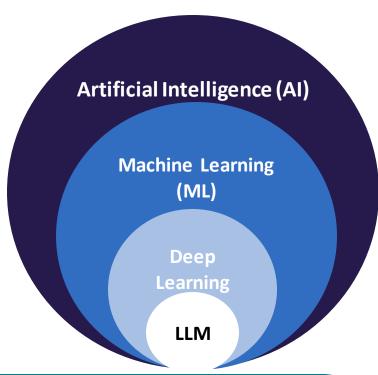


Developing Artificial Intelligence (AI) and Large Language Models (LLM) for scientific computing applications

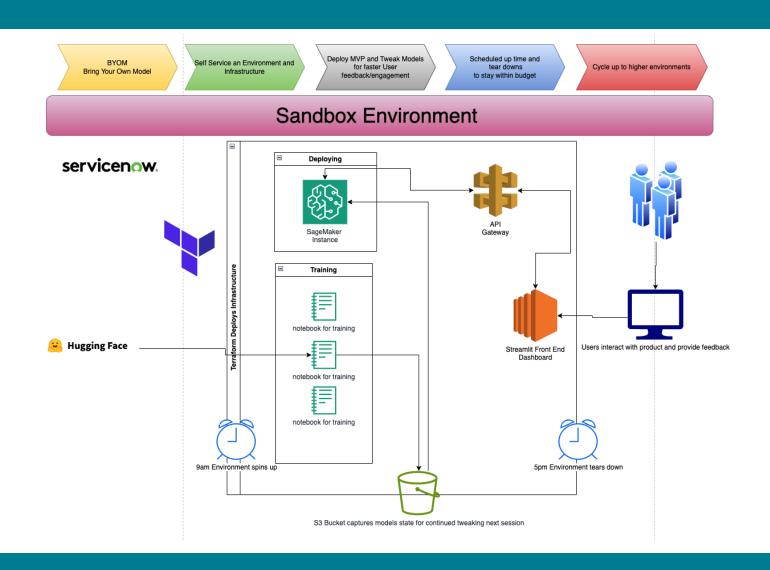
UKHSA Conference 2023

Introduction to Large Language Models (LLM)

- What are LLMs?
 - A Large Language Model (LLM) is a type of Artificial Intelligence (AI) algorithm that uses deep learning techniques and massively large data sets to understand, summarise, generate and predict content.
- How DevOps are supporting new Gen Al Products.
- What is the infrastructure behind it?
- How are we innovating at UKHSA?

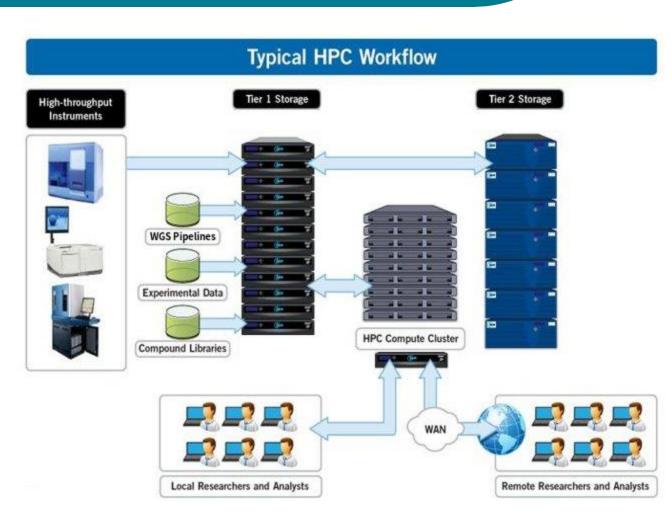


Infrastructure



UKHSA High Performance Computing (HPC)

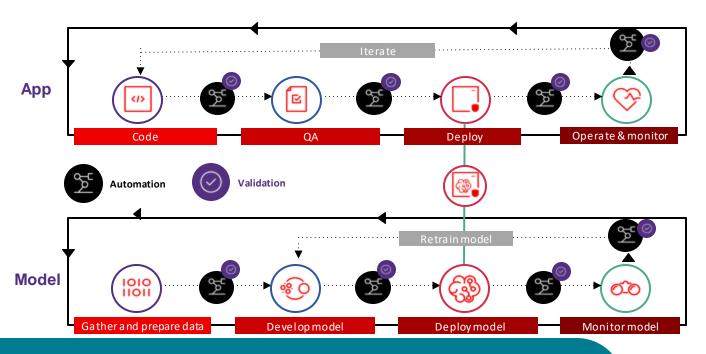
- What is HPC?
 - Supercomputing
 - Horizontal Scaling
 - GPU vs CPU
- UKHSA HPC Capabilities:
 - 2 large clusters: Porton and Colindale
 - 100k+ GPU cores on-prem capabilities optimised for Al and Machine Learning frameworks (PyTorch, Tensorflow & others))
 - Elastic HPC cluster on Azure possibility to scale (was able to be scaled 10x during 2022/23)
 - Covid pandemic
 - ~8,000 pathogen genomes per week



Making HPC Accessible to the Scientific Community

HPC is normally a complex platform to interact with – our platform requires us to:

- Aim to support and facilitate the learning curve.
- Provide front-end layers to abstract the complexity from users that come from pure scientific backgrounds to allow them to leverage the power of the HPC cluster (to further this, OpenShift layer can help make a simple front-end).
- OpenShift Al
 - Containerised AI workloads contributed by the wider scientific community.
 - Open-Source technologies enhanced with enterprise level support.
 - Cloud, on-premises, and the edge.

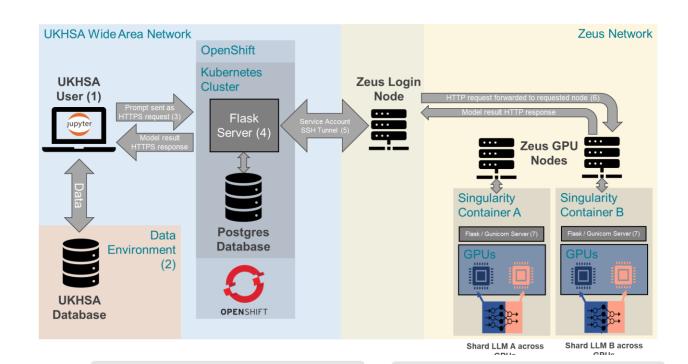


Janus Project

- Rise of LLMs (ChatGPT)
- Janus gateway LLM accessed via an API (OpenShift)
- Janus Python package
- Shared pool of compute resources
- Connect to UKHSA HPC infra
- Existing on-prem infra with A100 GPUs and NVLink Interconnect

Contributors

- Advanced Analytics Team: Leo Loman, Timothy Laurence, Joshua Harris
- Use case leads: Tim Laurence, Toby Nonnenmacher, Harry Long
- Technology: Sam Morris



Extraction

Text: I got food poisoning at this restaurant! I had fever, diarrhoea, and some headaches

fever diarrhoea headaches

Classification

Question: Does the following text refer to gastro-intenstinal illness, answer only with 'yes' or 'no'?

Text: I got food poisoning at this restaurant! I had fever, diarrhoea, and some headaches!

Janus use cases and validation



Legend

Key drivers of health security

Expert validation sets

Intervention validation:

- Guidance
- Adverse event reporting

Public health interventions:

- Pharmaceutical
- Non-pharmaceutical



Risk factors:

- Environmental / geographic
- Behavioural
- Genetic / biological

Risk validation:

- Country classification
- Sexual behaviour extraction
- Foods extraction

Diseases and symptoms:

- Infectious diseases
- Symptoms of inf. diseases

Disease validation:

- Pubmed disease extraction
- · GI illness classification
- · GI symptom extraction
- ProMed / News extraction
- quoref extraction

Advanced Analytics is taking a validation first approach to reassure users that any tools built on top of Janus should meet the required performance.

We are working with expert collaborators throughout UKHSA to build a range of validation sets for use cases across the breadth of UKHSA's activity.

We are also bringing in relevant published NLP benchmarks (like Pubmed) to ensure further testing of these models.

Future Planned Work



AI workload support

Support AI workload requirements on HPC and Kubernetes/OpenShift platfor ms.

e.g., hardware acceleration, GPU

Operator



Platform for AI-Enabled apps

Provide a consistent, hybrid/multicloud application platform for Data Scientists to build, train and deploy Al-enabled applications across multiple environments.

e.g. OpenShift AI (on AWS, Azure, GCP)



Working closely with the UKHSA community

Provide a unified platform for data scientists and developers to enable innovation in new Al fields and accelerate adoption of existing tools

e.g. Image classification, Paolo Alto Al-enabled firewall monitoring