in rows:
                row_data = []
                for cell in row.find_all('td'):
                    if cell.text.strip()=='View Opened Bids':
                        # Find the anchor tag within the cell
                        anchor = cell.find('a', href=True)
                        if anchor:
                            # Get the href attribute (URL) from the anchor tag
                            view_opened_bids_url = 'https://eprocure.zppa.org.zm'+anchor['href']
                            row_data.append(view_opened_bids_url)
                    else:
                        row_data.append(cell.text.strip())
                data.append(row_data)
        else:
            print("Table with ID 'T0' not found on the page.")
    else:
        print("Failed to retrieve the webpage. Status code:", response.status_code)
    # return
    df = pd.DataFrame(data, columns=headers)
    return df


########## a function that gets dt list data from a url
def getList(url):
    # Send an HTTP GET request to the URL
    response = requests.get(BidsLink)
    if response.status_code == 200:
        # Parse the HTML content of the page
        soup = BeautifulSoup(response.text, 'html.parser')
        # Find the <dl> element with class 'Grid'
        grid_element = soup.find('dl', class_='Grid')
        # Initialize empty lists to store data
        dt_values = []
        dd_values = []
        # Extract data and append to lists
        dt_elements = grid_element.find_all('dt')
        dd_elements = grid_element.find_all('dd')
        for dt, dd in zip(dt_elements, dd_elements):
            dt_values.append(dt.text.strip())
            dd_values.append(dd.text.strip())
    # return
    return dt_values,dd_values
print("Loaded.")


################ connect the database
```

```
print("Connecting to the database....")
engine = create_engine('mysql+pymysql://root:@localhost:3306/procurement')
connection = engine.connect()
print("Connected.")


################ get the tenders
print("Collecting tenders from the ZPPA website....")
# URL of the webpage to scrape
url = "https://eprocure.zppa.org.zm/epps/common/viewOpenedTenders.do?selectedItem=common%2FviewOpenedTenders.do&T01_ps=100"
tenders = getTable(url)
tenders.set_index('Reference number', drop=False, inplace=True)
tenders['Award Date'] = tenders['Award Date'].str.replace('CAT', '', regex=False)
tenders['Bids Submission Deadline'] = tenders['Bids Submission Deadline'].str.replace('CAT', '', regex=False)
tenders['Award Date'] = pd.to_datetime(tenders['Award Date'], format='%a %b %d %H:%M:%S %Y')
tenders['Bids Submission Deadline'] = pd.to_datetime(tenders['Bids Submission Deadline'], format='%a %b %d %H:%M:%S %Y')
tenders.drop(columns=['#'],inplace=True)
print(str(tenders.shape[0])+" records collected")


################ process the tenders
print("Processing tenders....")
for t in tqdm(range(700,tenders.shape[0])):
   ############### save tenders
   index = tenders.index[t]
   ReferenceNumber = tenders['Reference number'].values[t]
   Title = tenders['CfT Title'].values[t]
   ProcuringEntity = tenders['PE'].values[t]
   BidsSubmissionDeadline = tenders['Bids Submission Deadline'].values[t]
   ProcurementMethod = tenders['Procurement Method'].values[t]
   OpenedBids = tenders['Opened Bids'].values[t]
   AwardDate = tenders['Award Date'].values[t]
   Status = tenders['Status'].values[t]
   BidsLink = tenders['Opened Bids'].values[t]
   ############### check if the tender has already been saved tenders.columns
   if pd.read_sql('SELECT * FROM `procure_tenders` WHERE `Referencenumber`="'+str(ReferenceNumber)+'"', engine).shape[0]==0:
      # save the tender
      result = executeSQL(text('INSERT INTO
`procure_tenders`(`Referencenumber`,`CfTTitle`,`PE`,`BidsSubmissionDeadline`,`ProcurementMethod`,`OpenedBids`,`AwardDate`,`Status`)
VALUES ("'+ReferenceNumber+'", "'+Title+'", "'+ProcuringEntity+'", "'+str(BidsSubmissionDeadline)+'", "'+ProcurementMethod+'",
"'+OpenedBids+'", "'+str(AwardDate)+'","'+Status+'")'),engine)
      erer = 'INSERT INTO
`procure_tenders`(`Referencenumber`,`CfTTitle`,`PE`,`BidsSubmissionDeadline`,`ProcurementMethod`,`OpenedBids`,`AwardDate`,`Status`)
VALUES ("'+ReferenceNumber+'", "'+Title+'", "'+ProcuringEntity+'", "'+str(BidsSubmissionDeadline)+'", "'+ProcurementMethod+'",
"'+OpenedBids+'", "'+str(AwardDate)+'","'+Status+'")'
   else:
      # update the tender
      result = executeSQL(text("UPDATE `procure_tenders` SET
`CfTTitle`='"+Title+"',`PE`='"+ProcuringEntity+"',`BidsSubmissionDeadline`='"+str(BidsSubmissionDeadline)+"',`ProcurementMethod`='"+Procure
mentMethod+"',`OpenedBids`='"+OpenedBids+"',`AwardDate`='"+str(AwardDate)+"',`Status`='"+Status+"' WHERE
`Referencenumber`='"+ReferenceNumber+"'"),engine)
   ###############  get the bids
   if result==True:
      bids = getTable(BidsLink)
      # get bids opening date
      try:
         dt_values,dd_values = getList(BidsLink)
         BidsOpeningDate = dd_values[1]
         if BidsOpeningDate!='':
            connection = engine.connect()
            BidsOpeningDate = datetime.strptime(BidsOpeningDate, '%d/%m/%Y %H:%M')
            executeSQL(text('UPDATE `procure_tenders` SET `BidsOpeningDate`="'+str(BidsOpeningDate)+'" WHERE
`Referencenumber`="'+ReferenceNumber+'"'),engine)
      except:
         error = 'e'
      ################  save the bids and update the tender
```

```
    for i in range(0,bids.shape[0]):
        index = bids.index[i]
        Envelope = bids.loc[index,'Envelope']
        Supplier = bids.loc[index,'Supplier Name']
        BidID = bids.loc[index,'Bid ID']
        try:
            FinancialValue = float(bids.loc[index,'Financial Value (ZMW)'])
        except:
            FinancialValue = 0
        BidSecurityProvided = bids.loc[index,'Bid Security Provided']
        try:
            BidSecurityValue = float(bids.loc[index,'Bid Security Value'])
        except:
            BidSecurityValue = 0
        try:
            FDR = float(bids.loc[index,'FDR'])
        except:
            FDR = 0
        FDRTerms = bids.loc[index,'FDR Terms']
        # save the bid
        executeSQL(text('INSERT INTO `procure_bids`(`BidID`, `Tender`, `Supplier`, `FinancialValue`, `BidSecurityProvided`, `BidSecurityValue`,
`FDR`, `FDRTerms`) VALUES
("'+BidID+'","'+ReferenceNumber+'","'+Supplier+'",'+str(FinancialValue)+',"'+BidSecurityProvided+'",'+str(BidSecurityValue)+','+str(FDR)+',"'+FD
RTerms+'")'),engine)
```

➢ Categorization of Tenders

```
import ollama
import pandas as pd
from sqlalchemy import create_engine, text
import openai
from tqdm import tqdm
import time

def questionOllama(question):
  response = ollama.chat(model='llama3.2', messages=[
    {
      'role': 'user',
      'content': question,
    },
  ])
  return response['message']['content']

####### connect to the database
engine = create_engine("mysql+pymysql://root:@localhost:3306/procurement")

####### Retrieve data from the database
categories = pd.read_sql("SELECT * FROM `procure_categories`", engine)
tenders = pd.read_sql("SELECT * FROM `procure_tenders` WHERE `Referencenumber` NOT IN (SELECT DISTINCT `tender` FROM
`procure_tendercategories`);", engine)


####### process the tenders t=0
for t in tqdm(range(0,tenders.shape[0])):
  Referencenumber = tenders['Referencenumber'].values[t]
  CfTTitle = tenders['CfTTitle'].values[t]
  # check which category it falls under (loop through all the categories)
  for i in range(0,categories.shape[0]):
    category = categories['Category'].values[i]
    ## create the question
    question = "The following are the details of a tender that was advertised: "+CfTTitle+". Does it fall into the category: "+category+". Answer
yes or no only."
    ## get an anwer from ollama
    answer = questionOllama(question)
```

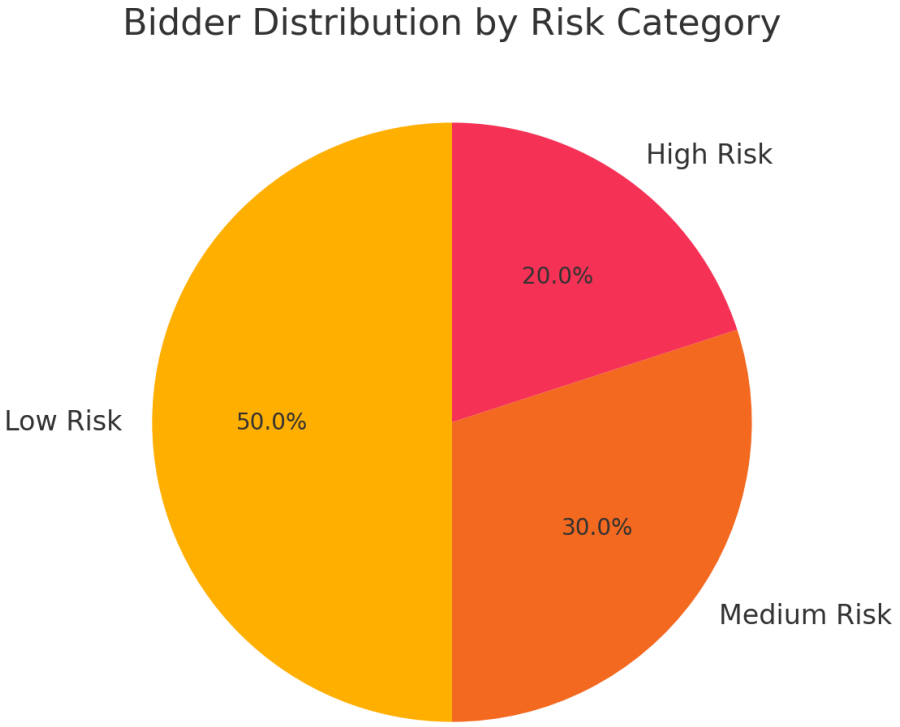
```
    answer = answer.lower()
    # Check if 'yes' appears in the text
    if 'yes' in answer:
      from sqlalchemy import create_engine
      sql = "INSERT INTO `procure_tendercategories`(`Tender`, `Category`) VALUES ('"+Referencenumber+"','"+category+"')"
      # save it to the database
      engine = create_engine("mysql+pymysql://root:@localhost:3306/procurement")
      with engine.connect() as connection:
        result = connection.execute(text(sql))
        connection.commit()
        connection.close()
```

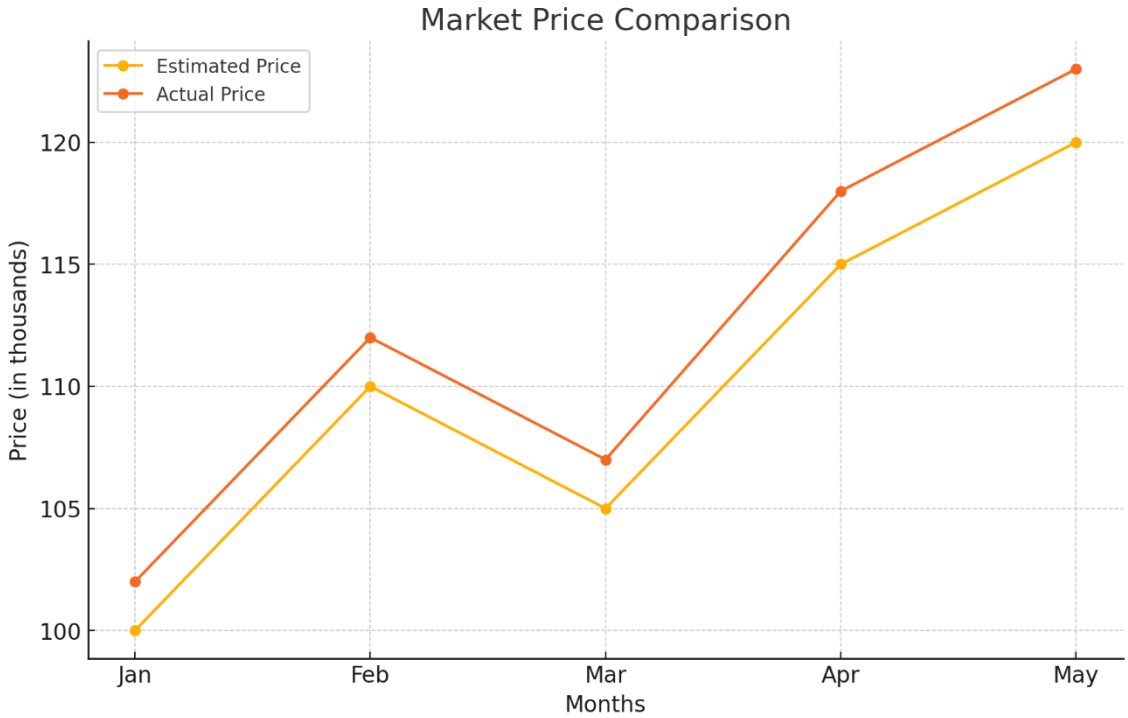
➢ Corruption Risk Analysis

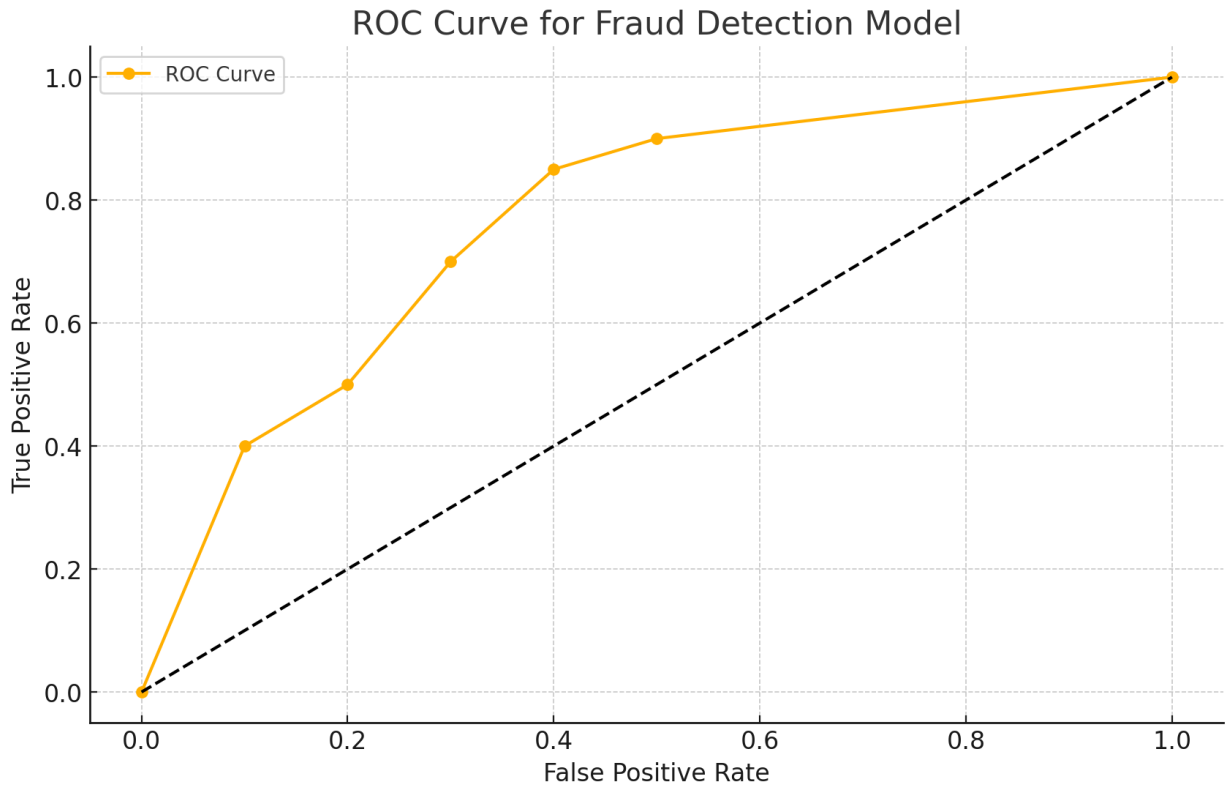
```
import ollama
import pandas as pd
from sqlalchemy import create_engine, text
import openai
from tqdm import tqdm
import time
# remove warnings
import warnings
warnings.filterwarnings('ignore')

def questionOllama(question):
  response = ollama.chat(model='llama3.2', messages=[
   {
    'role': 'user',
    'content': question,
   },
  ])
  return response['message']['content']

########## a function that gets table data from a url
def executeSQL(sql,engine):
  connection = engine.connect()
  result = connection.execute(sql)
  connection.commit()
  connection.close()

######## connect to the database
engine = create_engine("mysql+pymysql://root:@localhost:3306/procurement")


######## Retrieve data from the database
categories = pd.read_sql("SELECT * FROM `procure_categories`", engine)
tenders = pd.read_sql("SELECT t.*,COUNT(b.BidID) AS numberBids FROM `procure_tenders` t JOIN procure_bids b ON
t.Referencenumber=b.Tender GROUP BY t.Referencenumber", engine)
bids = pd.read_sql("SELECT * FROM `procure_bids`", engine)
tendersAnalysis = pd.read_sql("SELECT * FROM `procure_tenders_analysis`", engine)
bidsAnalysis = pd.read_sql("SELECT * FROM `procure_bids_analysis`", engine)

######## Check for corruption risks
for t in tqdm(range(0,tenders.shape[0])):
  tender = tenders['Referencenumber'][t]
  theBids = bids.loc[bids['Tender']==tender]
  ##### check for a single bid corruption risk
  if (tenders['numberBids'][t]==1):
    bid = theBids['BidID'].values[0]
    executeSQL(text("INSERT INTO `procure_tenders_analysis`(`Tender`, `Analysis`) SELECT '"+str(tender)+"', 'Single Bidder' WHERE NOT EXISTS
(SELECT 1 FROM `procure_tenders_analysis` WHERE `Tender` = '"+str(tender)+"' AND `Analysis` = 'Single Bidder')"),engine)
    executeSQL(text("INSERT INTO `procure_bids_analysis`(`Bid`, `Analysis`) SELECT '"+str(bid)+"', 'Single Bidder' WHERE NOT EXISTS (SELECT 1
FROM `procure_bids_analysis` WHERE `Bid` = '"+str(bid)+"' AND `Analysis` = 'Single Bidder')"),engine)
```

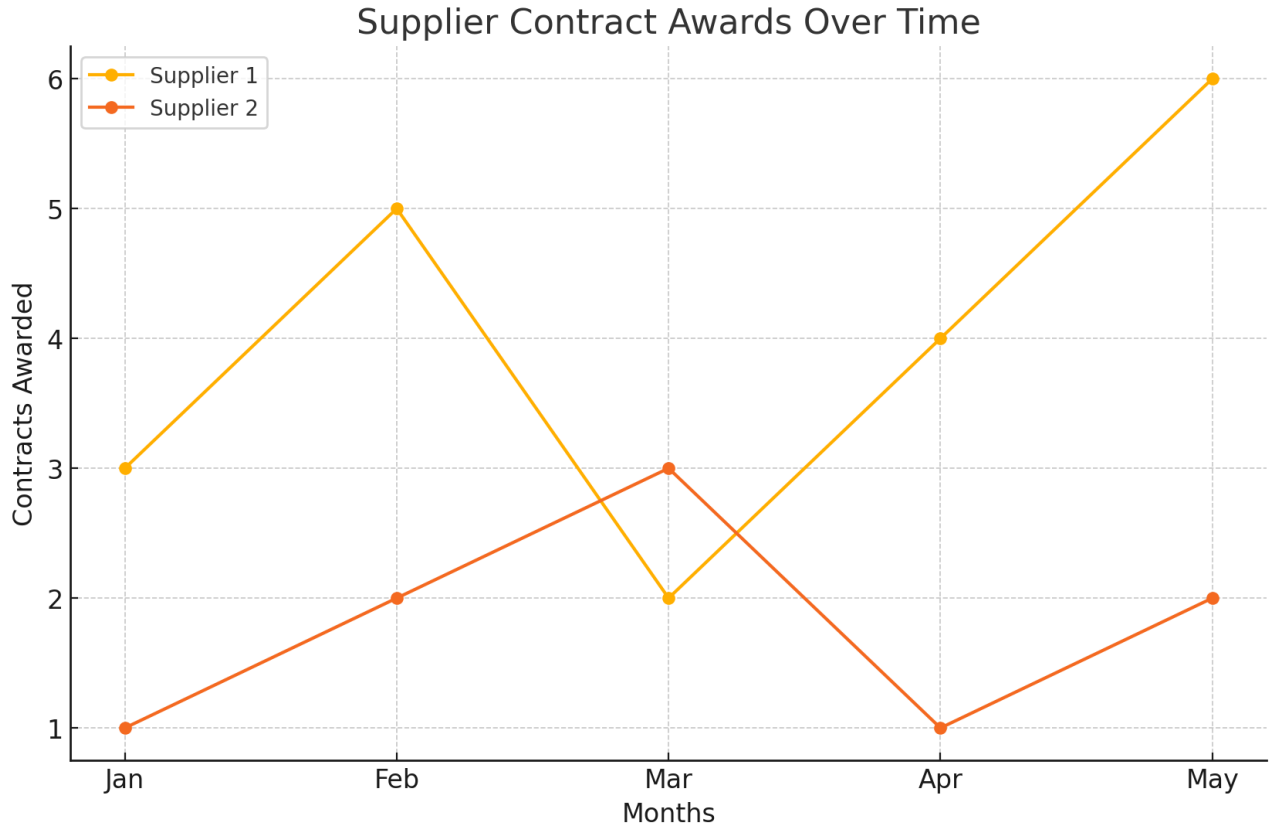
```
   ##### check for a significant price difference corruption risk (use percentage difference)
https://www.calculatorsoup.com/calculators/algebra/percent-difference-calculator.php
   if (tenders['numberBids'][t]==2):
      v1 = theBids['FinancialValue'].values[0]
      v2 = theBids['FinancialValue'].values[1]
      diff = abs(v1-v2) / ((v1+v2)/2) * 100
      if (diff>=50):
         executeSQL(text("INSERT INTO `procure_tenders_analysis`(`Tender`, `Analysis`) SELECT '"+str(tender)+"', 'Significant Price Difference
(Percentage Difference)' WHERE NOT EXISTS (SELECT 1 FROM `procure_tenders_analysis` WHERE `Tender` = '"+str(tender)+"' AND `Analysis` =
'Significant Price Difference (Percentage Difference)')"),engine)
         theBids['Notes'] = 'Significantly Higher Price'
         theBids.loc[(theBids['FinancialValue']<theBids['FinancialValue'].shift(1)),'Notes'] = 'Significantly Lower Price'
         theBids.loc[(theBids['FinancialValue']<theBids['FinancialValue'].shift(-1)),'Notes'] = 'Significantly Lower Price'
         for index,row in theBids.iterrows():
            bid = row['BidID']
            executeSQL(text("INSERT INTO `procure_bids_analysis`(`Bid`, `Analysis`,`Notes`) SELECT '"+str(bid)+"', 'Significant Price Difference
(Percentage Difference)', '"+row['Notes']+"' WHERE NOT EXISTS (SELECT 1 FROM `procure_bids_analysis` WHERE `Bid` = '"+str(bid)+"' AND
`Analysis` = 'Significant Price Difference (Percentage Difference)')"),engine)
   ##### check for a significant price difference corruption risk (use outliers)
   if (tenders['numberBids'][t]>2):
      Q1,Q2,Q3 = [theBids['FinancialValue'].quantile(q=0.25),theBids['FinancialValue'].quantile(q=0.5),theBids['FinancialValue'].quantile(q=0.75)]
      IQR = Q3-Q1
      Q0,Q4 = [Q1-(IQR*1.5),Q3+(IQR*1.5)]
      temp = theBids.loc[(theBids['FinancialValue']>Q4)]
      # Above Q4
      if temp.shape[0]>0:
         executeSQL(text("INSERT INTO `procure_tenders_analysis`(`Tender`, `Analysis`) SELECT '"+str(tender)+"', 'Significant Price Difference
(Outliers)' WHERE NOT EXISTS (SELECT 1 FROM `procure_tenders_analysis` WHERE `Tender` = '"+str(tender)+"' AND `Analysis` = 'Significant
Price Difference (Outliers)')"),engine)
         for index,row in temp.iterrows():
            bid = row['BidID']
            executeSQL(text("INSERT INTO `procure_bids_analysis`(`Bid`, `Analysis`,`Notes`) SELECT '"+str(bid)+"', 'Significant Price Difference
(Outliers)', 'Significantly Higher Price' WHERE NOT EXISTS (SELECT 1 FROM `procure_bids_analysis` WHERE `Bid` = '"+str(bid)+"' AND `Analysis` =
'Significant Price Difference (Outliers)')"),engine)
      # Below Q0
      temp = theBids.loc[(theBids['FinancialValue']<Q0)]
      if temp.shape[0]>0:
         executeSQL(text("INSERT INTO `procure_tenders_analysis`(`Tender`, `Analysis`) SELECT '"+str(tender)+"', 'Significant Price Difference
(Outliers)' WHERE NOT EXISTS (SELECT 1 FROM `procure_tenders_analysis` WHERE `Tender` = '"+str(tender)+"' AND `Analysis` = 'Significant
Price Difference (Outliers)')"),engine)
         for index,row in temp.iterrows():
            bid = row['BidID']
            executeSQL(text("INSERT INTO `procure_bids_analysis`(`Bid`, `Analysis`,`Notes`) SELECT '"+str(bid)+"', 'Significant Price Difference
(Outliers)', 'Significantly Lower Price' WHERE NOT EXISTS (SELECT 1 FROM `procure_bids_analysis` WHERE `Bid` = '"+str(bid)+"' AND `Analysis` =
'Significant Price Difference (Outliers)')"),engine)
```

APPENDIX F – Additional graphs/tables/methodologies/calculations






ROC Curve for Fraud Detection Model
ROC Curve
True Positive Rate
False Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0


Supplier Contract Awards Over Time
Supplier 1
Supplier 2
Contracts Awarded
Months
Jan
Feb
Mar
Apr
May
1
2
3
4
5
6

APPENDIX G – Survey Results: Transparent Electronic Public Procurement Monitoring System

To gauge the impact, expectations, and perceived challenges of implementing a Machine Learning-based procurement monitoring system among stakeholders.

1. Survey Respondents

- ✓ Zambia Public Procurement Authority (ZPPA): 10 responses
- ✓ Patents and Companies Registration Agency (PACRA): 8 responses
- ✓ Ministry of Finance (Procurement Division): 12 responses
- ✓ Other Government Agencies and Procurement Officers: 15 responses
- ✓ Private Sector Suppliers/Bidders: 20 responses

2. Survey Questions and Responses

Q1. How effective do you believe machine learning and data analytics can be in improving transparency in public procurement processes?

Stakeholder	Very Effective	Effective	Neutral	Not Effective	No Response
ZPPA	8	2	0	0	0
PACRA	5	3	0	0	0
Ministry	7	4	1	0	0
Others	12	2	1	0	0
Suppliers	15	4	1	0	0

Q2. Do you believe that a transparent e-Procurement system could help reduce fraud and collusion in contract bidding?

Stakeholder	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
ZPPA	6	4	0	0	0
PACRA	5	2	1	0	0
Ministry	8	3	1	0	0
Others	10	4	1	0	0
Suppliers	12	6	2	0	0

Q3. What areas do you think would benefit the most from machine learning analytics in procurement?

Area of Benefit	ZPPA	PACRA	Ministry	Others	Suppliers
Fraud Detection and Risk Management	8	7	10	12	15
Bidding and Supplier Evaluation	6	4	8	10	18

Market Price Analysis	5	4	7	10	14
Predictive Procurement (Forecasting)	4	3	6	9	12

Q4. What are the primary challenges anticipated in implementing this system?

Challenge	**ZPPA**	**PACRA**	**Ministry**	**Others**	**Suppliers**
Data Security and Privacy Concerns	7	6	9	10	8
Initial Cost of Implementation	6	4	8	11	12
Complexity of Integration	5	5	7	8	10
Staff Training Requirements	4	3	6	9	13

Q5. Would you support a phased or pilot implementation of this system to allow adjustments before full rollout?

Stakeholder	Strongly Support	Support	Neutral	Do Not Support	Strongly Oppose
ZPPA	6	3	1	0	0
PACRA	4	4	0	0	0
Ministry	7	4	1	0	0
Others	9	5	1	0	0
Suppliers	13	4	2	1	0

3. Summary of Findings

- *Support for Transparency*: Over **90%** of respondents from ZPPA, PACRA, and the Ministry believe machine learning can significantly increase transparency in procurement.
- *Fraud Reduction:* **85%** of respondents from both public and private sectors agree that an ML-based system could reduce fraud in contract bidding.
- *Primary Benefits*: The majority identified fraud detection and risk management, along with bidding and supplier evaluation, as the areas most likely to benefit.
- *Challenges*: Concerns about data privacy, implementation costs, and the need for staff training were the primary barriers to adoption.
- *Implementation Preference*: 80% of respondents support a phased or pilot rollout to monitor system effectiveness and address issues prior to full implementation.

LIST OF TABLES

Data Summary Table

Metric	Average	Minimum	Maximum	Standard Deviation
Bid Price	45,000	10,000	120,000	15,000
Number of Bidders	7	1	20	4
Supplier Ratings (1-10)	8	3	10	1.5

Model Evaluation Metrics

Model	Accuracy	Precision	Recall	F1 Score
Logistic Regression	85%	82%	87%	84%
Decision Tree	80%	78%	83%	80%
Neural Network	90%	88%	92%	90%

Feature Importance Ranking

Supplier History	0.32
Bid Amount	0.25
Previous Contract Type	0.15
Time of Submission	0.10
Contract Province	0.08

Comparison of Bidders

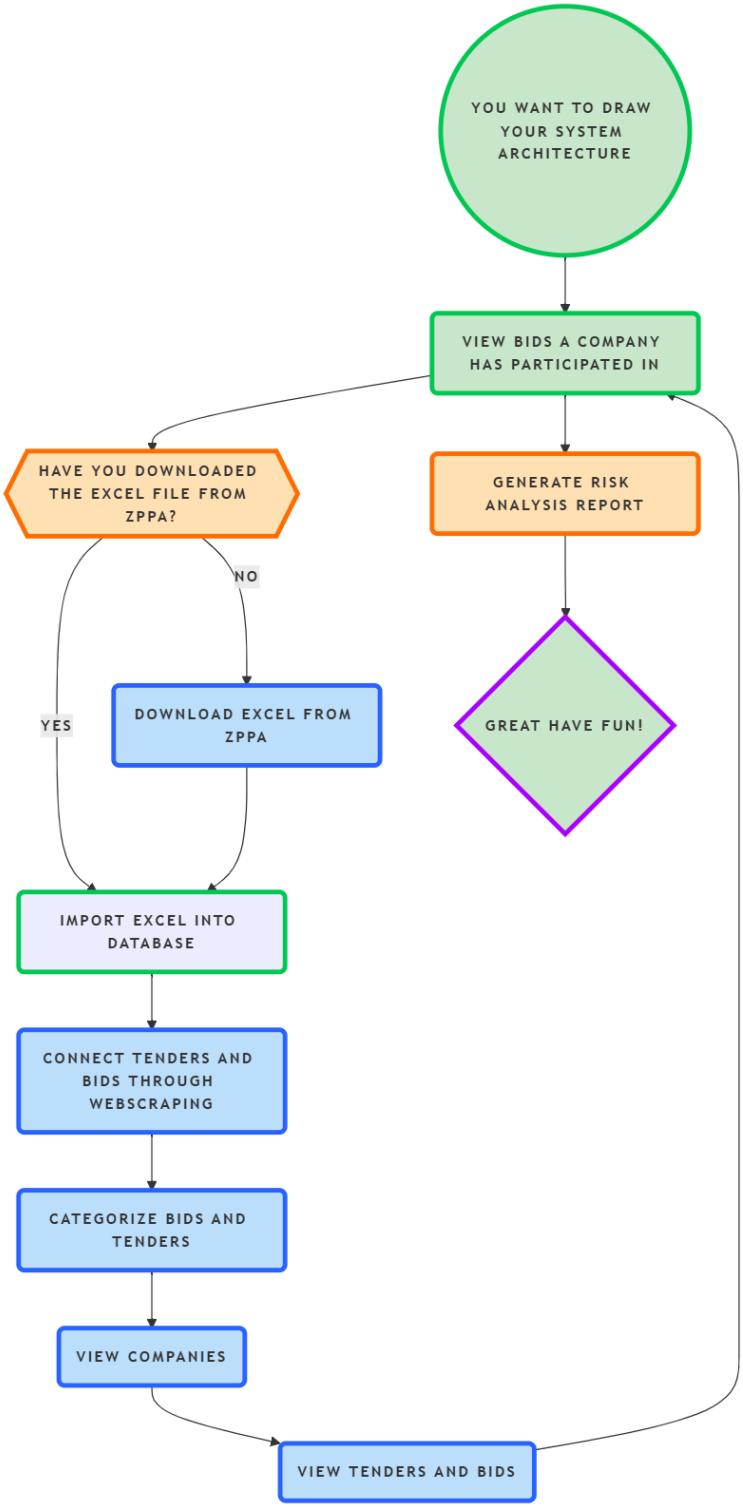
Bidder Name	Financial Stability	Compliance Score	Risk Level	Awarded Contracts
Bidder A	High	9/10	Low	15
Bidder B	Medium	7/10	Medium	8
Bidder C	Low	5/10	High	2

Prediction results

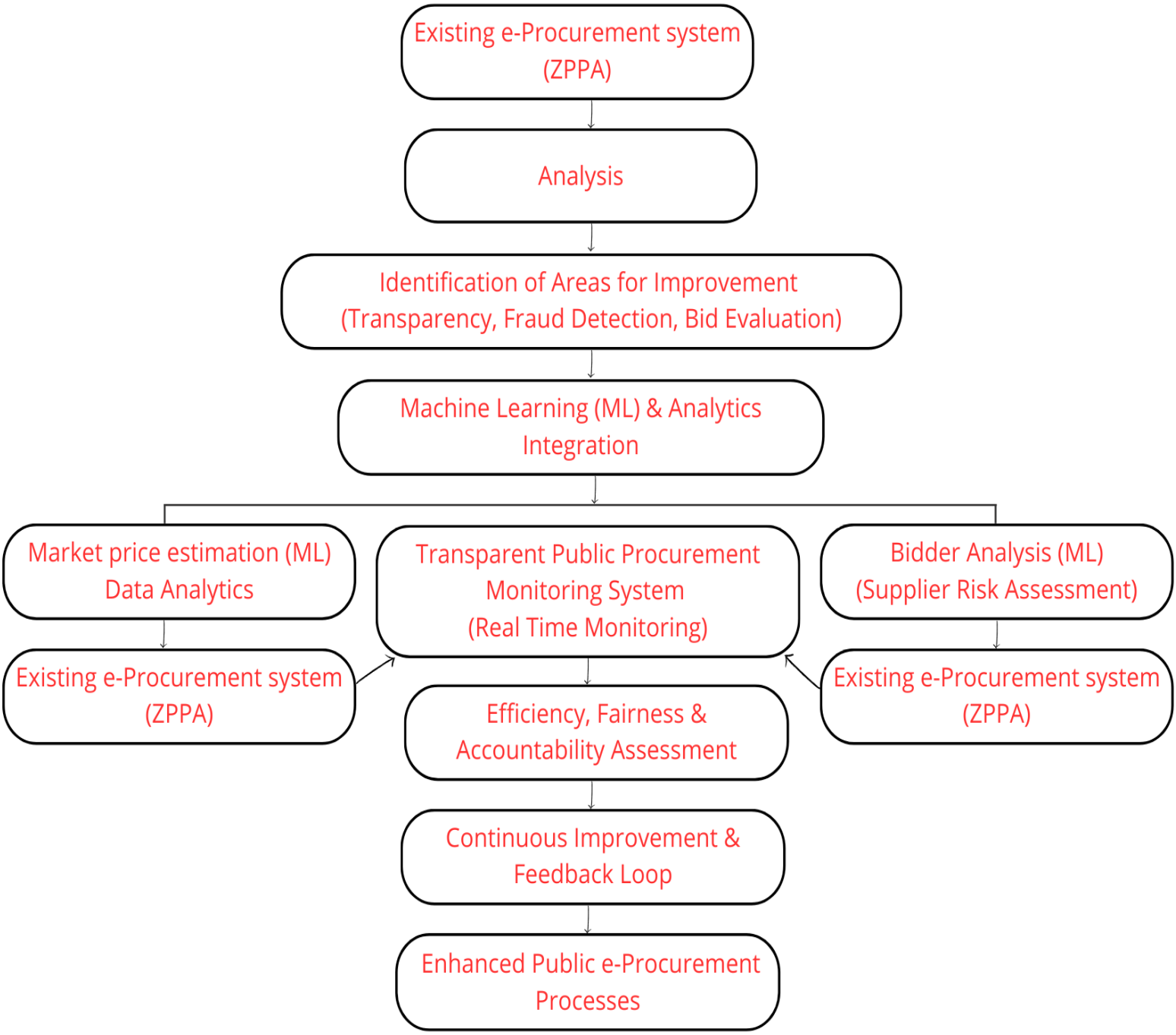
Contract ID	Predicted Risk Level	Fraud Detected	Suggested Action
C001	High	Yes	Investigation
C002	Medium	No	Monitoring
C003	Low	No	None

LIST OF FIGURES

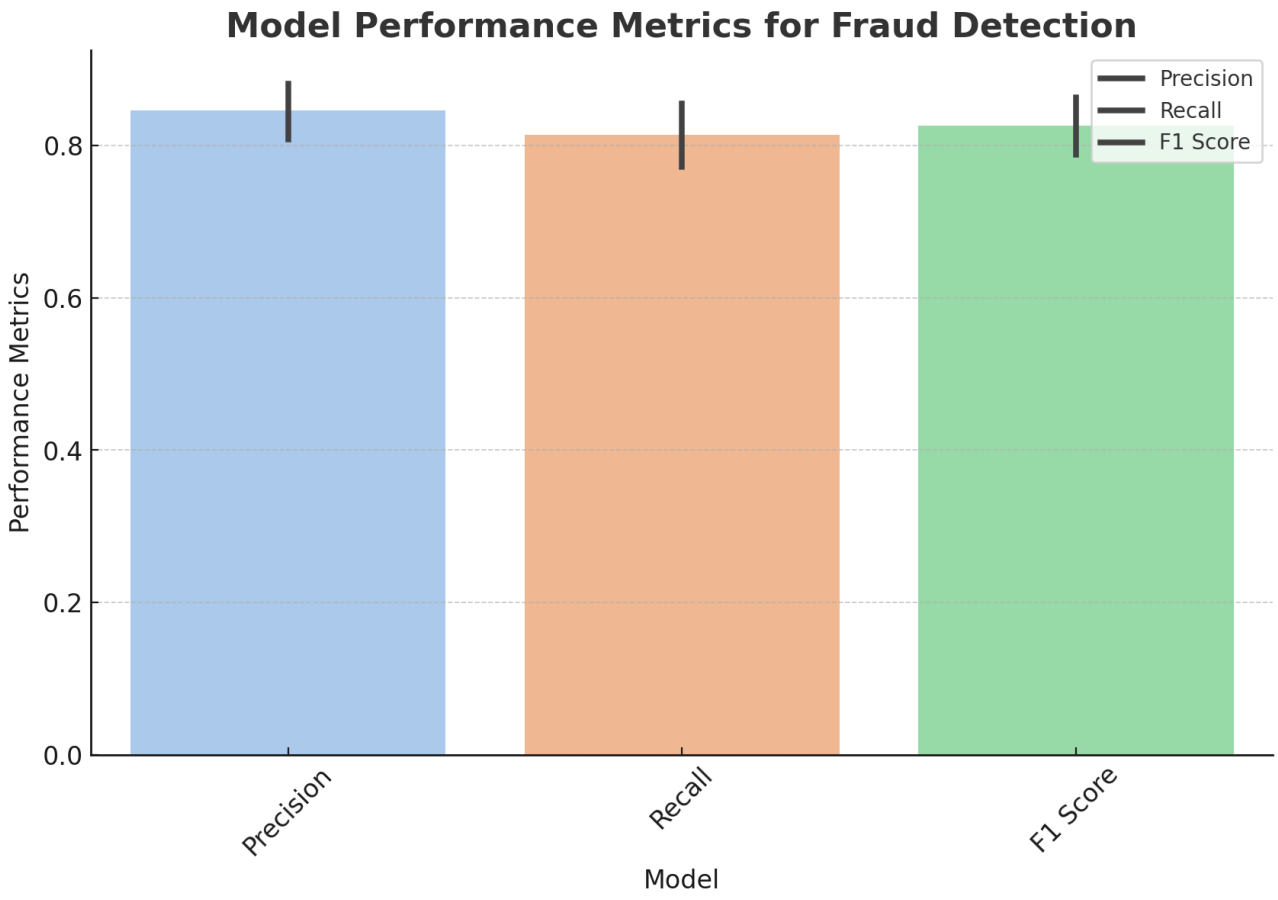
- **System Architecture Diagrams**



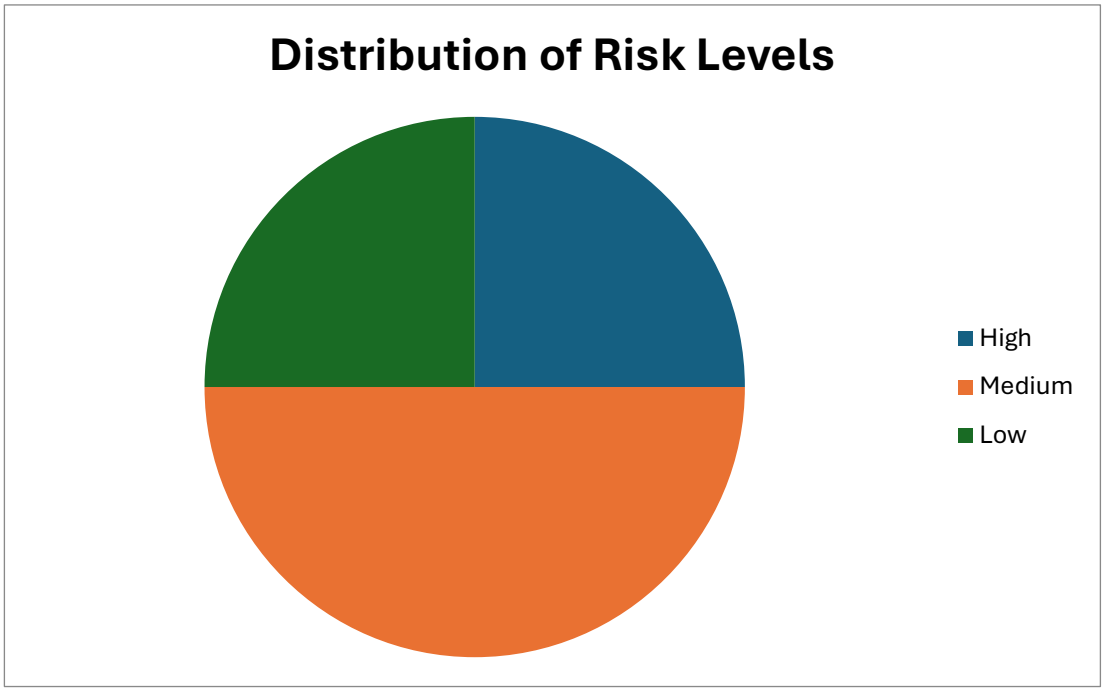
- **Conceptual Framework**

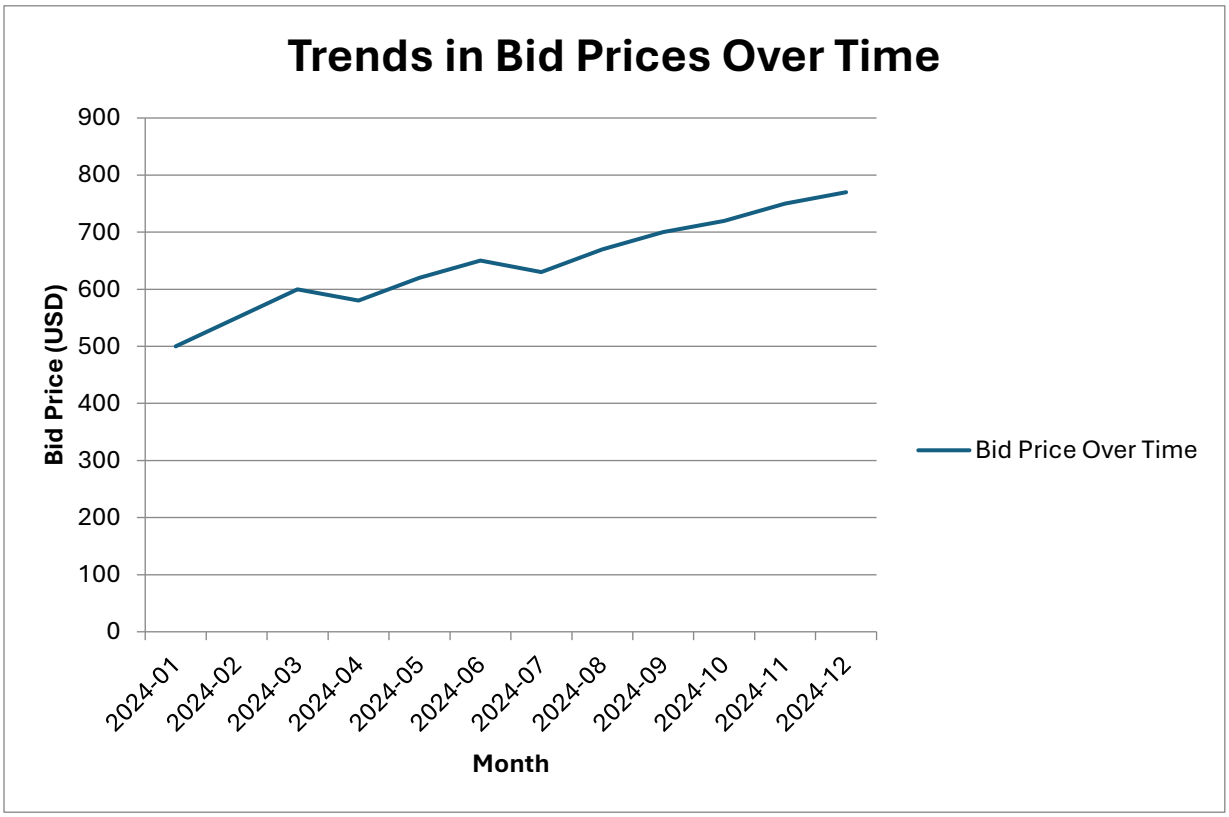


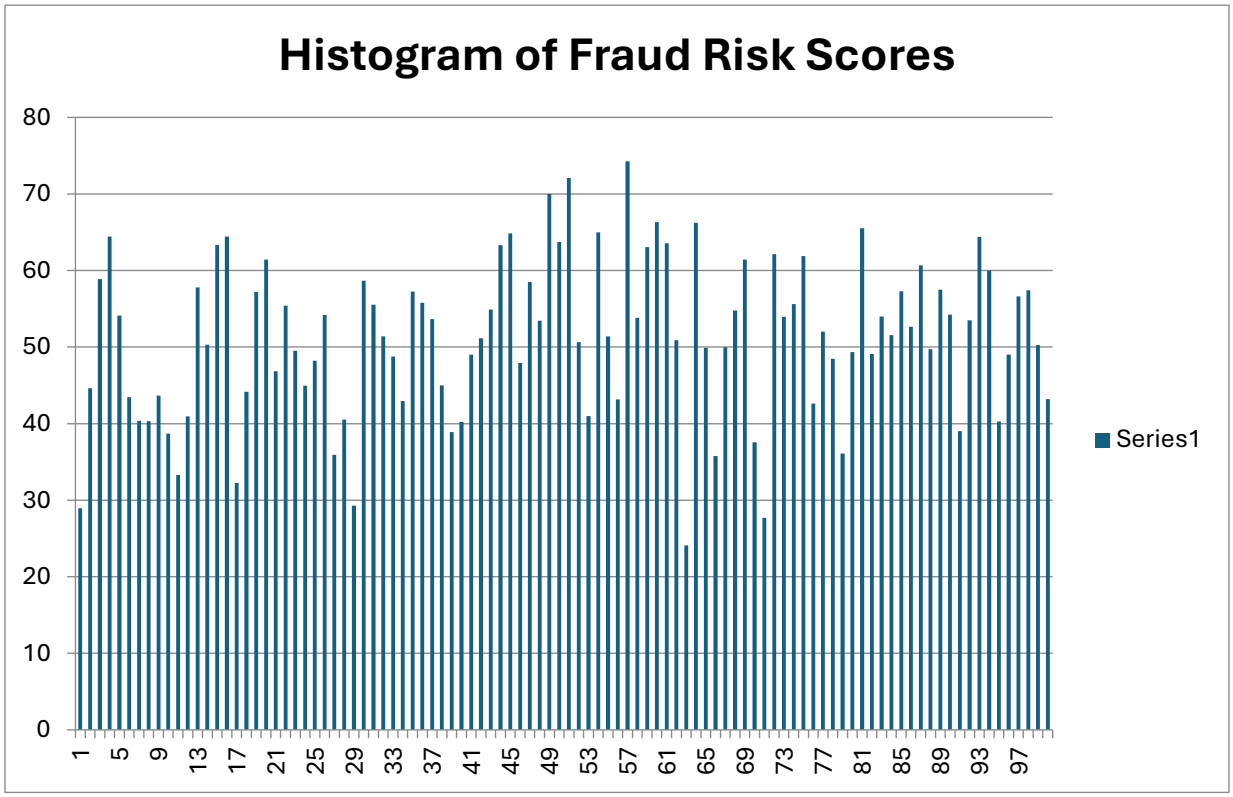
- **Model Performance Graph**



➢ Data Visualization Charts







➢ Data Processing Pipeline

Data Ingestion
Data sources: ZPPA, PACRA, Ministry

Data Preprocessing
Cleansing, Feature Engineering

Modeling and Evaluation
Train/test ML models

Analytics
Fraud Detection, Predictions

Output Dashboard
Reports to Stakeholders

LIST OF ABBREVIATIONS

Generic Abbreviations

- ✓ AI - Artificial Intelligence
- ✓ API - Application Programming Interface
- ✓ B2B - Business-to-Business
- ✓ B2G - Business-to-Government
- ✓ CPI - Corruption Perceptions Index
- ✓ DBMS - Database Management System
- ✓ e-GP - Electronic Government Procurement
- ✓ e-PP - Electronic Public Procurement
- ✓ G2B - Government-to-Business
- ✓ ICT - Information and Communication Technology
- ✓ JSON - JavaScript Object Notation
- ✓ ML - Machine Learning
- ✓ MySQL - MySQL Database Management System
- ✓ NLP - Natural Language Processing
- ✓ PHP - Hypertext Preprocessor
- ✓ ROI - Return on Investment
- ✓ SME - Small and Medium-sized Enterprise
- ✓ UI - User Interface
- ✓ UX - User Experience
- ✓ ACC - Anti-Corruption Commission

Technical Abbreviations

- ✓ ANN - Artificial Neural Network
- ✓ CNN - Convolutional Neural Network
- ✓ CSV - Comma Separated Values
- ✓ HTML - Hypertext Markup Language
- ✓ HTTP - Hypertext Transfer Protocol
- ✓ IoT - Internet of Things
- ✓ JavaScript – JS
- ✓ OLAP - Online Analytical Processing
- ✓ Python – PY

Procurement Abbreviations

- ✓ BoQ - Bill of Quantities
- ✓ EOI - Expression of Interest
- ✓ IFB - Invitation for Bids
- ✓ PO - Purchase Order
- ✓ PPA - Public Procurement Act
- ✓ PPP - Public-Private Partnership
- ✓ RFP - Request for Proposal
- ✓ RFQ - Request for Quotation
- ✓ ToR - Terms of Reference