Automating Documents Screening & Data Extraction Using Large Language Models for Enhanced Horizon Scanning

Muhinyia Ndegwa[1] Saleh Mohammed[2]
1 – muhinyiandegwa@gmail.com 2 – salehmohammed@ncl.ac.uk

BACKGROUND

Horizon scanning is the systematic process of identifying, filtering, analysing and monitoring therapies and research that may impact healthcare delivery in the future.

- Early detection of incoming discoveries helps inform policy, resource allocation and clinical guidelines.
- The NIHR Innovation Observatory (IO) has a national remit to deliver medical horizon scanning for the NHS.
- To achieve this, information analysts must sift through thousands of publications, reports, among others.
- This manual review creates a bottleneck: initial screening takes analysts months, delaying critical insights. The process is tedious and prone to human bias, resulting in errors. Critical innovations slip unnoticed due to the volume of the workload.

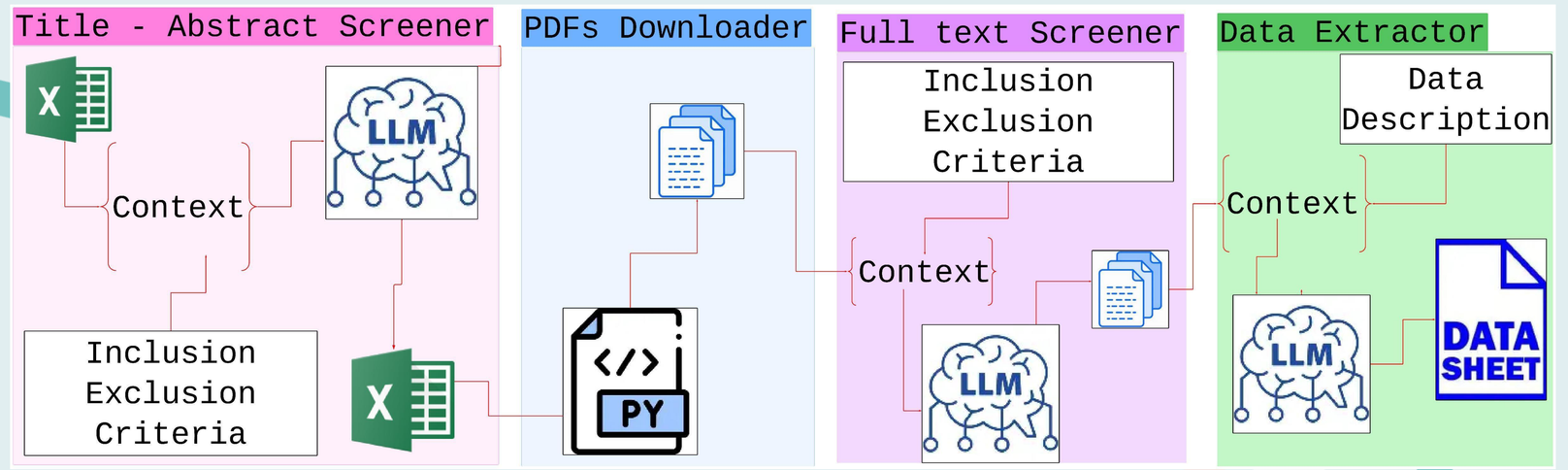

Figure 1: Pipeline architecture

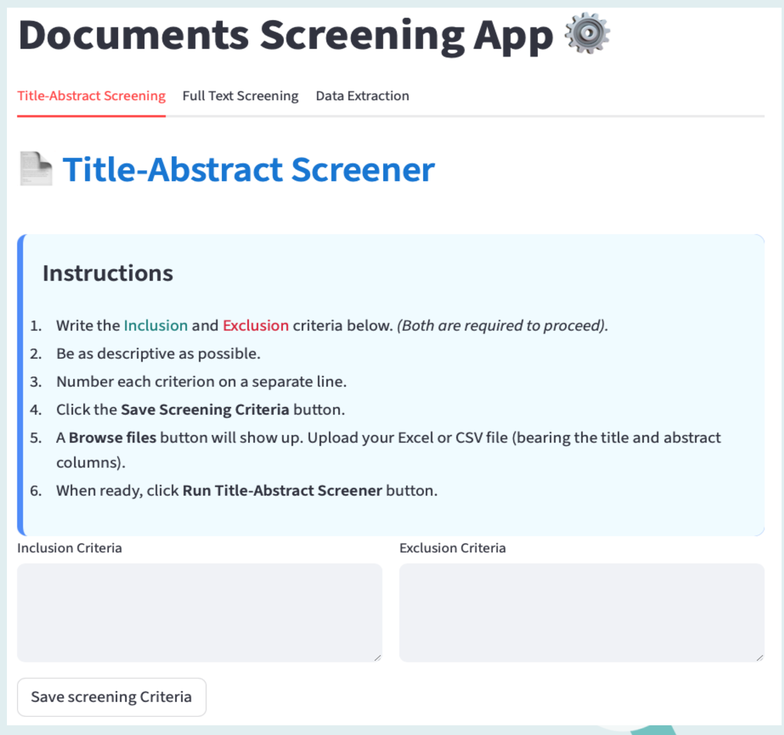

Figure 2: User Interface

Device Name	...	Developer
Flow Cytom...	...	Department...
BioFire® B...	...	bioMérieux
BioFire BC...	...	BioFire Di...
Blood-base...	...	Sadik Aref
Accelerate...	...	Accelerate...
Multiplex ...	...	Université...
Rapid Iden...	...	Hospital d...
EUCAST Rap...	...	European C...
Combined B...	...	Prenosis
Beckman Co...	...	Beckman Co...

Figure 4: Sample extracted data

METHODS

We present a pipeline that leverages LLMs to automate document screening and ranking as well as data extraction—transforming horizon scanning from a marathon into a sprint.

i. **Preliminary Filtering** – Filter the documents against a database of keywords to remove spurious studies.
ii. **Criteria & Schema Definition** – Analysts set inclusion/exclusion rules and extraction fields via a web UI.
iii. **Title - Abstract triage** – The LLM screens titles and abstract sections to quickly discard irrelevant studies.
iv. **Documents cleaning & chunking** – Clean up articles by removing unnecessary sections such as references, acknowledgements, appendices, and split the documents into smaller chunks for LLM calls.
v. **Full text analysis –** Make LLM calls to analyse the chunks against the analyst's criteria. Rank the chunks and aggregate the classifications for document-level ranking.
vi. **Data Extraction –** Analysts define their data target fields via the UI. LLM API extracts structured data into CSV for seamless downstream analysis.

RESULTS

- **Case Study** – We tested the pipeline using Sepsis & AMR-in-Sepsis publications for the period *2021–2025*; **~5200** documents averaging *18* pages per document; previously screened manually (**for 2 months 3 weeks**).
- **Screening time** – Reduced from 2 months 3 weeks to **under 7** hours **(>98%** reduction).
- **Classification accuracy** – The pipeline ranked documents with an accuracy of 90.56%.
- The pipeline flagged subtle cases by introducing ***UNSURE*** class for a review by the analysts.
- The pipeline provides justification for each ranking allowing analysts to review the process and update their inclusion/exclusion criteria accordingly.

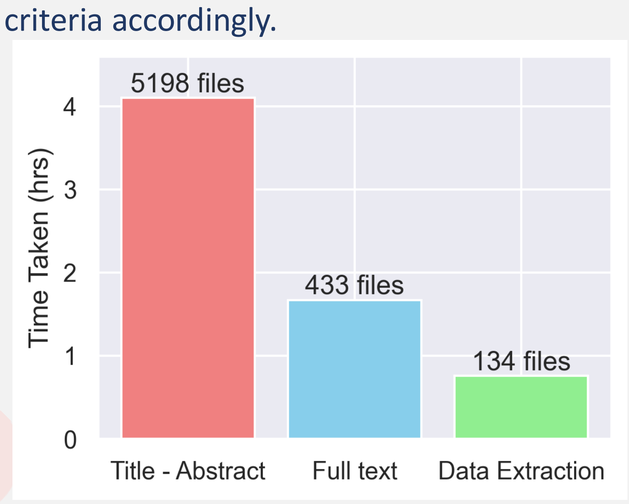

Figure 3: Pipeline Files Processing times

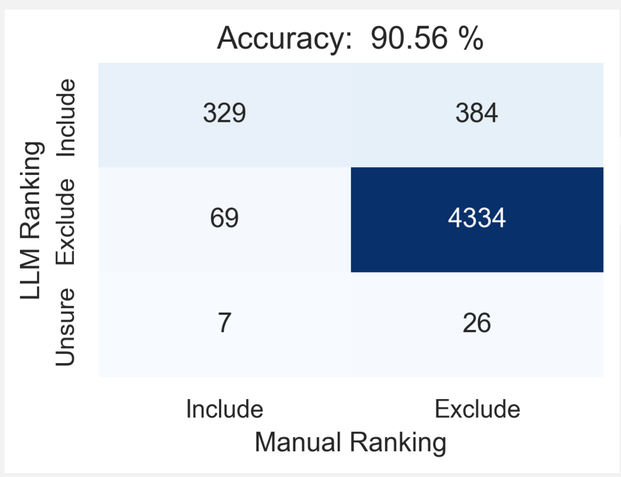

Figure 4: Evaluation of Pipeline results

WHY IT MATTERS

In this transformative era of Artificial Intelligence, adopting AI tools and methods responsibly has the potential to dramatically increase research throughput which translates to better and timely healthcare services. This tool showcases such efforts where LLMs amplify human expertise by transforming months of manual workloads into a few hours. Here are some of the specific benefits:

- **Key Impact –** Accelerates evidence-informed decision making in technology appraisal, clinical guideline updates and research. Ultimately, improving healthcare service delivery outcomes.
- Quicker actions – Fast-tracking horizon scanning delivers timely alerts on emerging treatments and safety concerns, allowing for quicker actions. This would mean more lives saved.
- **Resource Optimisation** – Frees analyst capacity for high-value interpretation, stakeholder engagement and policy development.

CONCLUSION

Leveraging large language models can revolutionise horizon scanning by transforming an overwhelming, error-prone manual process into an efficient, accurate, and scalable experience that enables researchers to focus on innovation rather than information processing.

ACKNOWLDEGEMENT

- NIHR Innovation Observatory
- HDRUK
- Black Interns Foundation

References:

1. Herbst, P., & Baars, H. (2023). Accelerating literature screening for systematic literature reviews with Large Language Models-development, application, and first evaluation of a solution. In LWDA (pp. 41-51).
2. Omelyanovskiy, V. V., Mukhortova, P. A., Zinadinov, S. I., Slabikova, A. A., Krekhtunova, L. O., Rukavitsyna, N. P., ... & Bogdanova, V. O. (2025). International experience in the application of horizon scanning systems for identifying innovative health technologies. Translational Medicine, 11(6), 546-561.
3. Michels, R. E., de Graaff, M. B., Abrishami, P., & Delnoij, D. M. J. (2024). Anticipating emerging medical technologies: The start of an international horizon scanning tool for medical devices. Futures, 156, 103326.
4. Saito, H. (2025). Sepsis in Global Health: Current global strategies to fight against sepsis. Acute Medicine & Surgery, 12(1), e70045.