Available online at https://asric.africa/engineering-sciences
ASRIC Journal on Engineering Sciences Vol.6(1)(2025) 171-177

GLISSA: A GenAI-Augmented Framework for Real-time Situation Awareness of power Infrastructure based on Crowd Sourced Data

Godfrey Kibalya[1], JohnBosco Ssemakula[2], Richard Musanje[1]

[1]Dept. Computer Engineering, Busitema University, Uganda, gkibalya.eng@busitema.ac.ug, musanjerichard70@gmail.com

[2]Dept. ICT, Cavendish University, Kampala, Uganda, jbssemakula@cavendish.ac.ug

Corresponding Author: gkibalya.eng@busitema.ac.ug

Received 10 July 2025; revised 11 August 2025; accepted 23 October 2025

Abstract

Monitoring critical infrastructure (CI) like national power networks in developing countries is hindered by the limited coverage and high cost of traditional data collection methods (e.g., sensors, SCADA), making real-time situational awareness difficult during service disruptions. In contrast, social media platforms like X (formerly Twitter) offer a rich but underutilized source of real-time user-generated data. However, its unstructured, noisy, and diverse nature challenges traditional AI methods. This paper proposes a fully automated pipeline leveraging Generative AI (GenAI), specifically Large Language Models (LLMs), to extract and synthesize actionable insights from social media. The system includes Twitter API-based data ingestion, LLM-driven filtering and classification, geolocation inference, and visualization. A case study in Uganda validates its ability to detect power disruption events, addressing crowd-source data challenges.

Keywords: Infrastructure Monitoring, GenAI, LLMs, Social Media Analytics, Real-Time Systems, Uganda

I. Introduction

Critical Infrastructure (CI), including roads, power lines, and underground water systems forms the backbone of a country´s operations, as their disruption can trigger cascading effects across vital sectors such as the economy, education, public health, and safety (Xu et al., 2024 , Huang et al., 2022). However, in practice, maintaining uninterrupted 24/7 operation of such infrastructure is often unfeasible, particularly in developing countries, due to their heightened vulnerability to natural disasters like floods and human-induced threats such as vandalism, in addition to the inherent technical faults. In the face of service disruptions, having real-time operational status and situational awareness of the CI becomes essential for fault localization and implementing effective recovery strategies to reduce service downtime. This is especially important for large-scale, nationwide systems such as power lines which are typically managed by a single service provider, as is common in many developing countries like Uganda (Uganda Manufacturers Association, 2025). Nevertheless, maintaining situational awareness of critical infrastructure remains a complex and demanding task. Traditional monitoring and data collection methods such as physical inspections, SCADA systems, and sensor networks often prove inadequate due to limited coverage, slow response times, and susceptibility to disasters and disruptions like power outages. In Uganda, for example, the vast power grid stretching 65,129 km and the planned 1,443 km oil pipeline make real-time situational analysis through conventional means impractical. Deploying sensor networks across such expansive infrastructure would be prohibitively expensive, necessitating new innovative, resilient, and scalable monitoring solutions.

The proliferation of internet connectivity and widespread use of social media platforms has led to the generation of vast amounts of user-shared content in form of images, text, and videos related to critical infrastructure (Chen et al., 2021). This crowdsourced data offers an untapped opportunity to overcome the limitations of traditional data collection methods, particularly restricted coverage and prohibitive costs for massive deployment. By tapping into this real-time infrastructure related information, situational awareness can be significantly enhanced, especially for geographically distributed infrastructure, thus turning human social media users into mobile, distributed, adaptive and intelligent sensor nodes.

However, several key challenges arise when attempting to extract meaningful and actionable insights from social media data especially for the adopted case study:1) the data is inherently diverse in format and is often cluttered with noise, including unstructured grammar; 2) Similar sentence structure could imply opposite operational state of infrastructure i.e. sarcastic text; 3) multiple locations can have similar names, complicating incident localization; 4) keywords related to incidents are diverse, complicating effective search and filtering (Ziaullah et al., 2024). These complexities render conventional human-centric approaches even rule-based AI approaches insufficient for real-time, CI monitoring, and situational insight generation especially when high responsiveness and actionable insights are crucial.

As infrastructure becomes increasingly vulnerable to environmental threats and natural disasters, ensuring its resilience and consistent functionality has become a central research focus. A wealth of existing work (He et al., 2024, Oguntoye et al., 2024, Xu et al., 2024, Deelstra et al., 2023, Gagliardi et al., 2023, Inam et al., 2023, Huang et al., 2022, Wahyuni et al., 2014, Tariq et al.,2013) has explored enhancing the robustness of buried systems like transit tunnels and roadways through sensor networks and mathematical modeling. However, these methods often struggle with scalability due to their limited geographic reach and significant resource requirements, making them less suitable for fast-changing, large-scale environments.

To address limitations of conventional data collection schemes, recent works have explored the potential of crowd-sourced-data, specifically social media data for disaster insight generation and user behavioral analysis (Caliskan et al.,2025, Imran et al., 2025, Jiao et al., 2025, Yin et al., 2025, Zguir et al., 2025, 2025, Ziaullah et al., 2024). Crowd-sourced data has also been investigated for infrastructure monitoring with notable work by (Ziaullah et al., 2024) integrating AI, specifically a single LLM model for generation of synthetic text data, albeit within controlled datasets pertaining to predefined regions, thus overlooking practical challenges related to insight extraction from crowd-sourced data.

To address this gap, this paper introduces pipeline leveraging Generative AI to extract human-interpretable insights and generate real-time alerts about CI anomalies from social media streams. The system leverages the reasoning and pattern recognition capabilities of LLMs to handle large-scale and heterogeneous datasets. This approach is validated through a case study in Uganda power system demonstrating its effectiveness in identifying power-related service disruptions, particularly in power infrastructure while addressing core challenges in social media data analysis. Our contribution is fourfold:

- Automated pipeline for end-to-end CI situational analysis generation and insight extraction.
- LLM integration into the pipeline for external tool invocation, search keyword refinement and extraction of CI related alerts and insights.
- Integration of visualization module for rendering actionable insights regarding the power supply infrastructure.
- Incorporation of external tools (e.g., maps) to disambiguate locations and enhance the LLM's knowledge base.

The rest of the paper is organized as follows: Section 2 presents proposed framework GLISAA framework while the performance evaluation of the framework is introduced in section 3. The paper is concluded in section 4.

2. Methodology

The proposed architecture of the pipeline is shown in Figure 1. As shown, the framework is consisted of components for the following functionality:

2.1 Data Collection and Pre-processing

The framework crowdsources data from platforms like Twitter (X) using APIs, focusing on infrastructure-related information specific to Uganda. The data collection and preprocessing pipeline involved the following aspects:

2.1.1 Keyword Refinement

Initially, keyword-based searches were manually crafted using terms such as “power outage” and “power blackout + in Uganda.” However, these broad keywords often retrieved tweets unrelated to Uganda or the power infrastructure, while also missing relevant ones. To improve the search precision, tweets tagged to UEDCL (Uganda Electricity Distribution Company Limited) the Ugandan power distribution company were retrieved. Then using these tweets, a large language model (LLM) agent was employed to identify power-related events.

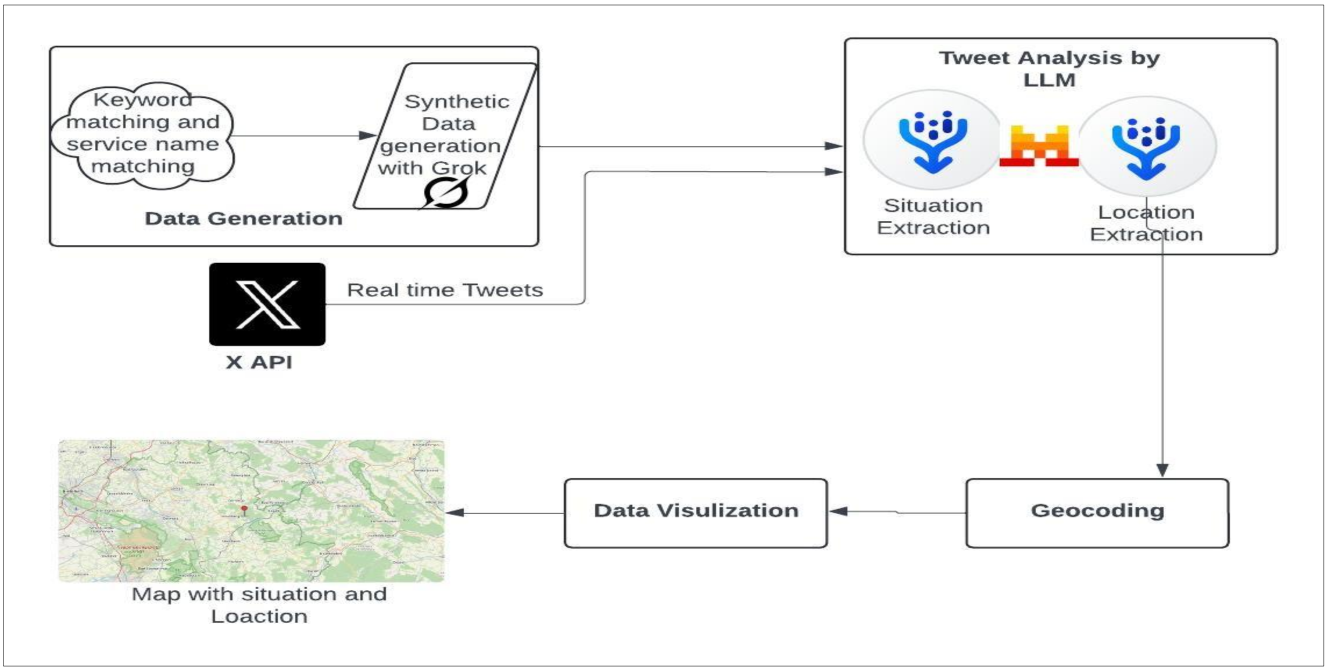


Fig. 1: . GLISSA Framework

From this curated set, the most frequently occurring and contextually relevant keywords across all tweets were extracted and used for improved tweet retrieval. It was identified that some of these incidents were specific to the Uganda context. For instance, " not possible to load Yaka " where Yaka refers to the prepaid tokens used to activate/credit the smart prepaid electric meters. This approach enhanced the relevance and geographic specificity of the data collected.

2.1.2 Synthetic Data Generation

Due to restrictions and high costs associated with the Twitter API, access to relevant Ugandan Twitter data was limited, which impeded the development of a comprehensive infrastructure monitoring system. To address this challenge, the study employed the Groq model to generate synthetic tweets that realistically simulate local scenarios (e.g., power outages, road disruptions, yaka loading issues, etc) to supplement the limited real data, facilitating effective training, testing, and validation of our pipeline under conditions reflective of Uganda’s infrastructure landscape.

2.2 Insight Extraction

This involved extraction of situation insights involving extraction of power service-related incident (such as outage, black out, falling pole, yaka loading failure, burning transformer, etc) and the associated locality for the incident. However, extraction of locations was complicated by fact that Names could be misspelt leading to failure to locate them to the visualization map. We applied fuzzy logic through RapidFuzz’s token_sort_ratio_scorer, which compares distorted location names to known ones by measuring similarity based on token sorting and edit distance. This allows the system to handle common distortions such as misspellings, causing errors, and spacing inconsistencies by selecting the closest match with a similarity score above a defined threshold.

2.3 Visualization

The incidents extracted by the pipeline are displayed on a visualization map highlighting high level insights regarding the infrastructure including incident name, location and where available cause of incidents.

3. Results and Discussion

This section introduces the key results of the proposed framework and their discussion.

3.1 Fuzzy Logic Performance

To validate the performance of the fuzzy logic, the locations in the tweets were subjected to random manipulation through character insertion, deletion, flipping and space insertion and case manipulation. The result shows that fuzzy logic predicts the correct location names to an accuracy of 0.82 which is 82% on average for all manipulations. An example of the character changes for the different places and the corresponding fuzzy score are shown in Table I.

Table 1: Fuzzy Matching results

True Location	Distorted Location	Predicted Location	Fuzzy Score
Kampala	Kampaa	Kampala	92.3
Kampala	Kam Pala	Kampala	80
Jinja	Jitnja	Jinja	90.9
Jinja	ji NJA	None	36.4
Mbale	Mblae	Mbale	80
Mbale	Mable	Mbale	80
Arua	AruA	Arua	75
Arua	Aua	Arua	85.7
Mbarara	Mbaara	Mbarara	92.3

The confusion matrix shown in Figure 2 shows that the fuzzy logic model performs well in identifying the different locations even after undergoing character-level distortions achieving an overall accuracy of 82%. This suggests that while the fuzzy logic model is robust to noise, certain locations may be more susceptible to ambiguity due to the similarity in their manipulated forms or shorter names.

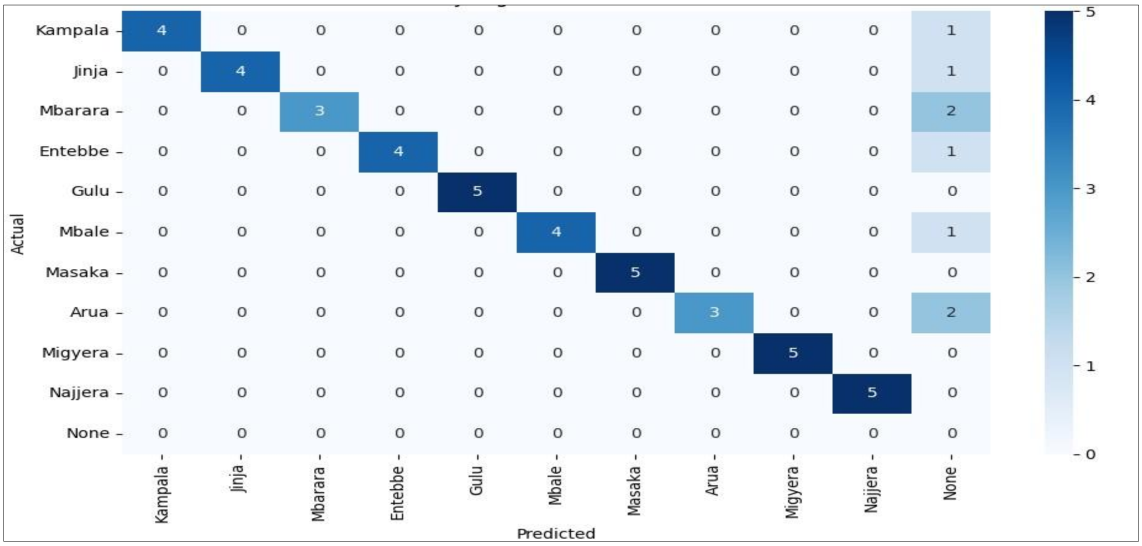
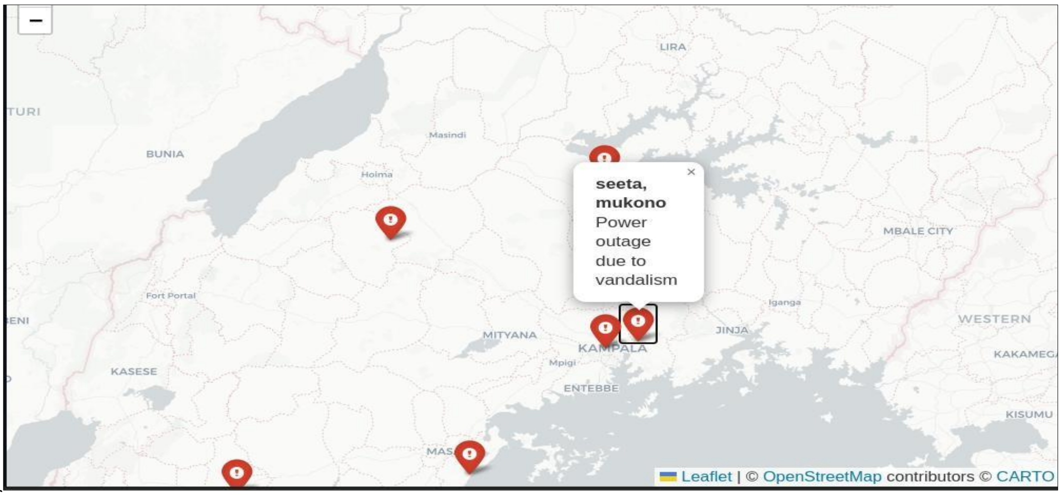


Fig. 2: Confusion Matrix for the Fuzzy matching step

3.2 Visualization

The output of the insight extraction process is illustrated in Figure 3. This visualization map highlights power-related incidents detected from social media data, including the affected locations and, where identifiable, the likely causes of the disruptions. By automatically surfacing this information, the system significantly reduces the manual effort typically required to sift through large volumes of unstructured online content for relevant incident alerts. This not only accelerates situational awareness but also enables quicker decision-making and response, especially in resource-constrained environments where timely intervention is critical.



s

Fig. 3: Visualization map for insight extraction

4. Conclusion

This paper has introduced a novel, fully automated framework that harnesses the capabilities of Large Language Models to extract actionable insights from social media data for realtime monitoring of critical infrastructure, specifically power networks in low-resource settings. By effectively addressing key challenges such as data noise, keyword diversity, and geolocation ambiguity, the system delivers scalable situational awareness with minimal dependence on expensive physical infrastructure. A case study conducted in Uganda demonstrates the framework's ability to detect and classify power-related incidents, such as outages, underscoring its practical value for infrastructure maintenance and disaster response. This work highlights the untapped potential of user-generated content as a complementary resource to traditional critical infrastructure monitoring tools, especially in regions where conventional methods are limited.

Future efforts will focus on integration of language translation capabilities into the framework to cater for the multilinguistic nature of Uganda social media users and integration of multi-modal data sources including sensory data, images and videos to generate insights.

Acknowledgments

This study was funded by the Carnegie Corporation of New York and Makerere University through The Consolidating Early Career Academics Program at Makerere University and Partner Public Universities (CECAP II) 2024—2027 Project (Grant Number G-PS-23-60690)

References

Caliskan, C., Kibris, E. S., & Patel, A. (2025). Victim assistance using large language models: The case of a massive earthquake in Hatay, Turkey. *SSRN.* https://doi.org/10.2139/ssrn.4787542

Chen, Y., & Ji, W. (2021). Enhancing situational assessment of critical infrastructure following disasters using social media. *Journal of Management in Engineering, 37*(6), 04021058. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000921

Deelstra, A., & Bristow, D. N. (2023). Assessing the effectiveness of disaster risk reduction strategies on the regional recovery of critical infrastructure systems. *Resilient Cities and Structures, 2*(3), 41–52.

Gagliardi, V., Tosti, F., Bianchini Ciampoli, L., Battagliere, M. L., D'Amato, L., Alani, A. M., & Benedetto, A. (2023). Satellite remote sensing and non-destructive testing methods for transport infrastructure monitoring: Advances, challenges and perspectives. *Remote Sensing, 15*(2), 418. https://doi.org/10.3390/rs15020418

He, X., Zhang, Z., Jing, L., Zhou, H., Zhu, Z., & Zhou, W. (2024). Design and application of wireless monitoring system for urban distribution network. In *Proceedings of the Fourth International Conference on Machine Learning and Computer Application (ICMLCA 2023)* (Vol. 13176, pp. 359–364). SPIE. https://doi.org/10.1117/12.3014872

Huang, H., Zhang, D., & Huang, Z. (2022). Resilience of city underground infrastructure under multi-hazards impact: From structural level to network level. *Resilient Cities and Structures, 1*(2), 76–86.

Imran, M., Ziaullah, A. W., Chen, K., & Ofli, F. (2025). Evaluating robustness of LLMs on crisis-related microblogs across events, information types, and linguistic features. In *Proceedings of the ACM Web Conference 2025 (WWW '25)* (pp. 5117–5126). Association for Computing Machinery. https://doi.org/10.1145/3696410.3714511

Inam, H., Islam, N. U., Akram, M. U., & Ullah, F. (2023). Smart and automated infrastructure management: A deep learning approach for crack detection in bridge images. *Sustainability, 15*(3), 1866. https://doi.org/10.3390/su15031866

Jiao, J., Park, J., Xu, Y., Sussman, K., & Atkinson, L. (2025). SafeMate: A modular RAG-based agent for context-aware emergency guidance. *arXiv preprint* arXiv:2505. https://arxiv.org/abs/2505.00000

Oguntoye, K. S., Laflamme, S., Sturgill, R., Salazar Martinez, D. A., Eisenmann, D. J., & Kimber, A. (2024). Probabilistic method to fuse artificial intelligence-generated underground utility mapping. *Sensors, 24*(11), 3559. https://doi.org/10.3390/s24113559

Tariq, A.-K., Ziyad, A.-T., & Abdullah, A.-O. (2013). Wireless sensor networks for leakage detection in underground pipelines: A survey paper. *Procedia Computer Science, 21,* 491–498. https://doi.org/10.1016/j.procs.2013.09.065

Uganda Manufacturers Association. (2025). *Uganda's manufacturing sector: A new era of electricity distribution with Uganda Electricity Distribution Company Limited (UEDCL).* Uganda Manufacturers Association. https://uma.or.ug/2025/04/01/ugandas-manufacturing-sectora-new-era-of-electricity-distribution-with-uedcl/

Wahyuni, R. T., Wijaya, Y. P., Nurmalasari, D., & Riau, P. (2014). Design of wireless sensor network for drainage monitoring system. *Innovative Systems Design and Engineering, 5*(5), 6–13.

Wang, A., Togo, R., Ogawa, T., & Haseyama, M. (2022). Defect detection of subway tunnels using advanced U-net network. *Sensors, 22*(6), 2330. https://www.mdpi.com/1424-8220/22/6/2330

Xu, F., Fang, D., Chen, B., & Wang, H. (2024). Resilience assessment of subway system to waterlogging disaster. *Sustainable Cities and Society, 113,* 105710. https://doi.org/10.1016/j.scs.2024.105710

Yin, K., Dong, X., Liu, C., Huang, L., Xiao, Y., Liu, Z., Mostafavi, A., & Caverlee, J. (2025). DisastIR: A comprehensive information retrieval benchmark for disaster management. *arXiv preprint* arXiv:2505.15856. https://arxiv.org/abs/2505.15856

Zguir, A. E. F., Ofli, F., & Imran, M. (2025). Detecting actionable requests and offers on social media during crises using LLMs. *arXiv preprint* arXiv:2504.16144. https://arxiv.org/abs/2504.16144

Ziaullah, A. W., Ofli, F., & Imran, M. (2024). Monitoring critical infrastructure facilities during disasters using large language models. *arXiv preprint* arXiv:2404.14432. https://arxiv.org/abs/2404.14432