



Artificial Intelligence in analysis of qualitative data from programme monitoring of human development programmes – Readiness criteria for application in low-income countries

Dr. Boniface Francis Kalanda

Introduction :

Development programmes in low-income settings attract a lot of funding from both bilateral and multilateral donors. It is estimated that in 2020, total foreign aid was \$204 billion by members of the Development Assistance Committee (OECD, 2023). While these resources are used to provide and buy services, a substantial amount of it is also used for programme monitoring. Such monitoring generate a lot of data. Most of it is qualitative data from service point such as hospitals and schools and from community level interventions such as localized advocacy efforts. These programme monitoring activities cost a lot of money. There are no known studies to estimate how much of global overseas development assistance (ODA) is used for monitoring. A rough estimate showed that one monitoring visit costs in excess of \$500 per monitoring visit to a community (Kalanda, 2021, unpublished)¹. Such data is not routinely analyzed, in the process losing opportunities for programme course corrections. Application of artificial intelligence (AI)² offers an opportunity for faster, smarter analysis of huge qualitative data sets of human development data to facilitate programme effectiveness. This paper explores readiness criteria for low-income countries, where most development aid lands, to exploit the potential of AI for analysis of qualitative data generated from monitoring of human development programmes.

Documented use of AI in analysis of qualitative data in human development programmes :

As a classic example of how a human development agency can use AI, the Independent Evaluation Office of UNDP is exploring the use of AI and machine learning to analyze evaluation reports. The main aim is to leverage this rich unstructured data, from Evaluation Reports, to generate insights, relationships and trends that might not otherwise be readily apparent and to visualize this information more quickly (Market Trends, 2022).

Since 2022, the United Nations Environmental Programme (UNEP)'s World Environment Situation Room (WESR) has employed AI to “curate, aggregate and visualize the best available earth observation and sensor data to inform near real-time analysis and future predictions on multiple factors, including CO₂ atmospheric concentration, changes in glacier mass and sea level rise” (UNEP, 2022). WESR is a demand driven data platform that is increasingly delivering data for decision making into the hands of users.

The United Nations Development Programme (UNDP) launched Artificial Intelligence for Development Analytics (AIDA), a platform that leverages AI and big data to provide insights and solutions for sustainable development. For example, In Bangladesh, UNDP and a2i partnered with a non-profit, GiveDirectly, to direct aid to those most in need. It uses AI tools to extract mobile phone and satellite metadata to identify vulnerable populations who may be eligible for cash transfers in a way cheaper (Market Trends, 2022)

The Text and Data Analytics unit at the World Bank is also leading in advancing frontiers of knowledge in this area. The Text and Data Analytics Unit at WB provides support by automatically extracting knowledge and insights from the wealth of documents and data inside and outside the World Bank (World Bank, 2018) for analysis using AI technology.

Siiman et al (2022) have reported that AI assisted qualitative analysis has the potential to improve transparency in the coding of qualitative data through assisting humans to report AI prompts that agree with their interpretations of the data, in the process, enabling re-use by other researchers. Again, in qualitative analysis of data, AI is able to identify categories amenable to automated coding, and to rapidly conduct that coding on the entirety of very large datasets. This saves time and money. It also supports researcher to avoids limitations inherent in limiting qualitative analysis to limited samples of a given dataset (Lennon, et al, 2022). Buluswar et al (2018) have summarized other areas where AI has been used for analysis such as predicting food prices from Twitter (now X) and analysis of demographics. Franzen et al (2022) also explores the question of using AI for advanced content analysis of data from programme monitoring of human development programmes with very promising results, although others, such as Others however, such as

¹ This is an estimate from a complex/humanitarian environment. The cost could be lower in normal development settings

² Artificial intelligence (AI) refers to the simulation of intellectual human behavior by computers. AI can be designed using algorithms related to machine learning (ML), deep learning (DL), natural language processing (NLP), support vector machine (SVM), and artificial neural network (ANN)

Mahto et al (2020) have emphasized that IA does better with human support.

Limitations of AI :

Like all technologies, AI has limitations. On the other hand, Halmilton et al (2023) have outlined a number of limitations for AI such limited contextual understanding, lack of common-sense knowledge, limited memory, and inability to generate original ideas amongst others. In addition, AI has the potential to pose challenges in analysis of data (Smith & Neupane, 2018) including mis information as was the famous case in the USA elections of 2016. Others have also warned against limitations of AI such ambiguity surrounding copyright, the lack of truthfulness and accuracy, and the increased potential for misuse and bias (Emenike & Emenike,2023).

Application of AI for analysis of human development programme monitoring data :

There is potential for applying AI for analysis of human development monitoring data in low-income countries. Application of AI for analysis in such settings will be driven by (i) Data Governance Structures; (ii) intentionality; (iii) understanding what AI is and is not; (iv) investing in public data infrastructure; (v) data science skills; (vi) identifying appropriate solution spaces and (vii) risk management amongst others.

- **Intentionality (including Data Governance³):** AI will be used in low-income settings when donors and implementing partners, led by governments, UN agencies, national and international NGOs are intentional in using AI for analysis. Development programmes generate a lot of data through programme monitoring. The present scenario, where such data is not analyzed constitutes loss of value – value which could have been realized from analysis leading to course correction. A good example of “intentionality” come from UNICEF which has developed AI driven tools to support programme analysis (UNICEF, 2022)⁴ and UNDP (2022) which is using AI to analyze evaluation reports. Part of “Intentionality” will be setting up Data Governance Structures⁵. WDR (2022) estimates that “*No low-income countries have a data governance entity, compared to 53 percent of high-income countries*”. Establishing Data Governance Structures will be one of the risk mitigation measures that will need to be taken by lo-income countries prior to using AI for analysis of qualitative data from human development programmes.
- **Bridging the know gaps between AI experts and human development experts:** AI can only appropriately be used if it is fully understood. There is a knowledge divide between experts in AI and human development experts. Bridging such a gap, through the later group reaching to the former will be the starting point for appropriately applying AI in the appropriate “solution spaces.”
- **Appropriate solution spaces.** Since the “AI for Good Global Summit Report” (2017), which stated that AI will be “*central to the achievement of the Sustainable Development Goals and could help solve humanity’s grand challenges*”, there has been optimism in the development sector to employ AI. However not all “spaces” are amenable to AI. Identifying appropriate solution spaces in data analysis in the human development sector is therefore a starting point in application of AI.
- **Public data infrastructure⁶:** AI can be used for analysis only if the correct infrastructure is available. Most low-income countries have “deficient” Data Density Index⁷ (Buluswar et. al 2018). This deficiency is not conducive to data analysis using AI. Donors and governments should progressively invest in “data” infrastructure if AI driven analysis is to become a reality. Appropriate infrastructure is of utmost importance as a risk mitigation measure. Poor infrastructure would lead to abuse of data, risking the very development data is supposed to support.
- **Data Scientists:** In low-income settings, even when AI matures for analysis of human development data, the lack of data scientists can make its use extremely limited (Chui & Harrysson, 2019). To exploit AI for analysis, low-income countries need to invest in analysis skills in its data scientists. Low-income countries will need nationals with experience in using deep learning algorithms, using massive amounts of data to model overly complex patterns if AI is to be used for analysis (Buluswar et al., 2018).
- **Risk driven:** AI has very well-known risks. Geoffrey Hinton, the God Father of IA has been quoted as saying “*These things could get*

³ Data Governance constitutes control of risks by ensuring security, integrity, and protection of data and systems, and to capture value by establishing rules and technical standards to enable data to be more effectively transferred, combined, and exchange (WDR, 2022)

⁴ PIDB Recommender and the End of Year Summary Narrative (EYSN) Quality Assurance Tool – These are tools for analysis of programmes codes and analysis of annual programme reports.

⁵ Data Governance constitutes control of risks by ensuring security, integrity, and protection of data and systems, and to capture value by establishing rules and technical standards to enable data to be more effectively transferred, combined, and exchange (WDR, 2022)

⁶ Data infrastructure refers to the various components—including hardware, software, networking, services, policies, and more—that enable data consumption, storage, and sharing. Having the right data infrastructure strategy is critical for organizations seeking to undertake data-driven digital transformation (Hewlett Packard Enterprise)

⁷ Data Density Index (DDI) is a measure of the strength of a country’s data infrastructure and its readiness to utilize data to advance its development objectives (Buluswar et al., 2018). DDI covers dimensions of business, people, government, and infrastructure.

more intelligent than us and could decide to take over, and we need to worry now about how we prevent that happening.” As analysis of data from human development programmes in low-income countries becomes imminent, appropriate risk mitigation measures need to be put in place. AI will be a big industry, low-income countries should invest, now, in appropriate risk governance, frameworks and tools to manage AI driven analysis. Bias, ethics, manipulation, and other issues have been identified as areas where risk management is imperative when using AI for analysis of data. A major part of risk management will be on “data protection,” yet WDR (2022) estimates that “Only 24 percent of low-income countries have established data protection authorities, compared to 81 percent of high-income countries”.

Conclusion :

The large amount of aid poured into low-income settings, apart from buying services, can be used to generate positive externalities through analysis to support course correction. There is potential for such positive externalities with AI supported analysis. However, to enjoy such externalities from programme monitoring data of human development projects and interventions, low-income countries and their donors need to be intentional in their data investments. This paper has identified data governance, knowledge gaps, public data infrastructure and skills amongst Data Scientists in low-income countries as some of the most imminent areas that need investment. Development agencies such as World Bank, UNDP, UNEP, and UNICEF have already started such investments. They need to be supported. In addition to such investments, using AI for analysis of qualitative data will come with risks. Low-income countries will need to invest in public data infrastructure and data governance structures to manage risks that come with use of AI.

REFERENCES :

1. AI for Good Global Summit Report (2017), Geneva Switzerland, 2017.
2. Buluswar S, Roger S, Sridhar P, and Aluri, R (2018). Artificial intelligence and Data Analytics for Human Development. Separating facts from hype on where AI and data can genuinely help, and where it is a distraction. Institute for Transformative Technologies.
3. Chui, M & Harrysson, M (2019). AI in Human development. McKinsey Global Institute.
4. Emenike, M. E., & Emenike, B. U. (2023). Was This Title Generated by ChatGPT? Considerations for Artificial Intelligence Text-Generation Software Programs for Chemists and Chemistry Educators. Journal of Chemical Education. <https://doi.org/10.1021/acs.jchemed.3c00063>
5. Franzen, Samuel, Cuong Quang, Lukas Schweizer, Alexander Budzier, Jenny Gold, Mercedes Vellez, Santiago Ramirez, and Estelle Raimondo. 2022. Advanced Content Analysis: Can Artificial Intelligence Accelerate Theory-Driven Complex Program Evaluation? IEG Methods and Evaluation Capacity Development Working Paper Series. Independent Evaluation Group. Washington, DC: World Bank.
6. Hamilton, L, Elliott, D., Choplin, V (2023). Exploring the Use of AI in Qualitative Analysis: A Comparative Study of Guaranteed Income Data. International Journal of Qualitative Studies (September 2023)
7. Kalanda, B (2021). Comparison of costs of monitoring through UNICEF Yemen staff, Facilitators and Third-Party Monitoring Companies (UNICEF Yemen, unpublished).
8. Lennon, R., Calo, W., Miller, E., Zgierska, A, Van Scoy, L, Robert Fraleigh, R. (2022). Using artificial intelligence to support rapid, mixed-methods analysis: Developing an automated qualitative assistant (AQUA). Ann Fam Med. 2022 Apr 1;20(20 Suppl 1):2893. doi: 10.1370/afm.20. s1.2893.
9. Longo, L. (2020). Empowering Qualitative Research Methods in Education with Artificial Intelligence. In Advances in Intelligent Systems and Computing (pp. 1–21). https://doi.org/10.1007/978-3-030-31787-4_1
10. Mahto, Hogan, S., Hatfield, S., Coppola, M and Kulkarni, A (2022). Looping in your new sidekick. The role of machine learning in qualitative data analysis (Deloitte Insights, 2022).
11. Market Trends (2022). UNDP’s Initiative To Build an Immersive and Safer World for us
12. OECD (2023), "Detailed aid statistics: Tying status of bilateral ODA (Edition 2022)", OECD International Development Statistics (database), <https://doi.org/10.1787/c0791fa5-en> (accessed on 22 October 2023).
13. Siiman L, Rannastu-Avalos M, Pöysä-Tarhonen J, Häkkinen P, Pedaste M (2023).
14. Opportunities and Challenges for AI-Assisted Qualitative Data Analysis: An Example from Collaborative Problem-Solving Discourse Data. In: Huang, YM., Rocha, T. (eds) Innovative Technologies and Learning. ICITL 2023. Lecture Notes in Computer Science, vol 14099. Springer, Cham. https://doi.org/10.1007/978-3-031-40113-8_9 (Accessed, October 2023)
15. Smith, M & Neupane, S (2018). Artificial intelligence and human development: toward a research agenda. IDRC Research Results.
16. Thomas, M (2023). [12 Risks and Dangers of Artificial Intelligence \(AI\) | Built In](#) (Accessed, October 2023).
17. UNEP (2022). How artificial intelligence is helping tackle environmental challenges - [How artificial intelligence is helping tackle environmental challenges \(unep.org\)](#) (accessed, October 2023).
18. UNICEF (2022) - PIDB Recommender and the End of Year Summary Narrative (EYSN) Quality Assurance Tool (Unpublished data).
19. World Bank (2018). Text and Data Analytics unit. Brief (December 2018).
20. World Development Report (WDR) (2021). Data Governance System and Services (DGSS) data.