AI Model for Improving Social Protection Delivery in Zambia

Farai Muzungaile
ICT Unit
Ministry of Community Development and Social Services (MCDSS)
Lusaka, Zambia
Email: faraimuzungaile@gmail.com

Chiyaba Njovu
Open, Distance and eLearning (ODel)
ZCAS University
Lusaka, Zambia
Email: chiyaba.njovu@zcasu.edu.zm

***Abstract*—Social Protection programmes in Zambia aim to reduce the levels of poverty and improve the lives of the vulnerable in the communities. However, the two most common challenges in the Social Protection sector in Zambia are fragmented data and poor targeting of beneficiaries. This report describes the development of an AI model prototype for improving Social Protection delivery in Zambia. AI models such as machine learning are discussed in helping solve the fragmented data problem through integration of data from different management information systems as well as improve the targeting challenge using algorithms. The legal, technological and ethical issues in relation to the prototype being developed are also discussed. Furthermore, the report discussed the future works on the enhancements of the model and its functionality.**

Keywords—artificial intelligence, social protection, machine learning, targeting

I. INTRODUCTION

Social Protection in Zambia consists of programmes which aim to reduce the levels of poverty and improve the lives of the vulnerable in the communities. Apart from helping poor and vulnerable families from being stagnant in poverty, Social Protection has greatly contributed to the economic growth by raising labour productivity and enhancing social stability [1].

The Ministry of Community Development and Social Services (MCDSS) is mandated to coordinate the provision of equitable social protection services to communities in order to contribute to sustainable human development [2]. The two main departments at MCDSS that deal with these social protection programmes are the Department of Community Development and the Department of Social Welfare.

The Department of Community Development is mandated to enhance the livelihoods of the poor and vulnerable but viable people, through collective efforts to improve the social, economic and cultural conditions of the communities for sustainability, poverty reduction and national development. Programmes implemented under this department include Food Security Pack (FSP), Supporting Women's Livelihood (SWL), Village Banking, among others [3].

On the other hand, the Department of Social Welfare is responsible for the provision and promotion of quality social welfare services aimed at alleviating poverty, reducing destitution, promoting family values and reducing juvenile delinquency. Programmes implemented under this department include Social Cash Transfer (SCT), Emergency Cash Transfer (ECT), Public Welfare Assistance Scheme (PWAS), Juvenile Welfare, Ending Child Marriage, among others [4].

Statistics provided by MCDSS reveal that, as of 2024, there were 1,311,101 beneficiaries under the SCT programme, 952,570 beneficiaries under ECT, 263,700 beneficiaries under the FSP programme and 139,351 beneficiaries under the SWL programme [5]. These beneficiaries are cut across all the 116 districts in Zambia.

Despite the attempts to enhance Social Protection delivery in Zambia, the country faces lack of integration leading to fragmented data systems and poor targeting of beneficiaries. In order to tackle these existing challenges, the application of Artificial Intelligence (AI) in Social Protection will help to provide solutions via AI models such as machine learning to improve efficiency in the Social Protection sector in Zambia.

II. LITERATURE REVIEW

A. Social Protection

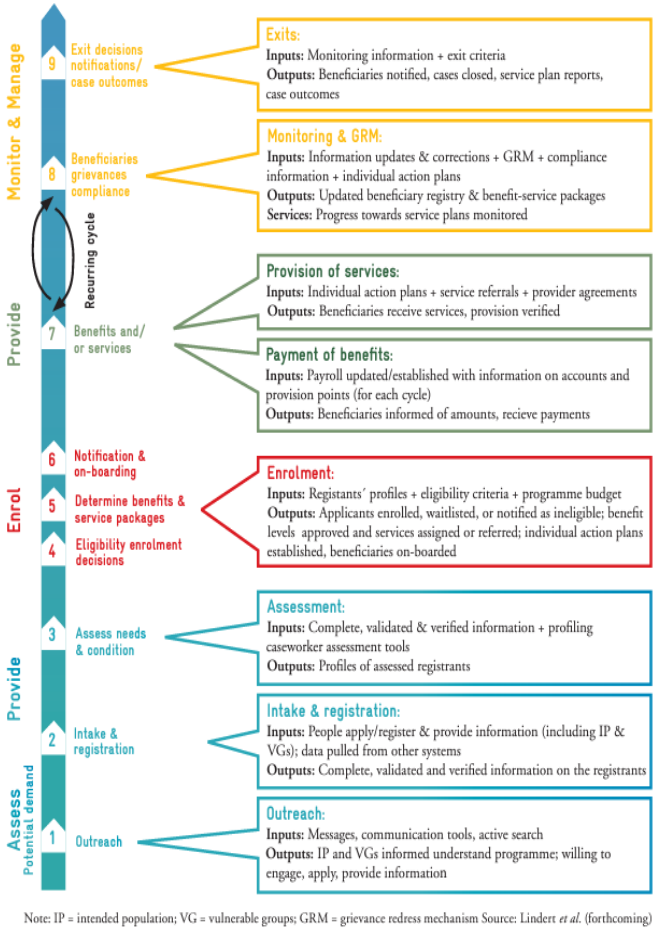
It is a set of programmes and policies by government that provide income or consumption transfers to the poor, protect the vulnerable against livelihood risks and enhance the social status and rights of the excluded; with the overall objective of reducing the economic and social vulnerability of poor and excluded groups.

Social protection can be both contributory and non-contributory. Contributory schemes are defined as plans that require both the employee and their employer to contribute a portion of their income in order to claim their benefits in the future. Non-contributory schemes are those that normally need no direct financial contribution from beneficiaries or their employers as a condition of entitlement to receive benefits. In Zambia, an example of a contributory scheme is the National Pension Scheme Authority (NAPSA) which provides social insurance for formal sector employees in the event of retirement, disability, death of a spouse or parent. Examples of non-contributory schemes are:

- **Social Cash Transfer (SCT)** - This programme is implemented by MCDSS through the Department of Social Welfare. The objective of the programme is to provide financial support to the poor and vulnerable households in order for them to meet their basic needs, such as health, education, food and shelter. This programme is being implemented in all the 116 districts of Zambia [6].

 ISBN: 978-9982-95-500-3

- **Food Security Pack (FSP) -** This programme is implemented by MCDSS through the Department of Community Development. The objective of the programme is to empower the poor and vulnerable but viable farmer households with agricultural inputs and livelihood skills to improve their productivity with the ultimate goal of enhancing their food, nutrition and income security for self-sustainability and poverty reduction. FSP has three components namely; Rainfed, Wetland and Alternative Livelihood Initiative (ALI). These components are implemented in all the 116 districts of Zambia [7].
- **Supporting Women's Livelihoods (SWL)** - This programme is implemented by MCDSS through the Department of Community Development. SWL is a component of the Girl's Education and Women's Empowerment and Livelihoods (GEWEL) project. The objective of SWL is to provide extremely poor women aged between 19 to 64 years with opportunities to increase the productivity of their livelihoods, and their economic empowerment, through training, mentoring, peer support, productivity grants and help with setting up savings clubs. This programme is implemented in 81 districts of Zambia [8]. Fig 1 illustrates the Social Protection delivery chain [9].



Note: IP = intended population; VG = vulnerable groups; GRM = grievance redress mechanism Source: Lindert *et al.* (forthcoming)

Fig. 1. Social Protection Delivery chain

a. Artificial Intelligence (A.I)

Artificial Intelligence (AI) is a technology that concentrates on building machines capability of executing tasks that are typically thought to require human intelligence. It is able to interpret and sort data at scale, solve complex problems and automate various tasks concurrently, which can save time and fill in operational gaps overlooked by humans [10].

B. Types of A.I Models

Machine learning (ML) is a subset of AI that uses algorithms trained on datasets to create self-learning models that are competent at predicting outcomes and organizing information without human intervention. A machine learning algorithm uses statistical techniques to help it "learn" and progressively better at a task, without necessarily having been programmed for that certain task. It uses historical data as input to predict new output values. There are four types of machine learning namely [11];

A. **Supervised Machine Learning -** It uses algorithms that are trained on labelled datasets whereby each example has an input and a corresponding output to make predictions on new, unseen data.

B. **Unsupervised Machine Learning** - It uses algorithms that are trained on unlabelled datasets to identify hidden patterns with no guidance.

C. **Semi-Supervised Machine Learning -** It uses algorithms that are trained on both labelled and unlabelled datasets. This technique operates by feeding algorithms a small quantity of labelled data to assist direct their development and then feeds much larger quantities of unlabelled data to complete the algorithm.

D. **Reinforcement Learning -** It It uses algorithms that are trained on trial and error basis within an interactive environment. During this process, the algorithms receive feedback in the form of rewards or penalties as they perform actions allowing them to learn and improve performance over time uses algorithms that are trained on unlabelled datasets to identify hidden patterns with no guidance.

The benefits of AI that can be applied in the Social Protection sector.

E. **Fraud Detection -** Machine learning and deep learning algorithms can analyse transaction patterns and flag anomalies, such as paying beneficiaries who are deceased [12].

F. **Automation -** AI can improve the integration process using machine learning algorithms that can automate the process of cleaning data, alleviating data inconsistences or discrepancies between the various Social Protection programmes MISs [13].

G. **Improved Targeting of Beneficiaries -** Machine Learning algorithms can improve the targeting

 ISBN: 978-9982-95-500-3

process of beneficiaries by using big data (age, gender, geographic location, etc.) to predict beneficiaries who are eligible or non-eligible [14].

III. METHODOLOGY

The adopted methodology for this research was the Agile Methodology. The Agile Methodology is associated with the project in the following ways [15,16]:

[1] **Collaborative Teamwork:** As mentioned earlier in the Management Information Systems (MISs) at MCDSS operate in silos leading to fragmented Social Protection data. This methodology encouraged teamwork by breaking down silos allowing the departments at the Ministry to collaborate with each other for the integration task to be achieved by the proposed AI model.

[2] **Adaptability:** Agile methods are adaptable and they can enable AI models to alter strategies quickly, without disrupting the design of the model.

[3] **Iterative Nature:** Iterative nature of development means that you can go back to a previous phase in the project and change something that was not considered carefully in that particular phase. The Agile Methodology is Iterative, which suits well for Supervised Machine Learning as iterations can improve its performance on targeting eligible beneficiaries as new data or unseen data emerges.

A. Reseearch Data and Datasets

The research data consisted of three datasets with beneficiary data for Chongwe district under MCDSS. The first dataset contained information on Social Cash Transfer (SCT) beneficiaries. The second dataset contained information on Supporting Women's Livelihood (SWL) beneficiaries. The third dataset contained information on Food Security Pack (FSP) beneficiaries.

B. Data Collection Methods

The data collection techniques that were used for this research were quantitative data. The data was obtained from the three Zambian Government Management Information Systems (MISs) under the Ministry of Community Development and Social Services (MCDSS that deal with Social Protection Programmes:

[4] **The Zambian Integrated Social Protection Information System (ZISPIS)** - deals with beneficiaries who are on Social Cash Transfer (SCT),

[5] **Supporting Women's Livelihood Management Information System (SWLMIS)** - deals with beneficiaries who receive grants on the SWL

[6] **Food Security Pack Management Information System (FSPMIS)** - deals with beneficiaries who receive agricultural inputs such as fertilizer and seed on the FSP programme.

The three datasets were downloaded from the three MISs and appeared as Excel sheets in .csv format.

C. Data Analysis Techniques

The data analysis techniques include the following [17]:

[7] **Data Pre-processing** - This technique entailed cleaning and pre-processing the collected data in order to make it suitable for the AI model to undergo training. The cleaning of data involved tasks such as handling missing values by eliminating them. The data was normalized by formatting it into a single dataset to be used for training and testing the algorithm. The name of the normalized dataset was **Beneficiaries.csv**

[8] **Model Training and Testing** - Supervised machine learning models such as Classification algorithms were trained by analysing the input data and learning to map it to the correct output labels. The algorithms were trained with the labelled dataset and the targeting of beneficiaries were classified as "Eligible" or "Not Eligible" based on the integrated data from the government MISs.

[9] **Model Evaluation** - The AI model was evaluated by performance metrics such as accuracy, precision, recall, F1-score to vizualize how well it performed on the dataset. of beneficiaries were classified as "Eligible" or "Not Eligible" based on the integrated data from the government MISs.

D. Ethical Concerns

The ethical issues that usually arise in Social Protection systems involve the violation of potential beneficiaries' personal, health, financial and social records, algorithmic bias and discrimination and the impact on employment.

Different departments under the Ministry have different Social Protection data in their systems. AI models can utilize large volumes of social protection data from these departments which would therefore raise privacy and security concerns. These security concerns involve the risk of data being misused, breached from unknown sources. Therefore, the data being collected should also abide by the Data Protection Act of 2021.

Another vital ethical issue to put into consideration is Bias and Discrimination. In this case, AI algorithms are trained on data and the downside to this, is that if the data is biased, then the AI algorithm will be biased too. This could lead to AI algorithms making unsatisfactory or discriminatory decisions towards the targeting of beneficiaries, which would lead to exclusion of beneficiaries who are meant to receive the benefits [18].

The impact on employment is a third ethical issue that needs to be critically addressed. The current rate of unemployment in Zambia is very high and since Machine Learning is able to automate the tasks being conducted by community and social workers, this might cause worries about job losses in the Social Protection sector [18].

IV. IMPLEMENTATION

Table 1 below shows the current challenges of the existing Social Protection delivery systems in Zambia.

 ISBN: 978-9982-95-500-3

MIS	Modules	Challenges
ZISPIS	•Beneficiary Registration •Targeting •Households •Payment •Grievance Redress Mechanism •Reports	•Missing values of NRCs for beneficiaries. •No integration with SWLMIS and FSPMIS.
SWLMIS	•Beneficiary Enrolment •Targeting •Savings •Training •Payment •Reports	•Missing values of NRCs for beneficiaries. •No integration with ZISPIS and FSPMIS.
FSPMIS	•Beneficiaries •Food Security Pack •Set-Up (Administration)	•Missing values of NRCs for beneficiaries. •No integration with SWLMIS and ZISPIS.

Table 1. Comparison of existing Social Protection delivery systems

The proposed model employed **"Google Colab"**, a cloud-based Jupyter Notebook environment that allows for writing and executing Python code. This platform was used to train, test, and visualize the AI model. The Classification algorithm "Logistic Regression" was used for the implementation of the AI model. Fig 2 below shows the design of the prototype.

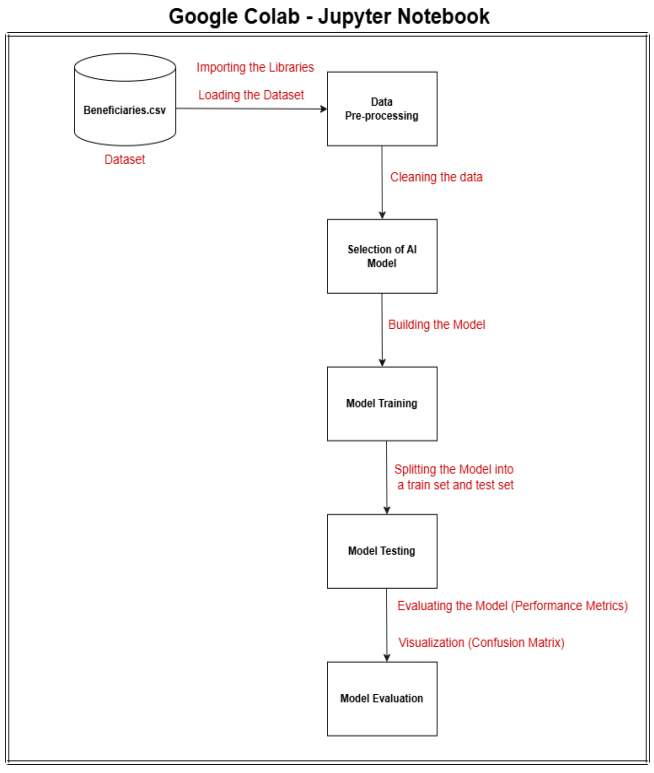


Fig 2. Prototype Design

V. RESULTS

The developed AI model was trained and tested by splitting the dataset into 75% for training and 25% for testing. Predictions were made on the test set to predict beneficiaries that are "Eligible" and those that are "Not Eligible.

A **Confusion Matrix** was used to visualize model performance by showing the number of correct and incorrect predictions made by the model. Correct predictions are inclusive of true positives and true negatives whereas incorrect predictions are inclusive of false positives and false negatives [19]. With reference to the AI Model for Social Protection, the predictions can be identified as follows:

1. **True Positive (TP)** – A beneficiary who is actually eligible (positive) and is correctly classified by the model as eligible.
2. **True Negative (TN)** – A beneficiary who is actually not eligible (negative) and is correctly classified by the model as not eligible.
3. **False Positive (FP)** – A beneficiary who is actually not eligible (negative) but is incorrectly classified by the model as eligible.
4. **False Negative (FN)** – A beneficiary who is actually eligible (positive) but is incorrectly classified by the model as not eligible.

A. Confusion Matrix

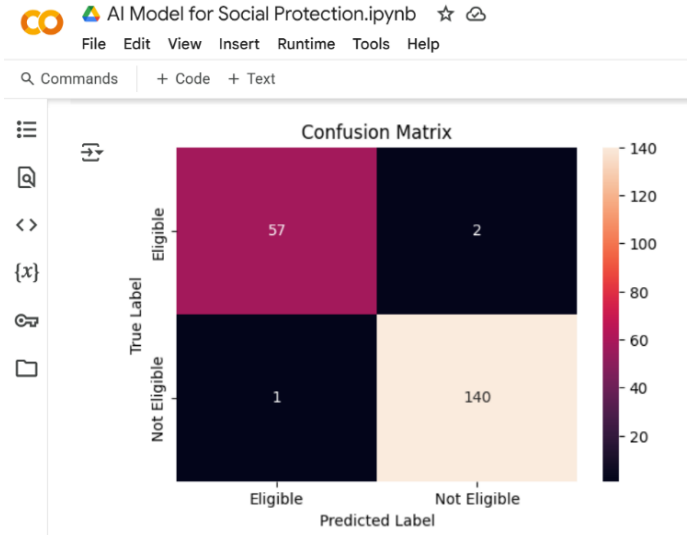


Fig 3. Confusion Matrix

57 beneficiaries were correctly classified as Eligible **(true positives).**

140 beneficiaries were correctly classified as Not Eligible **(true negatives)**.

1 beneficiary was incorrectly classified as Eligible **(false positive).**

2 beneficiaries were incorrectly classified as Not Eligible **(false negatives).**

 ISBN: 978-9982-95-500-3

B. *Performance Metrics*

The model was highly accurate at 98.5%, making it reliable and therefore this improves targeting of beneficiaries and reduces inclusion and exclusion errors.

The model had a precision of 98.6%; this means that when the model predicts the targeting of eligible beneficiaries, it is correct 98.6% of the time.

The model had a recall of 99.2%; this implies that the model captured a lot of eligible beneficiaries with only a shortfall of 0.8% of beneficiaries who are not eligible.

The model had a f1-score of 98.9%; which means that there was some strong balance between precision and recall.

VI. CONCLUSION

The system that hosts the National Registration Card (NRC) database in Zambia and is essential for the identification and authentication of beneficiaries, is the Integrated National Registration Information System (INRIS). It is managed by the Ministry of Home Affairs and Internal Security (MOHAIS). The INRIS can be used to validate the NRC details of the beneficiaries to ensure there are no errors, incomplete NRC numbers, inclusion of people that do not exist, validation of eligible beneficiaries, to ensure the right targeted beneficiaries are the ones that receive their benefits and reduce fraud [20]. According to [21], the World Bank recently granted Zambia an amount of $100 to digitize National IDs (NRCs). This financial support can help MOHAIS expedite the validation process in all the 116 districts in the country.

There are future works to scale up the SWL social protection programme under MCDSS from 81 districts to 116 districts which simply means that more beneficiaries will be added to the programme. Once scaled up, the dataset can be expanded for the AI model and INRIS can adopt Supervised Machine Learning techniques to integrate the government databases such as the Civil Registry, which contains the births, deaths, and marriages' records as well as the three Management Information Systems at MCDSS. The AI model can then be retrained to accommodate this newly integrated data. This would prevent fragmentation of Social Protection data as well as avoid paying beneficiaries who are deceased and henceforth improve Social Protection delivery in the long run [22].

ACKNOWLDEDGMENT

The authors would like to take this opportunity to thank the Almighty God for blessing them with the gift of life and wisdom in their ICT careers.

They would also like to express their gratitude to ZCAS University for providing resources to present the paper.

REFERENCES

[1] Zulu, B. (2021) The Status of Social Protection in Zambia
Available at: https://pmrczambia.com/wp-content/uploads/2021/07/OP-ED-The-Status-of-Social-Protection-in-Zambia.pdf (Accessed: 30 October 2024)

[2] About MCDSS - Vision, Mission and Values (2021)
Available at: https://www.mcdss.gov.zm/?page_id=4223 (Accessed: 30 October 2024)

[3] Department of Community Development (2021)
Available at: https://www.mcdss.gov.zm/?page_id=3228 (Accessed: 30 October 2024)

[4] Department of Social Welfare (2021)
Available at: https://www.mcdss.gov.zm/?page_id=1808 (Accessed: 30 October 2024)

[5] MCDSS Programme Data (2024)
Available at: https://www.mcdss.gov.zm/ (Accessed: 1 November 2024)

[6] Social Cash Transfer (2021)
Available at: https://www.mcdss.gov.zm/?page_id=2086 (Accessed: 16 November 2024)

[7] Food Security Pack (2021)
Available at: https://www.mcdss.gov.zm/?page_id=2936 (Accessed: 16 November 2024)

[8] Supporting Women's Livelihood. (2021)
Available at: https://www.mcdss.gov.zm/?page_id=3465 (Accessed: 16 November 2024)

[9] Ohlenburg, T. (2020) AI in Social Protection – Exploring Opportunities and Mitigating Risks Available at: https://socialprotection.org/discover/publications/ai-social-protection (Accessed: 14 November 2024)

[10] Glover, E. (2024) What Is Artificial Intelligence (AI)?
Available at: https://builtin.com/artificial-intelligence (Accessed: 19 November 2024)

[11] Coursera Staff (2024) What Is Machine Learning? Definition, Types, and Examples
Available at: https://www.coursera.org/articles/what-is-machine-learning (Accessed: 19 November 2024)

[12] Adams, D. (2024) 'Leveraging technology and data advances to improve social programme coverage and service delivery', Modernising Access to Social Protection, pp. 63-95
Available at: https://www.oecd-ilibrary.org/social-issues-migration-health/modernising-access-to-social-protection_e2bd6fd8-en (Accessed: 7 November 2024)

[13] Verhagen, A. (2024) 'Using AI to manage minimum income benefits and unemployment assistance: Opportunities, Risks and Possible Policy Directions', OECD Artificial Intelligence Papers, 21, pp. 1-54
Available at: https://www.oecd-ilibrary.org/content/paper/718c93a1-en (Accessed: 8 November 2024)

[14] Stryker, C., & Kavlakoglu, E. (2024) What Is AI?
Available at: https://www.ibm.com/topics/artificial-intelligence (Accessed: 19 November 2024)

[15] Laoyan, S. (2024). What Is Agile Methodology (A beginner's guide)?
Available at: https://asana.com/resources/agile-methodology (Accessed: 5 December 2024)

[16] Nagl, S. (2023). 5 Benefits of Agile Project Management.
Available at: https://sps.wfu.edu/articles/benefits-agile-project-management/ (Accessed: 5 December 2024)

[17] Kamiri, J., & Mariga, G. (2021) 'Research Methods in Machine Learning: A Content Analysis', International Journal of Computer and Information Technology (IJCIT), 10(2), pp. 78-91

Available at: https://www.ijcit.com/index.php/ijcit/article/view/79 (Accessed: 19 December 2024)

[18] Lowe, C. et al. (2023) 'Pathways toward digitalization in Social Protection and Labor (SPL) service delivery', Social Protection & Jobs, 2307, pp. 1-55.
Available at: https://documents.worldbank.org/curated/en/099100523151038686/pdf/P173530036a7f30f60899c06a6172040bf3.pdf (Accessed: 19 December 2024)

[19] Harikrishnan, N.B. (2019) Confusion Matrix, Accuracy, Precision, Recall, F1 Score
Available at: https://medium.com/analytics-vidhya/confusion-matrix-accuracy-precision-recall-f1-score-ade299cf63cd (Accessed: 16 March 2025)

[20] Barca, V., & Chirchir, R. (2019) Building an integrated and digital social protection information system
Available at: https://www.giz.de/en/downloads/giz2019-en-integrated-digital-social-protection-information-system.pdf (Accessed: 2 May 2025)

[21] Jere, J. (2025) World Bank Grants Zambia $100m to Digitize National ID's
Available at: https://znbc.co.zm/news/world-bank-grants-zambia-100m-to-digitize-national-ids/ (Accessed: 4 May 2025)

[22] Mwangwewo, L. (2023) The INRIS Project
Available at: https://www.mohais.gov.zm/?p=2637 (Accessed: 5 May 2025)

 ISBN: 978-9982-95-500-3