

Leveraging Local Somali Radio Stations for Real-Time Food Security Insights: An AI-Powered Approach

Preliminary Phase

1. Introduction

1.1 Background

Somalia continues to face severe food insecurity challenges due to recurring droughts, political instability, and armed conflict that have disrupted agricultural production across the country. Traditional food security monitoring systems often suffer from data gaps, delayed reporting, and limited coverage in conflict zones, creating critical information barriers for humanitarian response. The lack of timely and accurate data prevents early warning systems from functioning effectively, especially in remote and insecure regions where vulnerability to food crises is highest. Humanitarian organizations struggle to allocate resources efficiently without reliable, up-to-date information on local conditions, market prices, and population movements. These information gaps can lead to delayed interventions, allowing food security situations to deteriorate to emergency levels before adequate assistance reaches affected communities.

1.2 Research Question

"Do local radio stations offer useful new information to improve food security indicators' accuracy and/or timeliness?"

Local radio stations in Somalia could potentially provide crucial real-time information about food security conditions that traditional monitoring systems miss due to data gaps, delayed reporting, and limited access to conflict zones. By tapping into community broadcasts through an AI-powered approach, humanitarian organizations might gain timely insights about local agricultural conditions, market prices, and population movements in remote areas, potentially enabling faster and more effective interventions before food security situations deteriorate to emergency levels.

It all depends on the specific activities of radio stations and what they cover in their radio broadcasts; we will try to check if there are programs that can provide us with such information.

1.3 Objectives

This preliminary phase evaluates whether Somali radio broadcasts can serve as a novel data source for enhancing food security indicators through an AI-powered pipeline. I am identifying relevant radio stations, establishing data collection methods, and testing speech-to-text technologies specifically optimized for Somali language. After completing the 4-week plan, the results will include:

- List of Somali radio stations with relevance to food security.
- Sample data of audio recordings.
- Transcriptions using speech-to-text tools.
- Notebook with working code for testing.
- Summary report with insights and recommendations.

2. AI-Powered Pipeline for Extracting Food Security Insights

2.1 Proposed Workflow

Radio Broadcast Collection → Speech-to-Text (STT) Conversion → Large Language Model (LLM) Processing (translation and summarizing) → Food Security Insights Integration

The proposed four-step AI-powered workflow begins with collecting radio broadcasts from Somali stations, converting speech to text using specialized tools, processing the transcribed content through Large Language Models, and finally extracting actionable food security insights also using LLM. In this preliminary phase, I will focus on the first two steps: finding appropriate radio stations and broadcasts, establishing methods for radio broadcast collection and testing speech-to-text conversion capabilities optimized for the Somali language.

3. Four-Week Plan for a Preliminary Phase

During the first week, I conducted research on Somali radio stations to evaluate their relevance to food security. I compiled a list of online stations, identified available live streaming options, and assessed whether they covered important topics such as food security, agriculture, or climate conditions.

In the second week, I focused on evaluating the feasibility of accessing radio broadcasts and collecting sample audio data. I developed a script to download and store the samples of recordings from one of the stations, setting the stage for subsequent analysis. The prototype was developed using Python due to its capabilities for simplicity and also for fast implementation of speech-to-text models. To facilitate access, testing speed, and free GPU access, the prototype was created on Google Colab with data storage on Google Drive.

By the third week, I employed speech-to-text tools like Whisper to transcribe the recordings. I also evaluated transcription accuracy and compared different models for Somali language processing to determine the most reliable option for our specific needs.

In the fourth week, I synthesized all my findings and outlined comprehensive recommendations for further research. I provided detailed insights into potential next steps for the project, ensuring a clear path forward for continued investigation into how Somali radio broadcasts can contribute to food security monitoring and response.

4. A Preliminary List of Online Somali Radio Stations Categorized by Relevance to Food Security

4.1 Overview of relevant radio stations

Below is a list of radio stations in Somalia that provide information related to food security and those that potentially may provide such information. Only stations accessible via the internet have been included. The stations have been divided into three main categories: Highly Relevant to Food Security which have dedicated programs for this purpose, Moderately Relevant to Food Security which have covered food security topics in the past, and Less Relevant to Food Security where information related to food security appears sporadically. Each radio station has a brief overview, a link to its website, a link to its archive or live stream, its connection to food security, and its coverage area.

4.2 Highly Relevant to Food Security (Dedicated Programs on Food Security)

Radio Ergo

- **Overview:** A humanitarian radio station in Somalia delivering programs on health, farming, and livestock to support local communities.
 - **Website:** [Radio Ergo](#)
 - **Archive:** [Radio Ergo Archive](#)
 - **Food Security:** Provides comprehensive coverage on food security, featuring news and discussions on agriculture, climate conditions, market prices, and humanitarian aid updates.
 - **Coverage area:** Nationwide in **Somalia**, including **rural and hard-to-reach areas** via **shortwave radio** and **local partner stations**.
-

Radio Daar Dheer

- **Overview:** Community-based station in Dhusamareb, supported by the United Nations, covering news, health, education, and culture.
 - **Website:** [Radio Daar Dheer](#)
 - **Live Broadcast:** [Listen Live](#)
 - **Archive:** [Radio Daar Dheer Archive](#)
 - **Food Security Program:** "Barnaamijka Horumarinta Wax Soo saarka Beeraha iyo Xoolaha" ("Program on the Development of Agricultural and Livestock Production").
 - **Coverage area:** Dhusamareb and surrounding areas.
-

Isnaay Radio

- **Overview:** Community radio station in Jowhar, Middle Shabelle, delivering news, entertainment, and music.
 - **Website:** [Isnaay Radio](#)
 - **Archive:** [Isnaay Radio Archive](#)
 - **Food Security Program:** "Barnaamijka Hormarinta Wax Soosaarka Beeraha" ("Program on Agricultural Production Development").
 - **Coverage area:** Jowhar and surrounding areas.
-

4.3 Moderately Relevant to Food Security (Occasionally Covers Food Security Topics)

Radio Waamo

- **Overview:** Community radio station in Kismayo covering local news and educational programs.
 - **Website:** [Radio Waamo](#)
 - **Live:** [Listen Live](#)
 - **Archive:** [Radio Waamo Archive](#)
 - **Food Security Program:** Occasionally covers food security topics through climate change discussions and business programs.
 - **Coverage area:** Kismayo and surrounding regions.
-

Radio Galmudug

- **Overview:** State-run broadcaster in Galmudug focusing on news, education, and cultural programs.
 - **Website:** [Radio Galmudug](#)
 - **Live:** [Listen Live](#)
 - **Archive:** [Galgaduud TV Archive](#)
 - **Food Security Program:** Participates in educational content on agriculture, livestock, and nutrition.
 - **Coverage area:** Galmudug region, including Dhusamareb and Galkayo.
-

Radio Kismaayo

- **Overview:** State-run radio station of Kismaayo, Jubaland, covering news, education, and culture.
- **Website:** [Radio Kismaayo](#)
- **Live:** [Listen Live](#)
- **Food Security Program:** Occasionally covers food security through news reports and discussions on drought resilience.

- **Coverage area:** Kismayo and surrounding areas.
-

Radio Daljir

- **Overview:** Established in 2003, the largest FM network in Somalia operating in multiple locations.
 - **Website:** [Radio Daljir](#)
 - **Live:** [Listen Live](#)
 - **Food Security:** Reports on food security topics but does not have dedicated programs.
 - **Coverage area:** Puntland, Galmudug, eastern Somaliland, Somali region of Ethiopia.
-

4.4 Less Relevant to Food Security (Minimal to No Regular Coverage)

Radio Arlaadi

- **Overview:** Community engagement radio station in Baidoa.
 - **Website:** [Radio Arlaadi](#)
 - **Live:** [Listen Live](#)
 - **Archive:** [Radio Arlaadi Archive](#)
 - **Food Security Program:** No dedicated programs, but occasionally covers related topics.
 - **Coverage area:** Baidoa and surrounding regions.
-

Radio Sanguuni

- **Overview:** Community radio in Dhobley focusing on local issues.
 - **Website:** [Radio Sanguuni](#)
 - **Archive:** [Radio Sanguuni Archive](#)
 - **Food Security:** No dedicated programs but covers occasional topics in broader discussions.
 - **Coverage area:** Dhobley and surrounding regions.
-

Radio Dalsan

- **Overview:** Leading Somali news station with social and political discussions.
 - **Website:** [Radio Dalsan](#)
 - **Live:** [Listen Live](#)
 - **Food Security Program:** Occasionally covers related topics but has no dedicated program.
 - **Coverage area:** Mogadishu, Jowhar, Baidoa, Adado, Wanlaweyn.
-

Radio Mogadishu (Radio Muqdisho)

- **Overview:** Established in 1951, Somalia's state-run broadcaster.
 - **Website:** [Radio Muqdisho](#)
 - **Live:** [Listen Live](#)
 - **Food Security:** Occasionally participates in food security initiatives.
 - **Coverage area:** Mogadishu and surrounding areas.
-

4.5 Conclusions and choice

Based on the radio stations listed above, the choice was simple - Radio Ergo was selected as the primary source for our food security monitoring prototype. Radio Ergo offers several distinct advantages that make it ideal for this research:

1. It provides comprehensive coverage on food security topics, featuring dedicated programs on agriculture, climate conditions, market prices, and humanitarian aid updates.
2. Its nationwide reach across Somalia, including rural and hard-to-reach areas via shortwave radio and local partner stations, ensures broad geographical coverage.
3. The station has an accessible online archive of past broadcasts, particularly their farming programs, making historical data collection feasible.
4. As a humanitarian radio station explicitly focused on delivering programs related to health, farming, and livestock to support local communities, Radio Ergo's content directly aligns with our research objectives.
5. The consistency and reliability of their food security programming creates a more standardized dataset for our AI-powered speech-to-text analysis pipeline.

These factors combined make Radio Ergo the most pragmatic starting point for testing our hypothesis about extracting valuable food security insights from Somali radio broadcasts. While other stations like Radio Daar Dheer and Isnaay Radio also offer dedicated food security programs, Radio Ergo's broader coverage and more extensive archive provide optimal conditions for our initial prototype development.

5. Data Collection

5.1 Radio Ergo

Radio Ergo is a specialized humanitarian radio service that broadcasts critical information across Somalia and Somali-speaking regions through a network of 25 local correspondents, covering essential topics like health, agriculture, displacement, and education. Through its programming, Radio Ergo amplifies the voices and stories of ordinary Somalis, including farmers, pastoralists, women, and the displaced, enabling them to participate in public discussions about humanitarian crises and development challenges. The station employs innovative communication channels, including shortwave transmission and the Freedom Fone call-in platform, allowing listeners in remote areas to both receive vital information and provide feedback on local challenges, which is then analyzed and shared weekly with humanitarian organizations to inform their response efforts.

Also, a great thing about Radio Ergo is that on their website you can find officially translated English documentation of conversations with affected/vulnerable people. Their updates from 2022 provided critical humanitarian information and documented firsthand accounts from affected communities across Somalia during a particularly challenging period. The 2022 updates covered a range of pressing issues including drought conditions, disease outbreaks, displacement, and conflict situations, offering valuable insights directly from local Somali correspondents based in regions most impacted by these crises.

I think this is another great resource which could be used for Real-Time Food Security Insights.

5.2 Availability of Radio Ergo Broadcasts

Radio Ergo provides an archive of past broadcasts, which can be accessed through their official website: [Radio Ergo Archive](#) and on their SoundCloud profile: [Radio Ergo on SoundCloud](#).

After multiple attempts to access the audio files directly from the frontend, I discovered that the broadcasts are stored on **SoundCloud**, making it possible to download them directly from there. However, Python libraries such as **yt-dlp**, **requests**, and **BeautifulSoup4** did not work for direct downloads due to site restrictions.

To address this, I developed a function that successfully retrieves these files. The function, along with other related functionalities, is available in an easily accessible **Google Colab notebook**: [Google Colab Notebook](#). This prototype allows for easy testing and verification of the file download process.

5.3 Implementation of Audio File Downloading

I implemented a function to **download audio files from Radio Ergo's SoundCloud profile** based on a **specified date range**. The function works as follows:

- It scrapes the SoundCloud profile page and extracts track URLs.
- It filters the extracted URLs based on the **broadcast date** using **regular expressions**.
- It downloads the matching **MP3 files** using **yt-dlp**.
- If running in **Google Colab**, it can optionally save the downloaded files to **Google Drive** for easy access.

For a step-by-step guide on how to download the broadcasts, refer to the **Google Colab notebook** linked above.

6. Implementing Speech-to-Text for Somali Language Recordings

6.1 Initial Approach: Standard Whisper Model

The implementation begins with OpenAI's Whisper model, which offers multi-language support. The function:

- Processes MP3 files automatically from a specified directory
- Creates and organizes transcript files systematically
- Includes options to save outputs to Google Drive

However, evaluation reveals a critical issue: while Whisper detects Somali language, it incorrectly transcribes it using Arabic script instead of Latin script (which Somali uses in writing). The transcription consists almost entirely of the repeated Arabic word "موضوع" (meaning "subject"), making the output completely unusable.

6.2 Improved Approach: Specialized Somali Model

Then I found and implemented a fine tuned version of Whisper model specifically for somali language from Hugging Face [steja/whisper-small-somali](https://huggingface.co/steja/whisper-small-somali) which shows significant improvement:

Key advantages:

- Correctly transcribes in Latin script as used in written Somali
- Produces recognizable Somali vocabulary and content related to drought conditions, humanitarian aid, and local initiatives
- Successfully identifies proper nouns, locations, and contextual terms

6.3 Evaluation Results

So after conducting initial tests, the transcription of Somali language content using Speech-to-Text models remains a significant challenge. OpenAI's Whisper model incorrectly transcribes Somali using Arabic script instead of Latin characters, making the output unusable. A specialized Hugging Face model, "**steja/whisper-small-somali**," shows considerable improvement by correctly using Latin script. However, it still suffers from severe repetitive hallucination patterns that fragment otherwise meaningful content.

Analysis reveals persistent quality issues, including **224 instances of adjacent word repetitions** and excessive repetition of phrases such as "**iyo iyo iyo**" and "**dhul dhul dhul**." These challenges underscore the need for language-specific model development. Languages with limited digital resources often receive less attention in general-purpose speech recognition systems, leading to lower transcription quality.

Achieving effective Somali transcription will likely require **dedicated fine-tuning of models** with high-quality training data rather than relying solely on general multilingual models.

6.4 Key food security-related topics identified in the transcription

Highlights from the March 15th transcript reveal the multifaceted food security challenges facing Somali-speaking communities. The Radio Ergo broadcast documents a dire situation where water

scarcity has forced families to travel up to 40 kilometers for drinking water, with costs reaching seven dollars per barrel—an unaffordable sum pushing many into debt.

The broadcast features testimony from Abdullahi Mohammed Ali, who describes how local water sources have completely dried up, leaving his community dependent on distant water points they struggle to access. His account underscores the precarious situation of pastoralist communities whose livestock—their primary economic resource—are dying from lack of water and forage.

Concerns about impending "macluul" (famine) conditions dominate the discussion, with warnings that without rainfall or external assistance, the situation could deteriorate further. The broadcast mentions various humanitarian interventions including nutrition programs, mobile health clinics, and supplementary feeding initiatives, though these appear insufficient given the scale of need.

Agricultural disruption due to consecutive failed rainy seasons has depleted local food reserves, driving up prices and reducing access to basic food commodities. As drought conditions persist across multiple regions, this March 15th broadcast from Radio Ergo highlights the urgent need for a coordinated humanitarian response to address the worsening food security crisis in these vulnerable communities.

7. Challenges & Considerations

7.1 Data

In the case of Somalia, a significant advantage and good news is the existence of Radio Ergo, which is an ideal source of information with very good and easy access to broadcasts. However, without it, I think it would be necessary to conduct an experiment in which we would need to verify periods when widespread famine occurred in Somalia and then transcribe their main broadcasts to see how much information about food security is shared there.

7.2 Speech to text models

As I mentioned earlier, the problem here is that despite clear and good quality recordings, models perform poorly with the Somali language. Here a decision needs to be made about the next step. Potentially, there is still the possibility of other fine-tuned models that can be found on Hugging Face.

8. Conclusion

So this preliminary phase has demonstrated the potential of leveraging local radio broadcasts for real-time food security insights, but also highlighted key challenges that need to be addressed.

The AI-powered pipeline offers a promising approach to tap into this valuable data source, as shown by the identification of relevant radio stations like Radio Ergo and the successful implementation of a system for collecting and downloading broadcast archives.

However, the major bottleneck remains the accuracy of speech-to-text transcription for the Somali language. While specialized models show improvement, significant issues with transcription quality persist, hindering the extraction of reliable food security information.

Moving forward, dedicated efforts must focus on refining Somali speech-to-text technology, potentially through fine-tuning models with high-quality training data or exploring alternative solutions.

Overcoming this technical hurdle is crucial to fully unlock the potential of local radio as a powerful tool for enhancing food security monitoring and enabling more timely and effective humanitarian interventions in Somalia.

9. Recommended next steps

1. Finding or fine-tuning speech-to-text models to be better and more accurate with the Somali language
2. After solving the problem with the model, conducting an experiment with a larger amount of data from Radio Ergo and testing with updated food security indices.
3. Considering the utilization of documentation already produced by [Radio Ergo on ReliefWeb website](#) and contacting them or the Swedish Development Agency (SIDA) and the Danish Development Agency (DANIDA) which fund Radio Ergo to initiate collaboration and ask how their processes look like right now. This could potentially allow for immediate analysis of their existing transcripts to explore how they might be used to update food security indicators.