

Program Title: Leveraging AI for Food System Data

PI: Laila Cure, **co-PI:** Souvika Sarkar

Agency/sponsor: Kansas Health Foundation – Mass Action

Final Report

1. Project Summary

Goal: To develop the prototype of an application that (1) acquires and preprocesses web data about food opportunities in Wichita (food pantries, soup kitchens and other food outlets that provide free or low cost food to residents), (2) ensures that food opportunity data is accurate and up to date, and (3) answers questions about these food opportunities using natural language.

Activity 1: Collecting and Preprocessing Food System Data: web data scraping will be used to periodically extract data from the internet. This project will focus on the food opportunities listed by the Kansas Food Source, but we are flexible on modifying this scope based on recommendations received during the strategy lab.

Activity 2: Enhancing Food System Data Quality through Augmentation: In instances of missing/incomplete information about a food opportunity, we will augment the dataset by generating the missing information based on additional searches or context from similar entries.

Activity 3: Developing a Prototype Food Opportunity AI Application using Open Source LLMs.

2. Final Expenditures

Funds for this grant were entirely spent on student salaries and tuition remission. A student from the School of Computing (TH) was hired to carry out the stated project activities under the supervision of Dr. Sarkar and Dr. Cure. Here is the information that was instructed to be entered into the system on behalf of the student:

- Dates: 8/17/2025-12/20/2025 (would be 9 pay periods (18 weeks))

- 9 hours per week would total 162 hours. The salary of \$2,777/162 hours would equal \$17.14/ hour
- Graduate Research Assistant

In addition, the student was provided with a tuition remission of \$1,372.

3. Goal Accomplishment Analysis

A food assistant chatbot was developed based on updatable data scraped from the Kansas Food Source using an open-source model, which is free to users. The student worked under the guidance of two faculty members. Dr. Sarkar provided expertise in artificial intelligence, chatbot development and coding, and technical evaluation. Dr. Cure provided guidance on the problem of identifying reliable sources of data on food opportunities and on evaluating the practical usefulness of the tool for community members. The team met on a weekly basis to assess progress and guide research activities. A white paper with the technical details of the research activities is attached to this report (titled *“Kansas Food Pantry Assistant Project: A Pilot Study”*). Below, we provide a summary of the activities proposed and their accomplishment level. Note that all the studies performed during the Fall semester were inherently pilot studies given the timeframe and exploratory nature of the analyses.

Activity 1: After evaluating several food opportunity repositories, the team decided to focus on the Kansas Food Source. This repository presented several data-related challenges, despite appearing to be the most up-to-date and structured repository from the three evaluated based on the City of Wichita, KS.

Activity 2: The team focused on identifying inconsistencies with the structure of the Kansas Food Source data for its use in the chatbot. Most of the inconsistencies were related to how the website is structured and how contents are tagged. Besides using census data to generate lists of counties and zip codes, no data augmentation was done beyond the contents of the Kansas Food Source dataset to assess whether food opportunities were valid.

Activity 3: The prototype Food Opportunity AI Application was developed using Llama 3.1 8B Instruct, accessed via a HuggingFace pipeline based on the preprocessed Kansas Food

Source dataset. We evaluated the chatbot based on how accurately it answered 50 queries asking for food opportunities with different characteristics, locations, or requirements. The open source chatbot was correct on 23 of 50 queries tested, a 46% accuracy. For comparison, a similar chatbot based on a proprietary AI model (GPT 3.5 Turbo) answered 26 of the same 50 user queries correctly, for a 52% accuracy.

In conclusion, this exploratory project supported student learning and understanding of challenges and opportunities in retrieval augmented generative AI and in validating web-based food system data, which will inform future research objectives. However, the accuracy of the prototype chatbot is not yet ready to support the community in identifying and accessing food opportunities.

In addition to the stated objectives and activities, Dr. Cure started a study evaluating the freely accessible versions of commercial language models, such as ChatGPT, Gemini Flash, and Copilot (report attached *"Preliminary Exploration of the Capabilities and Limitations of Publicly Accessible Language Models to Characterize a Food Opportunity System"*). This preliminary study found that these commercially available systems retrieving data directly from the web appear to have a higher accuracy than we could achieve with our chatbot. The preliminary study also highlighted structural issues with data beyond the Kansas Food Bank that we anticipated using in augmenting our food system list.

These results facilitated by the Kansas Leadership Center will inform streamlined research objectives in a subsequent project funded by WSU's MURPA with the same research team during the Spring 2026. The new project will build on the results to overcome the challenges identified and get closer to achieving the vision of leveraging AI to obtain an accurate up-to-date picture of a food opportunity system.