Summary of the UK Open Multimodal AI Network (UKOMAIN)

UK Open Multimodal AI Network

Engineering and Physical Sciences Research Council

Funded Value: £1.82 million
Funded Period: January 2025 – December 2027
Funder: EPSRC
Project Reference: UKRI396
Principal Investigator: Haiping Lu
Co-Investigators: Nicola Morley, Nataliya Tkachenko, Peter Charlton, Dezong Zhao, and Yao Zhang

---

Multimodal Artificial Intelligence (AI), which integrates diverse data modalities such as text, image, and sound, is transforming our interaction with technology and data. A 2024 MIT Technology Review Insights article predicts the global multimodal AI market will grow at an annual rate of 32.2% from 2019 to 2030, reaching US$8.4 billion. This growth, primarily focused on multimedia data, only scratches the surface of multimodal AI's potential beyond common multimedia modalities. By leveraging diverse data sources and domain expertise holistically, multimodal AI has a crucial role to play in addressing Tomorrow's Engineering Research Challenges (TERCs), from health and wellbeing, transportation systems, and robotics to materials discovery, space research, nature-based engineering, global engineering solutions, and responsible engineering.

The 2022 EPSRC TERC report identified the above eight key challenge areas, all requiring cross-disciplinary collaborations and integrated solutions. However, current research efforts are often fragmented and siloed, with researchers from different TERC areas working within disconnected boundaries of their own disciplines. Solutions developed in one discipline face significant barriers to being transferred to other disciplines, although multimodal AI solutions often share common challenges in their development pipeline. This highlights the need for findable, accessible, interoperable, and reusable solutions to benefit multiple disciplines.

Over the past two years, we have created a vibrant multimodal AI community in the UK, comprising over 200 researchers and practitioners from diverse scientific and application domains. This laid the foundation for addressing shared multimodal AI problems in TERCs collectively, productively, and sustainably.

The UK Open Multimodal AI Network (UKOMAIN) aims to connect multimodal AI stakeholders and solutions across disciplines to address such problems in TERCs holistically and systematically for sustained and thriving growth. Our objectives are to:

1) Establish and expand a diverse and interdisciplinary network of researchers, industry partners, policymakers, and end users to promote collaboration and inclusivity, integrate diverse perspectives and contributions, and unleash Multimodal AI's potential in critical TERC areas such as health and wellbeing, transportation systems, and climate change.

2) Build a knowledge exchange and collaboration platform through a streamlined governance structure, carefully designed incentives, and well-connected engagement activities to foster a continuous flow of innovative solutions and interdisciplinary research.
3) Fund small-scale, high-impact feasibility studies to address specific TERCs that share common multimodal AI challenges and develop novel engineering approaches beyond their traditional

boundaries, leveraging expertise from diverse related domains and generating preliminary outcomes for larger-scale projects.

4) Engage industry and policy stakeholders to align our research with real-world needs and relevant policy initiatives that are key to deployment-centric research and potentially inform policy development.

5) Embed principles of environmental sustainability and social responsibility in network activities and funded projects to invest in our future and integrate diverse perspectives.

6) Enhance our research capabilities and outputs by leveraging research facilities and expertise across the network, providing training and upskilling opportunities and promoting open science practices.

To achieve these objectives, we will establish eight thematic and three community-focused special interest groups, organise workshops, hackathons, tutorials, sandpits, and grant-writing retreats, and build open-source resources. We will fund cross-TERC feasibility studies, community-led network activities, and early career researcher development with carefully designed incentives aligning with our objectives. These and related activities will transform cross-cutting multimodal AI research towards deployment-centric and user-centric innovations addressing the multifaceted TERCs and delivering much-needed economic and social benefits.