

UKHSA Presents Advancing Public Health: UKHSA's application of Al in Health Security

Chair: Lee Bailey, Director of Communications, UKHSA Thursday 8 May 2025



#UKHSAPresents



Director of Communications Lee Bailey, UKHSA











Agenda

UKHSA's Strategic approach to data and AI : Prof. Steven Riley, Chief Data Officer

Al use cases:

- a. Large Language Models Tim Laurence, Principal Data Scientist
- b. Environmental Monitoring Emma Marczylo, Principal Toxicologist

Future applications of AI in Public Health: Nick Watkins, Chief Data Scientist

Q & A

Scientist oxicologist

#UKHSAPresents



Chief Data Officer Professor Steven Riley, UKHSA











Our objectives



Prepare.

Be ready for future health hazards



Respond.

Save lives and reduce harm through effective health security response

Build.

Build the UK's health security capacity

Achieve more equitable outcomes



Our data strategy



We have expertise in the complete data pipeline from acquisition, curation, processing, analysis, sharing, publishing to inform public health policy









Data products and services

Governance



Supporting health equity





Some values have been rounded.

Crown copyright



All is changing our approach to data analysis in public health





Al Strategy

•

empower the agency to embed AI through reskilling our teams to ensure safe and effective implementation

embed enterprise level Al use across UKHSA in alignment with recognised policies and risk management practices

priorities

- 🗆





develop and learn from others' uses of AI technologies that are aligned to our strategic

allocate funding to AI tools and products which tangibly improve business productivity, efficiency and data driven knowledge, policy and actions



Adopting Al

Enabling safe and responsible access to public Al tools

Evaluating enterprise level solutions for deployment at scale

Developing mission specific, public health use cases











Mission specific public health use cases UKHSA







OF COLOR OF COLOR



Evaluating Large Language Models for Public Health Use Cases Tim Laurence, Advanced Analytics

The second second second

Janus – deploying and evaluating Large Language Models on UKHSA's systems

UKHSA Network

Send text

Provide a comma separated list of the symptoms mentioned in the text. I got food poisoning at this restaurant! I had fever, diarrhoea, and some headaches!

Receive processed text

fever, diarrhoea, headaches

Evaluating Large Language Models for Public Health

Burden	Risk Factors
Diseases Symptoms Injuries	Environment Behaviour Biology
[5] Public I	Health [5]
Evalua	tions
Interve	ntions
Non-Pharm Pharmac	naceutical ceutical
[6	1

	Mistral 7B	Llama-3 8B	Llama-3.1 8B	Flan-T5 XXL	Gemma-2 27B-It	Command R	Stable Beluga-2	Llama-3 70B	Llama-3.1 70B	Llama-3.3 70B	Mistral Large*
Task Name											
NCBI Disease Extraction	0.63	0.80	0.87	0.72	0.83	0.83	0.82	0.84	0.83	0.88	0.83
GI Classification	0.86	0.82	0.84	0.89	0.82	0.91	0.90	0.92	<u>0.94</u>	<u>0.94</u>	<u>0.94</u>
GI Symptom Extraction	0.58	0.87	0.71	0.76	0.90	0.87	0.90	0.91	0.90	0.90	0.90
ICD-10 Description Classification	0.94	0.93	0.85	0.92	0.94	0.87	0.94	0.95	0.94	0.94	0.94
News Headline Classification	0.90	0.95	0.95	0.44	0.95	<u>0.97</u>	0.95	0.93	0.93	0.94	0.90
Guidance Topic Classification	0.73	0.87	0.82	0.94	0.91	0.87	0.91	0.94	0.93	0.96	0.89
BioDex Drugs Extraction	0.25	0.32	0.32	0.33	0.36	0.30	0.30	0.33	0.35	0.34	0.33
Health Advice	0.75	0.68	0.78	0.77	0.74	0.78	0.66	0.78	0.81	0.82	0.82
Guidance Recommendations	0.75	0.83	0.86	0.83	0.82	0.88	0.82	0.87	0.86	0.86	0.87
Causal Relations	0.38	0.26	0.39	0.50	0.45	0.42	0.35	0.51	0.53	0.62	0.58
PubMedQA	0.33	0.74	0.76	0.76	0.49	0.73	0.73	0.75	<u>0.77</u>	<u>0.77</u>	0.59
Contact Classification	0.24	0.23	0.26	0.28	0.42	0.36	0.46	0.48	0.49	0.53	0.22
Country Disambiguation	0.82	0.86	0.89	0.83	0.91	0.92	0.86	0.92	<u>0.93</u>	0.92	0.90
Food Extraction	0.87	0.87	0.88	0.83	0.88	0.88	0.88	0.88	0.88	0.89	0.87
MMLU Genetics	0.59	0.70	0.72	0.59	0.90	0.78	0.68	0.91	0.91	0.86	0.90
MMLU Nutrition	0.60	0.68	0.72	0.60	0.81	0.74	0.71	0.85	0.88	0.86	0.88
Mean Task Rank	9.50	8.44	6.56	7.75	5.75	5.88	7.31	3.12	2.94	2.50	5.31

Table 3: Zero-shot Results (Micro-F1 Scores) for Open-Weight Models. Bold indicates the highest open-weight model micro-F1 score on a given task. Underlined indicates the highest micro-F1 score across all models evaluated (inc. GPT-4 series for a subset of tasks). Mean Task Rank is calculated as the rank of the open-weight model on each task averaged over all tasks. Models ordered by number of parameters. *Mistral-Large was run in INT-4 AWQ

Harris et al (2024) Evaluating Large Language Models for Public Health Classification and Extraction Tasks https://arxiv.org/abs/2405.14766

Improving consistency in Public Health Guidance

Provisional Results - Write Up Under Development

Improving consistency in Public Health Guidance

Janus Guidance Checker Betav2 ()?

Text Entry	PDF Upload	Discrepancy detected
Document Title		Document Title
Risk assessment and immediate manageme high consequence infectious diseases) in ac	ent of viral haemorrhagic fevers (contact	Risk assessment and immediate management of viral haemorrhagic fevers (contact high consequence infect dataset on Thu, 09 Jan 2025 01:53)
Guidance Text Chunk		Guidance Text Chunk
An unwell patient categorised as 'minimal ris side room to limit contact until the possibilit out. The side room does not need to have de dedicated commode.	sk of VHF' should be isolated in a single y of transmissible infection has been ruled edicated <u>en</u> -suite facilities or at least a	Risk assessment and immediate management of viral haemorrhagic fevers (contact hidiseases) in acute hospitals Section 3: management of a patient categorised as 'minimal risk of VHF' Infection prevention and control measures An unwell patient categorised as 'minimal risk of VHF' should be isolated in a single side room to limit contact infection has been ruled out. The side room should have dedicated en-suite facilities or at least a dedicated or all staff should be using SICPs. Additional precautions (TBPs) should be added for splash-inducing procedure. The medical procedures that are considered to be aerosol generating and associated with an increased risk of England's 'National Infection Prevention and Control manual for England (NIPCM)' as: • awake bronchoscopy (including awake tracheal intubation)
User input text	here omit Success	LIM Reasoning Does the above text conflict with the input text? Yes. TEXT 1 states that the side room for a patient categorised as 'minimal risk of VHF' does not need to have dedicated commode, whereas TEXT 2 states that the side room should have dedicated en-suite facilities or at contradiction in the recommendations for infection prevention and control measures for patients with minimal Reference of the state of th

Hypothetical Example - Not a change that is being co

Feedback Documentatio	n
tious diseases) in acute hospitals (added to	
igh consequence infectious	1
et until the possibility of transmissible commode.	
es and AGPs.	
of respiratory transmission are listed in NHS	
e dedicated en-suite facilities or at least a t least a dedicated commode. This is a direct al risk of viral haemorrhagic fever.	
Result 1 of 16 Previous Next	
Filter conflicts: Displaying all results Show/Hide Audit Options	
nsidered	

UKHSA's Al application in Health Security: Real-time pollen (bioaerosol) monitoring **Emma Marczylo**

Analysis

PolenoJupiter instrument: Pollens

🔟 pinaceae (100% ?)

🔀 🔟 quercus (100% 🖓)

Pine

Oak

Future development: fungal spores

In planning

Status:

Pollen monitoring since April 2024, and data analysed as a collaboration alongside Met Office, and Universities of Birmingham, Manchester and Leicester

Next:

- Bring fungal spore identification on-line
- Validate system for pollen and fungal spore monitoring as part of the new Health Protection Research Unit with the University of Leicester and the Health and Safety Executive (Apr-2025-Mar-2030)
- Establish UK Poleno user group and work towards building a UK real-time monitoring network that can feed data into the European network
- Develop platform to share monitoring data with the public & health care professionals.
- Inform further experimental work in airway models on fungi of concern •
- Inform other areas of work such as impact of changing climate on bioaerosols/aeroallergens and thunderstorm asthma event preparedness

Roadmap to a UK real-time monitoring network

Future applications of Al in Public Health Nick Watkins, Chief Data Scientist

Where will we apply AI in the future?

Unlocking insights with earth observation data

Integrating AI and Genomics for public health

Unlocking Insights with Earth Observation Data

- Earth observation data helps identify environmental factors affecting health.
- Satellite imagery provides realtime information on disease outbleaks.
- Data visualization enables better decision-making for public health interventions.
- Machine learning algorithms identify patterns in environmental data.

UK Health Security Agency

Al can identify structures of interest from aerial imagery

Rapid identification of structures and objects of interest is crucial for disease surveillance and incident response.

Large-scale surveys are impractical.

- Advances in highresolution imagery and computer vision offer new solutions.
- We can quickly and efficiently detect 'new' items of public health interest

Integrating AI and Genomics for Public Health

the second second

Machine learning models analyse genomic data to detect disease outbreaks

Al tools identify mutations and patterns in pathogens' genetic sequences

Genomic data enhances predictive models for public health responses

Integration allows for tailored vaccination strategies based on genomic analysis

Real-time monitoring of genomic changes supports swift public health interventions

Inferring zoonotic risk using genomic **Machine Learning**

- Pathogen genomics is a core component of the UKHSA mission
- Machine learning can be trained on genome sequencing data to predict future zoonoses
- Working with the University of Glasgow, we have trained 5 machine learning models on ¬ 50,000 avian influenza genetic features
- Aim to identify which measures of genome composition and functional protein chemistry are the strongest predictors of zoonotic capability

W Health Security Agency (EDAP)

EDAP is UKHSA's strategic data platform. It brings key UKHSA data and tooling in one safe & secure place and enables data analysis at scale and at pace.

Objectives

Build the UK Health Security capacity by offering a secure and scalable data and analytics platform capable of lawfully and securely ingesting, transforming and presenting health data to the experts skilled at interpreting and analysing it, generating new insights to inform public health action.

Enable national and international **collaboration on public health** data by offering a single and secure shared system with a standardised set of tools.

Provisioning **pre-processed and transformed data marts** that can be used to create high quality data products at pace.

Key benefits

- Multiple data sources held in a secure single platform
- Improved and standardised analytics tooling
- Increased confidence in the provenance of data
- A clear curated data catalogue
- Making data an enhanced enterprise asset
- Evidence of compliance with laws and regulations

With Health Security Agency Our responsible adoption of Al supports the management of associated risks

Data privacy concerns regarding sensitive health information

Cybersecurity threats can compromise AI systems and data integrity

Bias in Al algorithms may lead to unequal health outcomes

Al Enabled Public Health Agencies

Global initiatives in this field include GI-AI4H which aims to:

enable, facilitate, and implement AI in healthcare, establishing governance structures, policies, and standards for AI adoption in health

PHIs globally are already using AI for data analysis and disease surveillance. In future, AI has the potential to:

- Enhance data-driven decision making for health interventions
- Enable real-time monitoring of health trends through integrated systems
- Drive predictive analytics to anticipate outbreaks and public health needs
- Streamline communication and response efforts

Acknowledgements

Safe and effective adoption of AI is a cross-UKHSA effort. We would like to thank the following for their contributions

D & 198 Y & DO M Broke Wanter And mouse and the ballance V Marthe Contract grant the state of the

- **Advanced Analytics**
- **AI Steering Committee**
- **AI** Taskforce
- **All Hazards Intelligence National**
- **Central Guidance Team**
- **Chief Data Officer, Private Office**
- **Data Science Hub** •
- **Data Strategy**
- **Health Analytics and Automation**

- **Health and Safety Executive**
- **Health Protection Research Unit**
- **High Performance Computer Cloud & DevOps** Team
- **Strategic Delivery Team**
- Toxicology
- **University of Glasgow**
- **University of Leicester**
- **University of Oxford**

Q&A

UKHSA Presents Thank you

Find out more: <u>UKHSA Presents - GOV.UK</u>

#UKHSAPresents