NHS
Fellowship in Clinical AI
NHS Digital Academy Programme

Supporting a Global Community of Practice through Evaluation Clinics

INTRODUCTION

- Artificial Intelligence has the potential to revolutionise healthcare, but it is essential these technologies are safe, effective, and equitable, for both patients and society to benefit.
- The University of Birmingham (UoB) supported an established global Community of Practice (see Figure 1) by launching evaluation clinics to provide expertise to research teams who were undertaking evaluations of AI technologies involving Large Language Models (LLMs).
- Each evaluation clinic session was designed to provide tailored guidance and support from academics, to enhance the quality of the evidence base and strengthen the evaluation methodologies being used in AI research.

METHODS

- Figure 2 describes the evaluation clinic timeline that was conducted.
- Teams interested in receiving support from an evaluation clinic were asked to complete an online survey, where information was gathered on the product, the challenges that were faced, the objectives wished to be achieved and the timeframe for support.
- Figure 3 describes the support the UoB team were able to provide.
- Following a review of all survey responses (n=15), eight teams were notified of being successful for support in an evaluation clinic.
- All identified teams were invited for a virtual pre-meetings via Microsoft Teams as an information gathering exercise, to ensure the UoB team could tailor the guidance provided to meet the specific needs and objectives of each team.
- The virtual evaluation clinics then followed the pre-meeting.
- An online survey was requested to be completed by teams following the evaluation clinic to establish the usefulness of the process.

CONCLUSIONS

This opportunity generated interest from a wide range of teams across the globe to gain support for individual research priorities. Mutual benefit was generated by **producing a useful resource for the Community of Practice** where collective needs of all engaged teams were identified.

The use of the evaluation clinics allowed diverse teams to **understand and work towards the implementation of best practice** to evaluate AI technologies on an international scale.

The request for support from all teams to **understand how to assess the quality of responses generated by an LLM** identifies a specific knowledge gap when evaluating AI technologies.

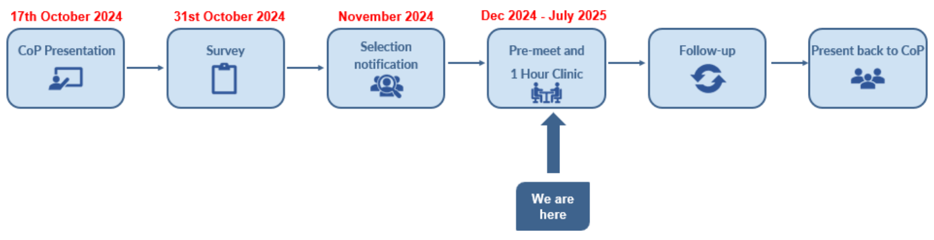


Figure 2: Evaluation clinic timeline

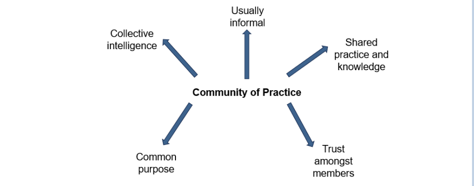
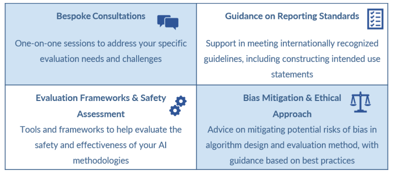


Figure 1: What is a Community of Practice?

RESULTS

- As of June 2025, virtual evaluation clinics have been performed with 5 different teams spanning the globe, including from Nigeria, India and Bangladesh.
- All teams (n=5) requested support to systematically assess the quality of the responses generated by an LLM across multiple axes (e.g. knowledge recall, logical reasoning, possible extent of harm etc).
- Guidance on assembling, training, and managing evaluation panels to ensure reliable assessment of an LLM was also shared with all teams (n-5).
- Two teams requested guidance on the regulatory requirements of their product. Additional support was provided to one team (n =1) on translation issues that can occur when using LLMs.
- One response was received to the online survey following the evaluation clinic. The clinic was evaluated favourably and described as "extremely useful to our team."



Bespoke Consultations One-on-one sessions to address your specific evaluation needs and challenges	**Guidance on Reporting Standards** Support in meeting internationally recognized guidelines, including constructing intended use statements
Evaluation Frameworks & Safety Assessment Tools and frameworks to help evaluate the safety and effectiveness of your AI methodologies	**Bias Mitigation & Ethical Approach** Advice on mitigating potential risks of bias in algorithm design and evaluation method, with guidance based on best practices

Figure 3: Support provided in the evaluation clinic

Authors: Natalie Davison, Vaishnavi Menon, Jaspret Gill, Alastair Denniston

UNIVERSITY OF BIRMINGHAM
NHS University Hospitals Birmingham NHS Foundation Trust
PATH