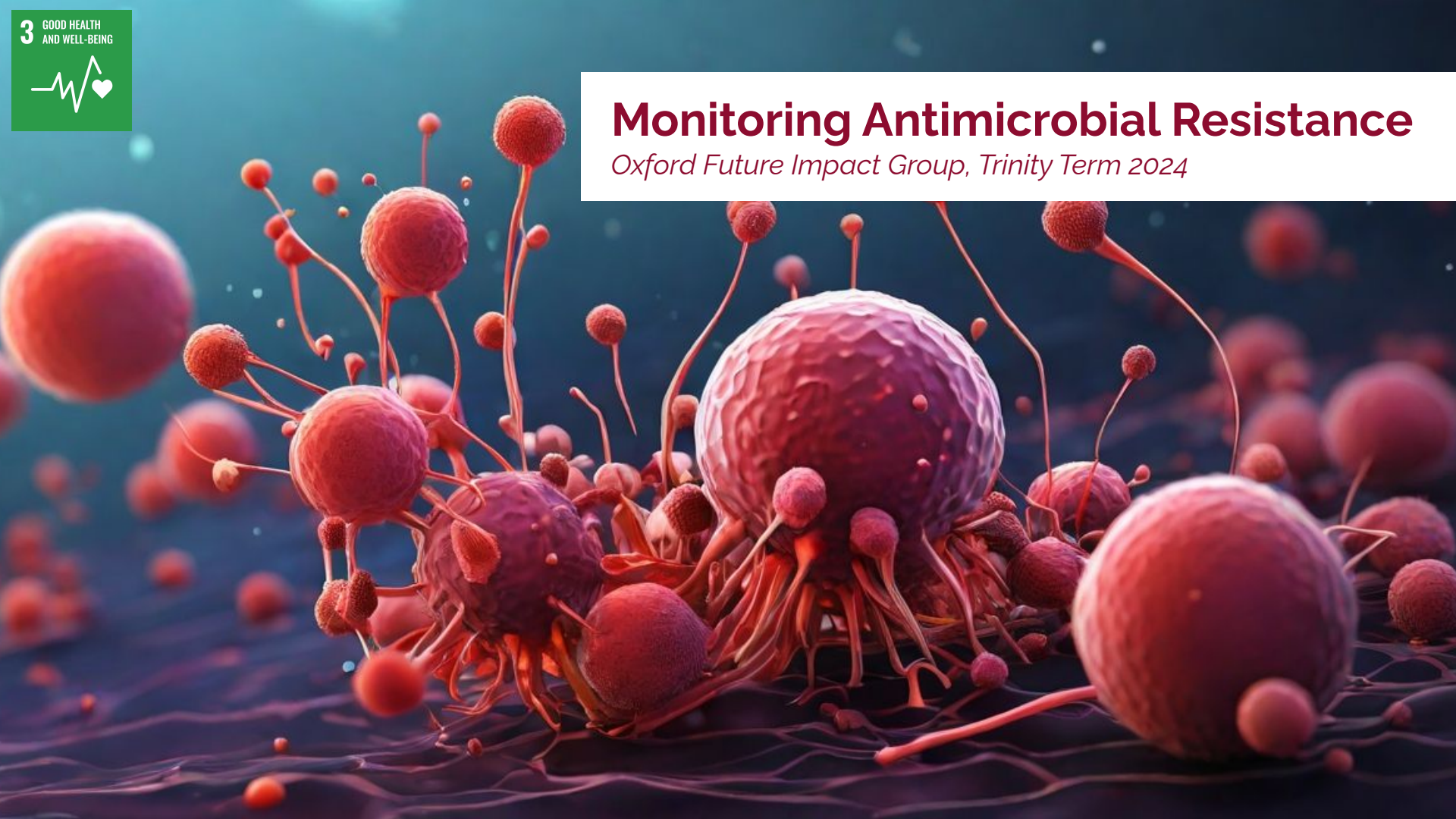


Monitoring Antimicrobial Resistance

Oxford Future Impact Group, Trinity Term 2024

Research Question:

How can large language models (LLMs) be leveraged to improve our knowledge of antimicrobial resistance (AMR)?

Research Overview

The goal of the project is to produce a tool that will scrape the internet for local AMR news and output a CSV

Problem

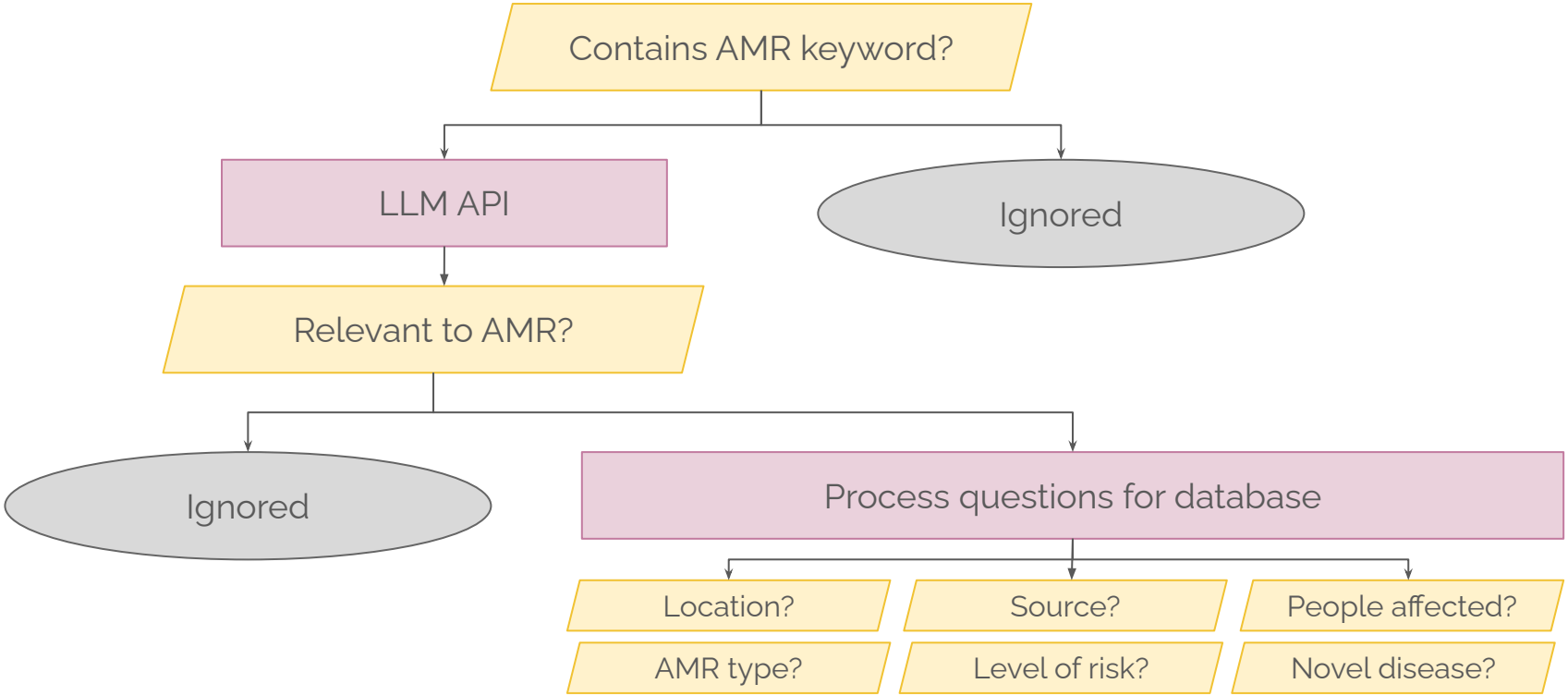
Local news has the potential to aid in detection of antimicrobial resistance (AMR) risks. However, articles may be in local languages or be too numerous for humans to read in full

Goal

Use Selenium to scrape the internet for AMR-related stories around from regions of interest then apply public large language models to extract relevant information from the articles (type of threat, location, number of people affected, antibiotics administered etc.) for storage in a global AMR surveillance spreadsheet

Model Outline

The tool will scrape articles from the internet in local Indian languages to identify AMR instances and risk



Example Article

Title

Indian man is world's first person to contract fungal infection from a plant

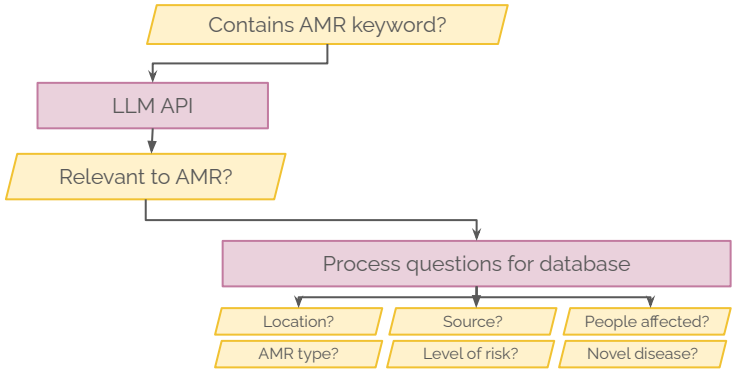
The mycologist was infected with Chondrostereum purpureum – the same fungus that causes silver leaf disease in plants

Joe Wallen IN MUMBAI *and* Sarah Newey, GLOBAL HEALTH SECURITY CORRESPONDENT, IN BANGKOK
30 March 2023 • 6:38pm

te

It is believed the man had come into contact with the fungus while carrying out research in his job | CREDIT: DENNIS KUNKEL MICROSCOPY/SCIENCE PHOTO LIBRARY

Article is detected by the keyword web scraper, and then also is interpreted by ChatGPT to be relevant to AMR – which flags specific concerns about **novel fungal diseases** (though it notes that there is limited person-to-person transmission), **the need for effective anti-fungal options** (it even catches that two different medications were required for a two month period). This would produce a flag for **novel disease risk in Kolkata**.



Example Article

Title

JN.1 Covid variant: WHO charts its rapid global spread

20 December 2023

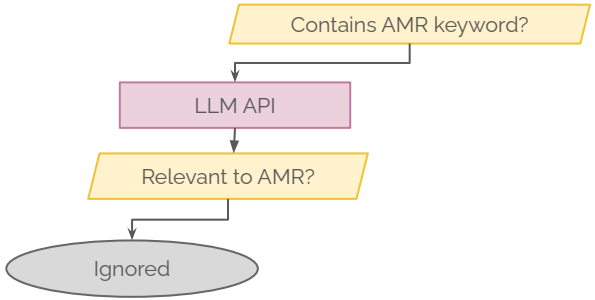
Coronavirus

By Philippa Roxby

Health reporter

A sub-variant of the Omicron strain of coronavirus has been classified as a "variant of interest" by the World Health Organization, because of "its rapidly increasing spread".

Article is detected by the keyword web scraper, but is **NOT interpreted by ChatGPT to be relevant to AMR** as it relates to a viral infection and does not mention the use of antibiotics.



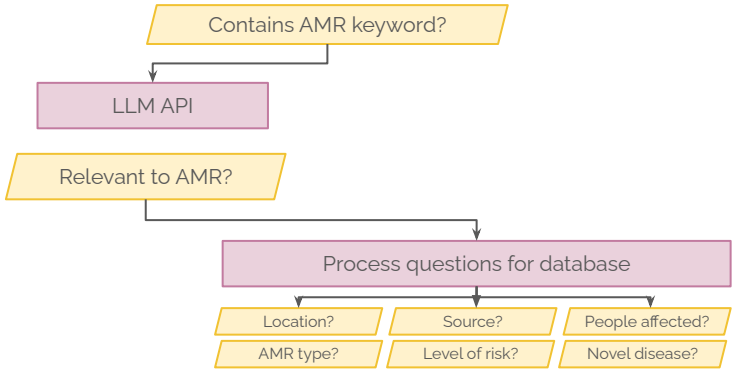
Example Article

Title

Example Article

Title

Article is detected to raise indirect concerns about AMR due to the **excess use of antibiotics** described in the article, though it notes that it is difficult to determine the level of risk from the text. This would produce a flag for **excess antibiotic use in Surat and Bhavnagar**.



Research Plan

Tasks will be due at each weekly meeting, where team members will present results and discuss challenges

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6
Literature review	Everyone					
Cost research	Caleb, Jess					
Theory of change	Richard, Sharon					
Database targets		Richard, Sharon				
Article selection			Richard, Sharon			
Keyword selection				Richard, Sharon		
Prompt engineering					Richard, Sharon	
Review database						Richard, Sharon
Scraping pipeline		Caleb, Jess				
LLM API			Caleb, Jess			
Standardise results				Caleb, Jess		
Run program					Caleb, Jess	

Optional Follow-Up Work

There are additional opportunities for further work on this product, dependent on research outcomes

Reports

- Report on dispersion and progression of AMR-related news in India
- Policy recommendations for specific international bodies or regional governments
- Describe research methods for non-technical audience

Collaboration

- Ensure GitHub repo is reproducible and user-friendly for open-source contributions
- Work with AMR organisations to use tool and data

Presentation

- Present work to reporters at G20 Rio Summit
- Potential future opportunities to present to other interested parties