

# Non-conventional data sources and artificial intelligence applications at FAO Statistics

**Piero Conforti**

Deputy Director  
Statistics Division



**The Eleventh Meeting of the Steering Committee of the Arab Statistics Initiative (Arabstat)**

November 13 - 14, 2024

# Introduction: statistics at FAO



## A strong mandate

**Article I of the FAO Constitution** mandates the Organization to “collect, analyse, interpret, and disseminate information relating to nutrition, food and agriculture”

## Main functions

- Collect, analyse and disseminate agrifood systems data from member countries, generating statistics and accompanying dissemination materials
- Work with countries to develop technical capacities and improve statistical systems.
- Develop and promote international food and agricultural statistical methods standards
- Ensure governance of FAO's statistical system and position FAO in statistical fora and discussions at global level

## Statistics



## Data centers



### FAOSTAT

FAOSTAT provides free access to food and agriculture statistics (including crop,



### SDG Indicators

FAO's Sustainable Development Goals (SDG) Data Portal provides access to



### AMIS

The Agricultural Market Information System (AMIS) is an inter-agency platform to



### Food and Agriculture Microdata Catalogue (FAM)

The Food and Agriculture

# A number of platforms and projects

## Main areas of work

- Agricultural Censuses
  - Agricultural Surveys
  - Crops, Livestock and Food Balances
  - Environment Statistics
  - Food Security and Nutrition Statistics
  - Social and Economic Statistics
- 
- Data and Statistical Coordination
  - Data Quality
  - Dissemination and Outreach
  - Methodological Innovation





## FAOSTAT

[Home](#)
[Data](#)
[Selected Indicators](#)
[Compare Data](#)
[Rankings](#)
[Definitions and Standards](#)
[FAQ](#)

### Data

DOMAINS    DOMAINS TABLE

- ▶ Production
- ▶ Food Security and Nutrition SDG indicators
- ▶ Food Balances
- ▶ Trade
- ▶ Prices
- ▶ Cost and Affordability of a Healthy Diet
- ▶ Food and Diet
- ▶ Land, Inputs and Sustainability
- ▶ Population and Employment
- ▶ Investment SDG indicator
- ▶ Macro-Economic Indicators
- ▶ Food Value Chain
- ▶ Climate Change: Agrifood systems emissions
- ▶ Forestry
- ▶ Fisheries
- ▶ SDG Indicators
- ▶ World Census of Agriculture

# FAOSTAT

- Mostly official data, some analytical datasets
- Global scope: 245 countries and territories; national, regional and global
- Long-term (1961–present)
- Over 20 thematic domains (e.g., production, trade, food balances, environment, climate change, employment, investment)

<https://www.fao.org/faostat/>



<b>GOAL 2:</b> Food security, Nutrition, Sustainable agriculture	2.1.1	2.1.2	2.3.1	2.3.2	2.4.1	2.5.1	2.5.2	2.a.1	2.c.1
<b>GOAL 5:</b> Gender equality	5.a.1	5.a.2							
<b>GOAL 6:</b> Use of water	6.4.1	6.4.2							
<b>GOAL 12:</b> Sustainable consumption and production	12.3.1								
<b>GOAL 14:</b> Oceans	14.4.1	14.6.1	14.7.1	14.b.1					
<b>GOAL 15:</b> Life on land	15.1.1	15.2.1	15.4.2						

status of the monitoring

**TIER I** – Established methodology exists and data already widely available

**TIER II** – Methodology established but insufficient coverage (>50% country coverage)

**TIER III** – Internationally agreed methodology not yet developed

# **Non-conventional sources and data science methods at the FAO Data Lab**

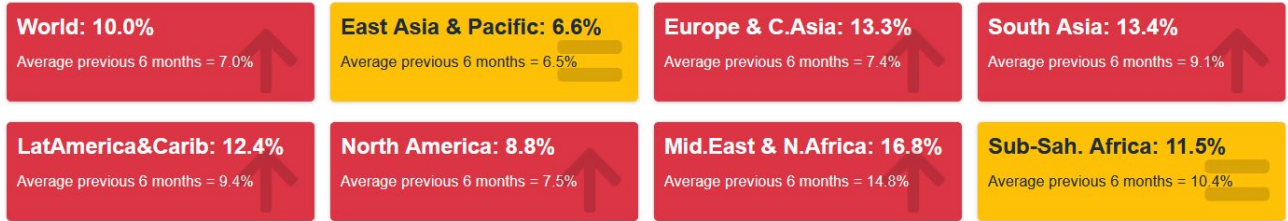
# Food inflation nowcasting and monitoring

Regional aggregation of food inflation

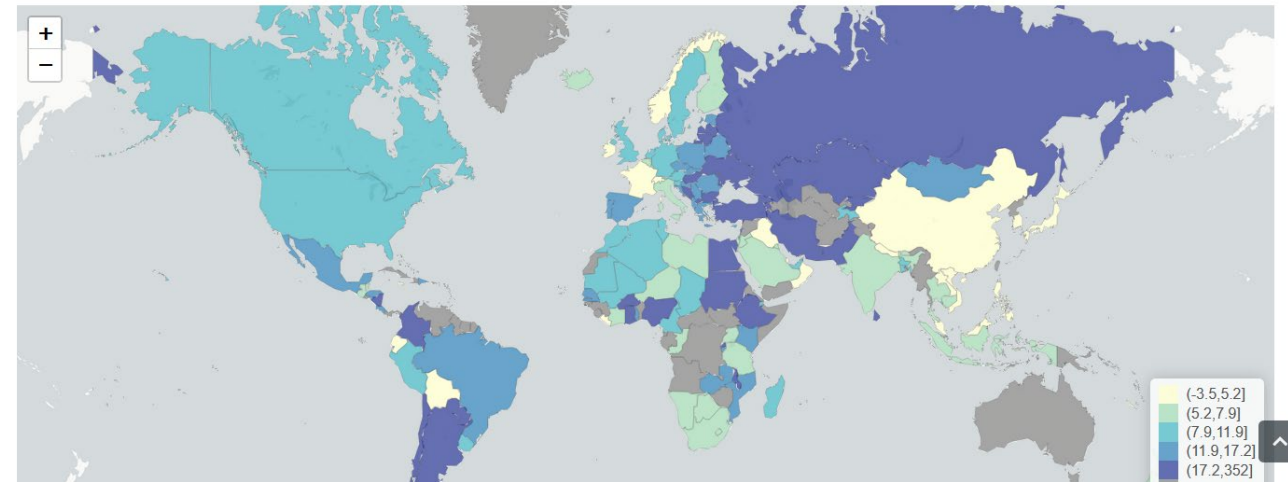
Interactive map with nowcasts

Table with details

Regional Food Consumer Price Index average variations



Expected Year-over-Year Food Prices Variations Map



Country	Region	Nowcasts			Moving mean
		point	lower	upper	
Albania	Europe & Central Asia	14.27	14.10	14.44	7.37
Algeria	Middle East & North Africa	8.53	7.12	9.95	8.57
Antigua and Barbuda	Latin America & Caribbean	7.92	6.72	9.13	6.55
Argentina	Latin America & Caribbean	63.13	62.78	63.49	54.82
Armenia	Europe & Central Asia	18.30	18.10	18.50	13.42
Aruba	Latin America & Caribbean	7.52	6.81	8.25	5.64
Austria	Europe & Central Asia	9.38	9.20	9.56	4.45
Bahrain	Middle East & North Africa	10.39	9.77	11.00	7.19
Bangladesh	South Asia	11.24	11.07	11.42	5.87
Barbados	Latin America & Caribbean	9.16	6.37	12.02	2.81



# FAO Hand-in-Hand Initiative

## Sub-national data from a variety of sources

### Sub-national data



Scraping Sub-National  
Agricultural Production  
and Harvested Area

- Data are scraped by text-mining documents coming mainly from National Statistical Offices or Ministries.
- Data from **33 countries** were harvested at several administrative levels and variety commodities. For some countries, it was possible to obtain data from **1976**.
- The FAO Data Lab releases the sub-national data, but also supports the production team / **FAOSTAT** to validate their data through scraped sub-national data, and contributed also to increase the amount of countries covered.
- **Database Integration with Hand-in-Hand Geospatial Platform:** data is automatically imported into the platform every day, so any new data added to DL database gets added to the GIS platform seamlessly. This is also true for any updates to existing data.

# Food Loss and Waste Database

## FLW Database

DataLab Logout Logged In User: christian.mongeau@fao.org

Pages ▾ Previous 1 2 ... 44 45 46 47 48 49 50 ... 56 57 Next

Page 47

Table 18 indicates percentages of potato waste in developing countries. Asia stands out with 15 percent losses, followed by Latin America and the Caribbean with potato losses of 11 percent.

In Table 19 total loss estimates are presented for some countries. United States registers the smallest loss at 1.5 percent and Dominican Republic the largest with 27 percent. It is important to note that loss averages fluctuate between 25 and 27 percent.

In case of Peru, no total losses for potato are quantified, but according to officials of the Peruvian Ministry of Agriculture losses in an improved storage facility can be 1.5 percent.

This contrasts to up to 50 percent in a traditional warehouse over a 6 month storage period. Wholesale Market N°1 in Lima, Peru reports from 1 000 to 1 200 t/day potatoes incoming, which is supplied with potatoes coming from the Andean uplands (Sierra) (9 percent) and coastal desert plains (Costa) (3 percent).

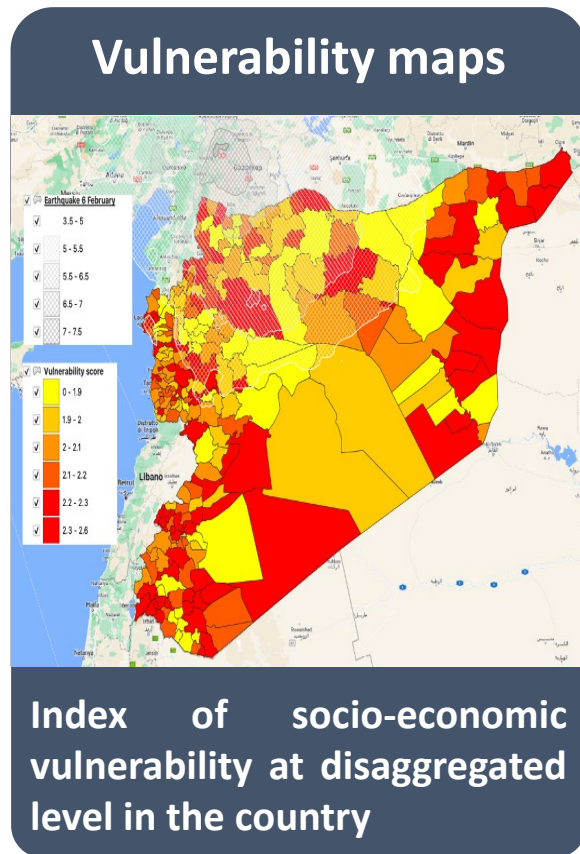
**Table 18. Potato waste percentage in developing countries. Averages in 1991-1992 years.**

COUNTRY and REGION	% WASTE
ASIA	
Bangladesh	15
China	10
India	5
Indonesia	17
Iran	6
Korea RPD	10
Nepal	10
Pakistan	15
Syria	10

Food loss and waste data from non-conventional sources.

- The FAO Data Lab scrapes the web to **automatically identify** reports and studies **measuring food loss and waste**, and uses Natural Language Processing (NLP) procedures to **identify food products, stages of the value chain**, and geographical areas, and quantities.
- The Fao Data Lab developed an user-interface that allows the user to build queries, and extract relevant information from the documents.
- **Dissemination: the Food Loss and Waste database is open access.**

# Vulnerability maps



The FAO Data Lab was engaged in geospatial analysis and built maps with **estimates at district and sub-district level** of:

- **Land Cover (LC)**
- **Vulnerability maps (LC + Socioeconomic indicators)**

Featured to:

- **Data:** costless data acquisition by leveraging online public sources
- **Speed:** get the results in a few hours (country-size-dependent)
- **Data poor countries**
- **Expert validation:** maps were submitted to country teams, through the FAO Offices.

# **Examples of applications of AI methods and tools at the FAO Data Lab**

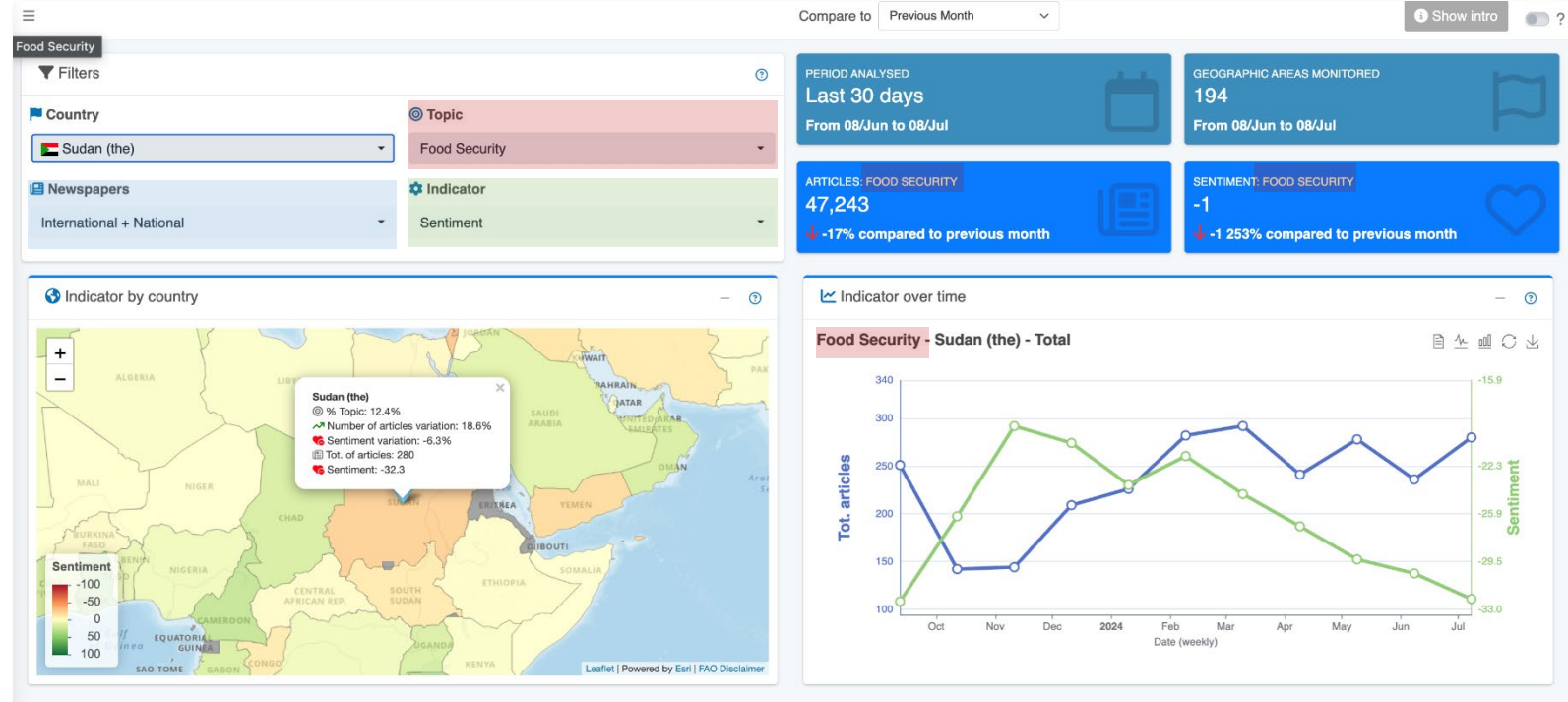
# Topics Explorer

- **Monitors global press topics related to Sustainable Development Goals (SDGs)**

*Food Security – Food Losses & Waste Climate Change  
– Social Unrest – COVID*

*Food Prices – Healthy Diets*

- Provides **sentiment** and **popularity** metrics for topics in **various countries**;
- Helps analysts and policymakers track how topics are treated in the news and **identify emerging issues**.
- **National and International newspapers perspectives**



<https://www.fao.org/datalab/early-warnings/topics-explorer>



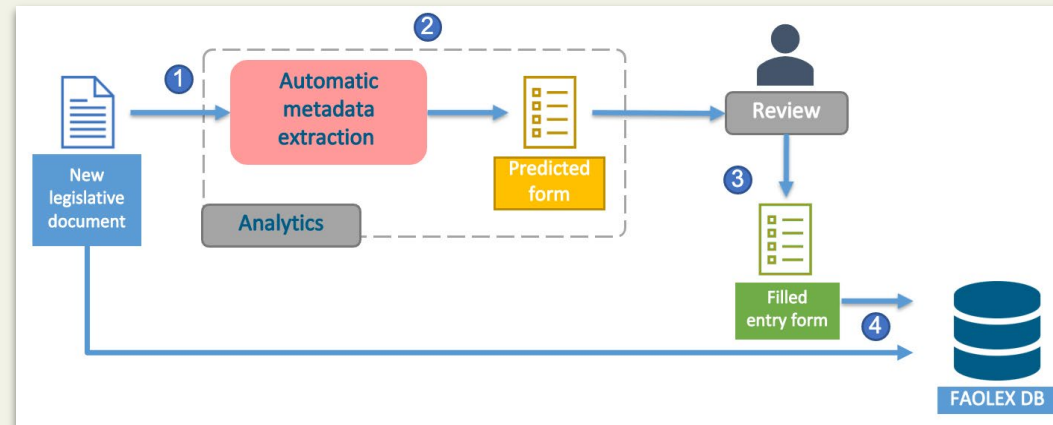
# Analysis and classification of legal and policy documents

The screenshot shows the FAOLEX Project website. At the top, there is a dark blue header with the text "FAOLEX Project" in white. Below this is a navigation bar with the FAO logo and the text "Food and Agriculture Organization of the United Nations". A search bar is visible with the text "UNHANCED BY Google". Below the navigation bar, there is a section titled "FAOLEX Database" with a list of links: "Background", "Country Profiles", "Thematic Collections", "Associated Databases", "Highlights Archive", "Glossary", "COVID-19", and "Open Data". A description of the database is provided: "FAOLEX is a comprehensive and up-to-date legislative and policy database, one of the world's largest online repositories of national laws, regulations and policies on food, agriculture and natural resources management. Users of FAOLEX have direct access to the abstracts and indexing information about each text, as well as to the full text of the legislation and policies contained in the database." Below this is a search bar with the text "FAOLEX Document Search" and an "ADVANCED SEARCH" button. A "Browse by Domain" section follows, with a grid of icons and labels for various domains: "AGRICULTURAL AND RURAL DEVELOPMENT", "CULTIVATED PLANTS", "ENVIRONMENT", "FISHERIES AND AQUACULTURE", "FOOD AND NUTRITION", "FORESTRY", "LAND AND SOIL", "LIVESTOCK", "MINERAL RESOURCES AND ENERGY", "SEA", and "WATER". At the bottom, there is a "Highlights" section with a list of recent legislative acts and a "Related Links" section with a list of links: "Legal Office", "Development Law Service", and "Legislative Studies".

**Automatic metadata prediction**

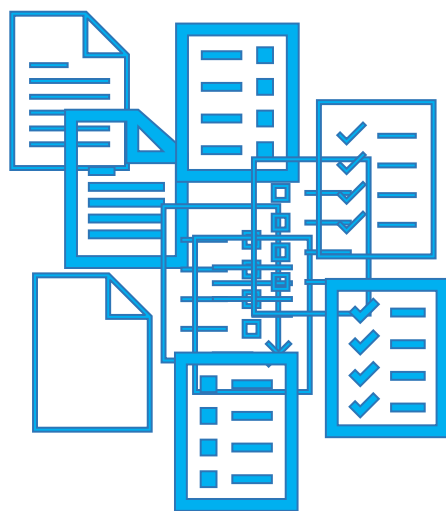
FAOLEX is one of the world's largest online repositories of national laws, regulations and policies on food, agriculture and natural resources management.

The goal of the **FAOLEX project** is to have **metadata predicted automatically** by machine learning algorithms.

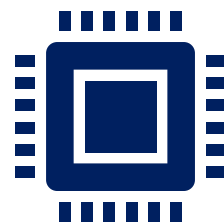


## Data-driven Taxonomies through AI

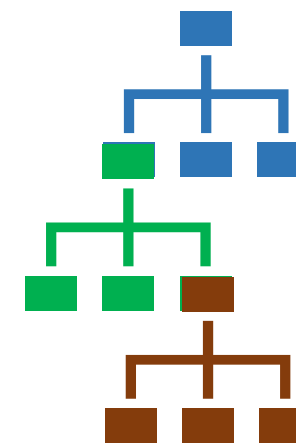
The FAO Data Lab is collaborating with the FAO's Office of Innovation to create a taxonomy of technologies built with AI methods on patent data.



Millions of patent documents

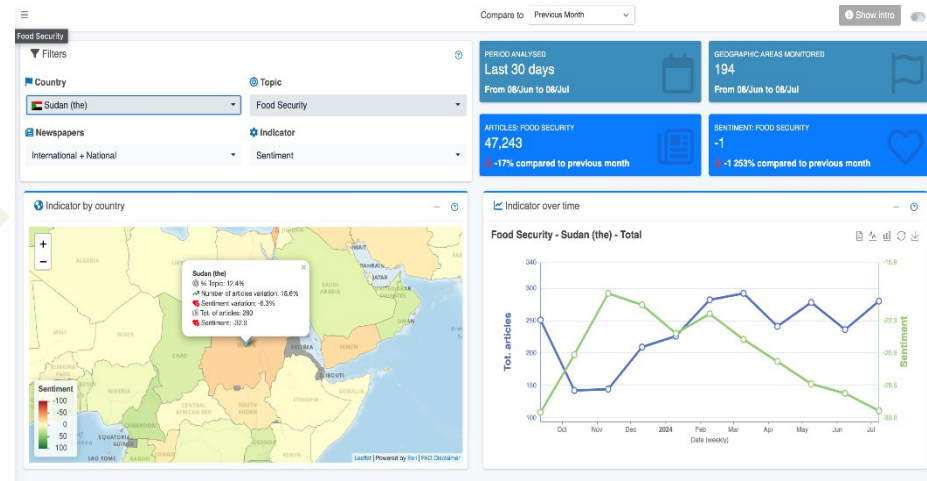
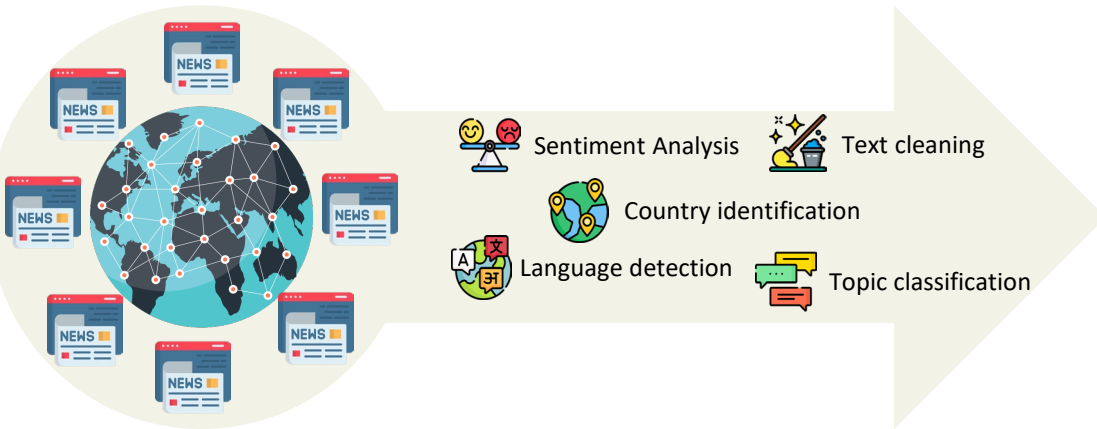


AI model



Taxonomy

# Narratives behind the numbers



## What's behind?

- Food Security + Sudan
- Why sentiment ↘ ?
- Why food security ↗ ?
- Reason?

### Report on Food Security Issues in Sudan:

Sudan is currently facing a severe food security crisis with at least 14 areas at the highest level of acute food insecurity, raising concerns of widespread famine due to ongoing civil war. Reports indicate that the conflict between the army and the paramilitary Rapid Support Forces is exacerbating the situation, leading to displacement of millions and acute malnutrition among over 3.6 million children in the country.

The geographic areas within Sudan experiencing food security problems are undefined in the articles, but the conflict is widespread across the country. The underlying causes of these issues stem from the civil war, which has disrupted food production and distribution systems, leading to limited access to food and essential resources for the population.

Responses and interventions addressing the food security problems in Sudan include calls from aid organizations for immediate humanitarian action to prevent famine. The United Nations has urged parties in the conflict to allow for the delivery of humanitarian aid to those in need to avert a potential humanitarian catastrophe. UN agencies and the international community are monitoring the situation closely and advocating for a cessation of hostilities to facilitate the delivery of vital assistance to vulnerable populations.

This report highlights the urgent need for humanitarian action to address the escalating food security crisis in Sudan, emphasizing the importance of immediate interventions to prevent further deterioration of the situation and mitigate the risk of famine.

Report generated automatically through ChatGPT



### Top 20 news on Food Security in Sudan

Al-Naba: Crisis in Sudan: AU mechanism convenes new talks

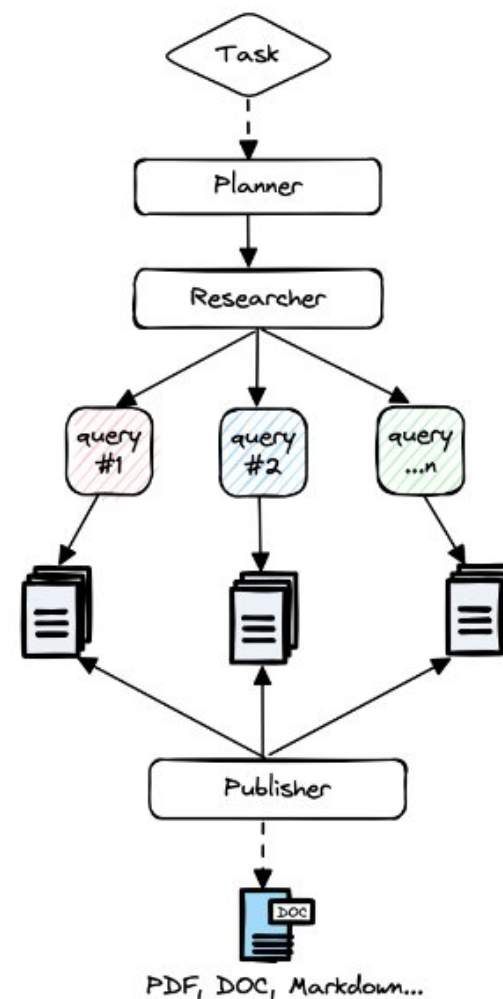
TCHADINPOS: 780 millions de personnes ne mangent pas à leur faim dans le monde

The Arabi Post: Famine risk is real for 14 areas of Sudan amid ongoing fighting

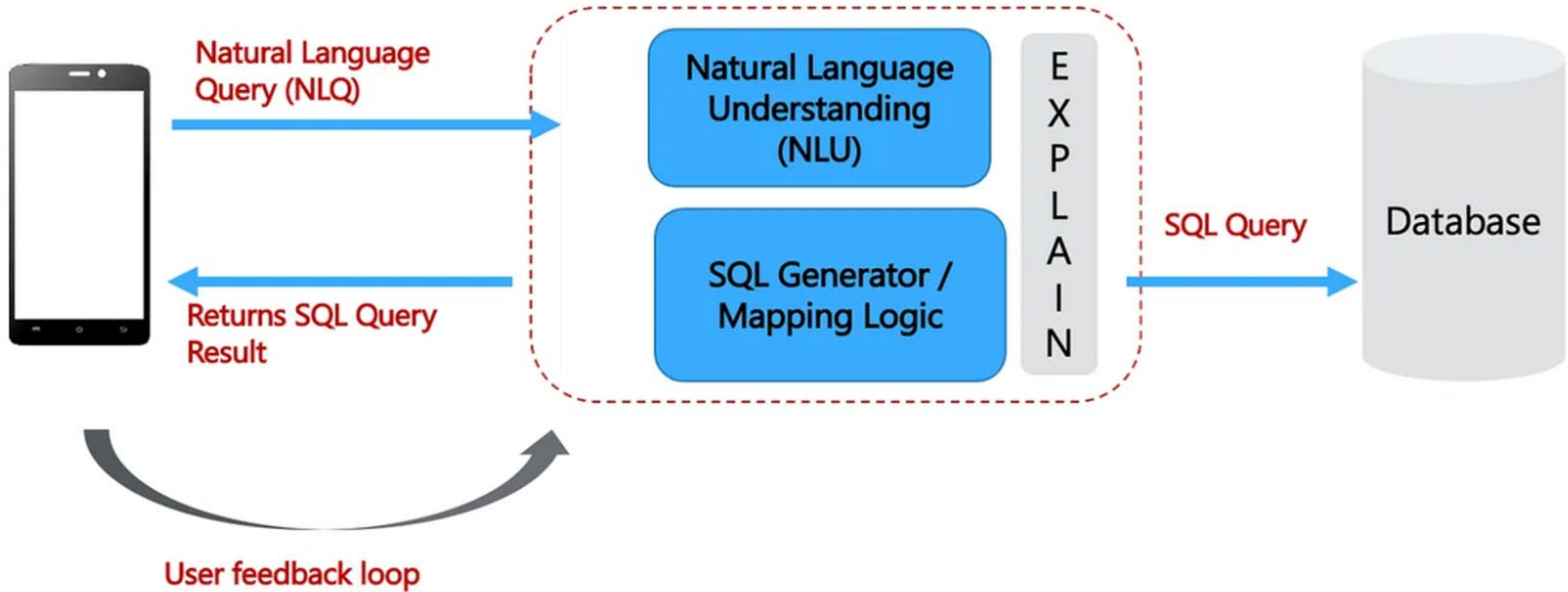
The Guardian: Internal displacement in Sudan exceeds 10 million

# Generation of draft reports with RAG systems and LLM-agents

- **Retrieval-Augmented Generation (RAG)** is a method that grounds LLM responses in defined knowledge bases to reduce the likelihood of "hallucinations" (incorrect or fabricated information)
- **LLM-agents** are systems that can use an LLM to reason through a problem, creating and executing a plan to solve it with the help of a set of tools
- The **FAO Data Lab** uses RAG and LLM-agents to create draft research reports, helping analysts quickly gather and organize insights from trusted sources, ensuring that these drafts pull accurate, verified information, allowing analysts to focus on refining and interpreting key findings efficiently



## User Interface





# Thank you

## for your attention

FAO Statistics <https://www.fao.org/statistics/en/>

FAO Data Lab <https://www.fao.org/datalab/en>

Contact us at [FAO-Statistics@fao.org](mailto:FAO-Statistics@fao.org)

Piero Conforti (PhD)

Deputy Director

Statistics Division <https://www.fao.org/about/who-we-are/departments/statistics-division/en/>

Food and Agriculture Organization of the United Nations (FAO)

Viale delle Terme di Caracalla, 00153 Rome, Italy