



Connecting UK Research to Responsible AI

Gopal Ramchurn, CEO

Techworks AI Launch, Bletchley Park, 16th April 2024



RAI UK Team



RAI UK Leadership and Delivery Teams



Muffy



Tom



Gopal



Wendy



Kate



Gina



Jack



Maire



Mac



Matt



Joel



Elvira



Caitlin



Prokar



Sana



Shoab

Why should technologists care About Responsible AI?

- New cyber vulnerabilities in systems
- more complex systems for humans to manage with one or multiple agents
- Fail with no explanation and no measures of recourse
- Trillions of data points at machine speed that humans can hardly comprehend, making safety critical systems less resilient.



Predictive Policing

© Predpol



Warehouse Robots



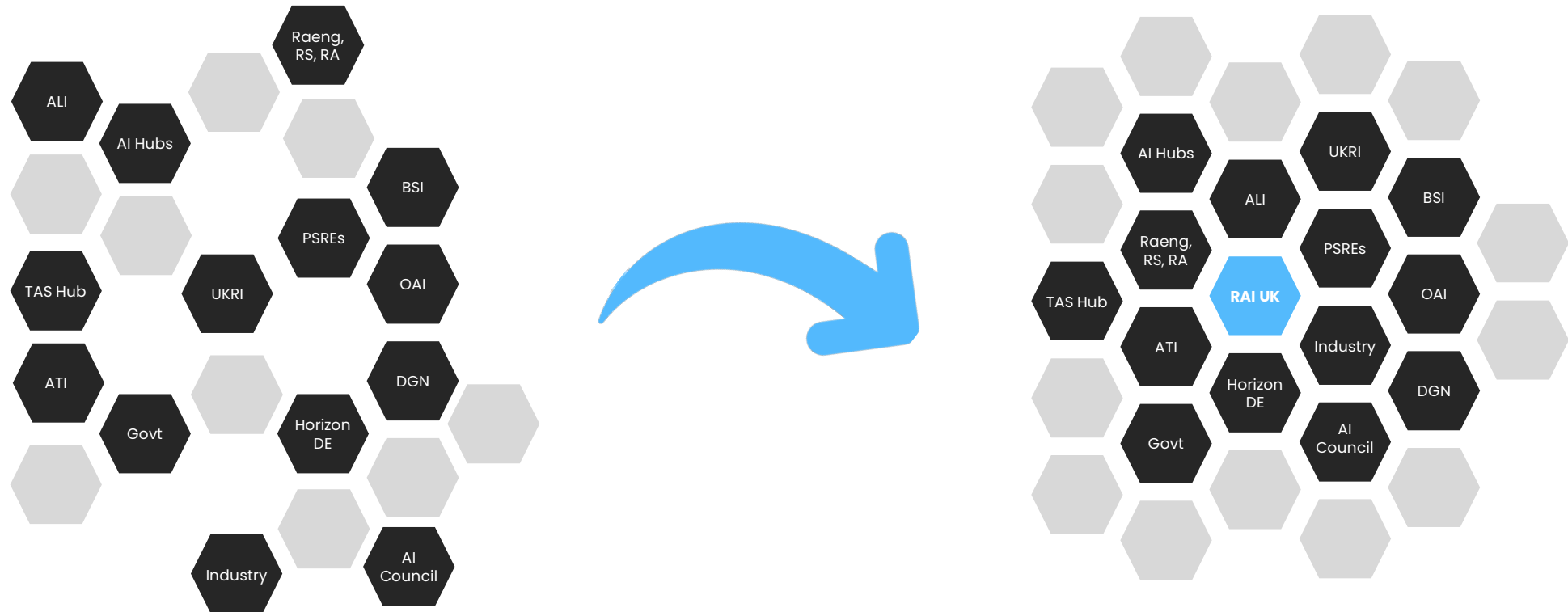
Autonomous Vehicles



Facial Recognition

It doesn't mean we can or should fix all these issues ourselves. But surely we can help government, industry, and the public prepare for such challenges.

Creating an Ecosystem... ...in a Fragmented Landscape



Our Vision



RAI UK will enable
Responsible and
Trustworthy AI to power
benefits for everyday life



Our Mission



Build a national AI ecosystem, where all voices are heard, Respected and debated, regardless of seniority. Or loudness.

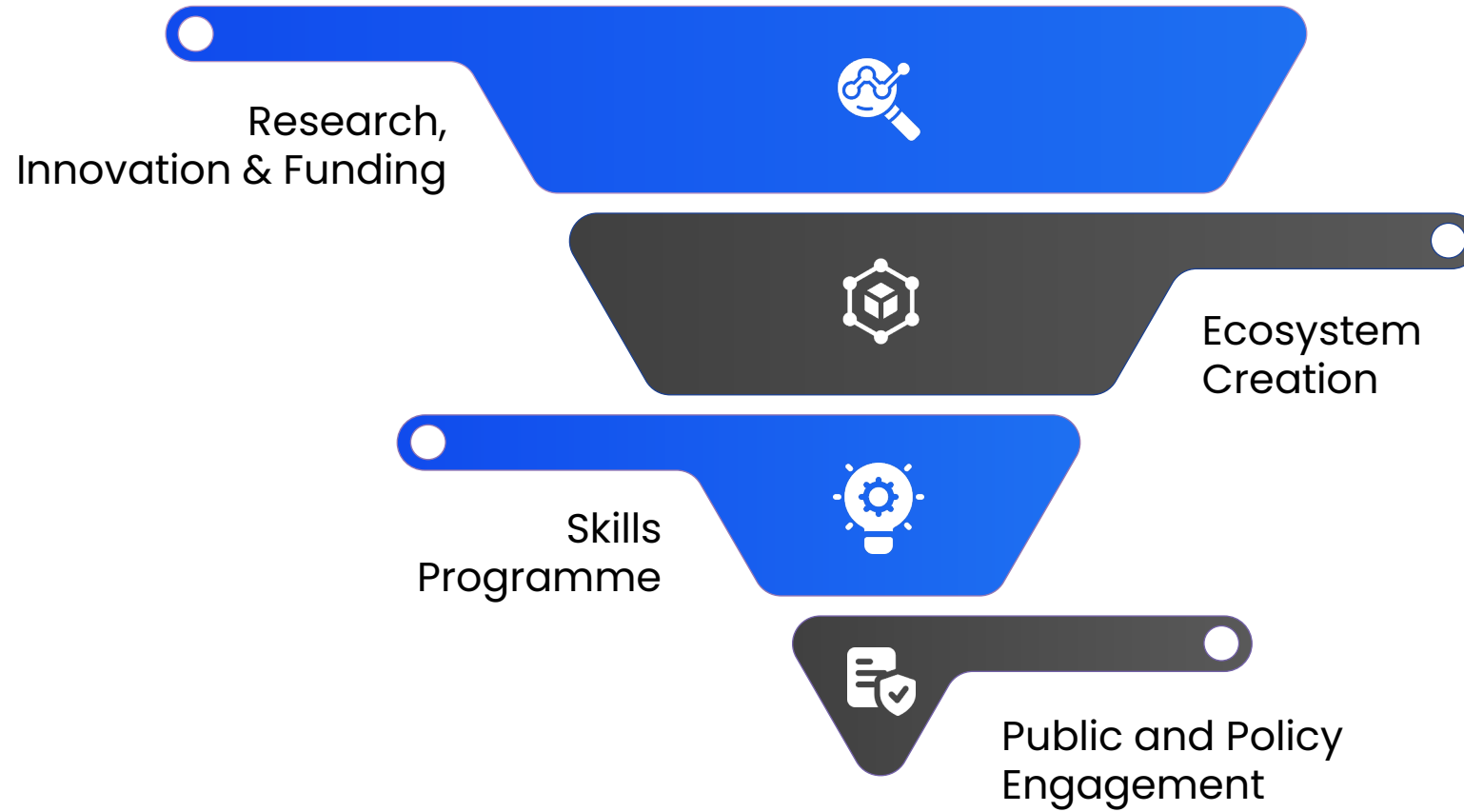


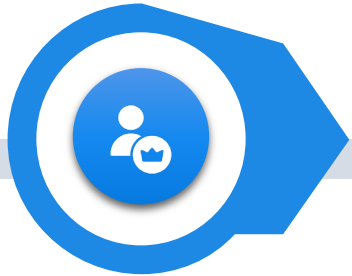
Deliver research that addresses societal and economic challenges.



Develop national conversations around AI – informed by research, not hype.

Our Work Covers **Four Areas**





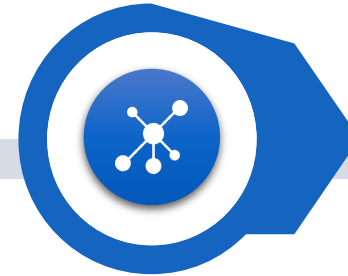
TRUSTED EXPERTISE:

RAI UK provides **research-driven expertise** on Responsible AI building on over 20 years of interdisciplinary research. Our 20-strong research team is ready to work with government and industry to deliver advice and undertake research on emerging challenges.



RESEARCH FUNDING:

RAI UK is deploying **more than £15m into research projects** to fill the gaps in the RAI research landscape and amplify the work of aligned programmes. Our research projects will involve over 300 researchers by end of 2024.



CONNECTOR & CONVENER:

RAI UK **convenes working groups, events, workshops**, and other activities to glue the AI Ecosystem in the UK and ensure the community is engaged internationally to construct a Responsible AI future.

Our Guiding Principles

1. Convening, not competing
2. Multidisciplinary by design
3. Committed to excellence
4. Failing better — responsibly
5. Fair and transparent
6. Committed to public benefit



WHAT ARE
WE DOING?

Highlights in Year One



Community Building

4 Townhalls in London, Glasgow, Belfast, Cardiff

22 Roundtables to influence research priorities

AI Hubs and CDTs



AI Fringe

Organised events at the AI Fringe in 2023,

Outcome: recently published a white paper.



Funding Research

7 Impact Accelerator Projects

11 International Partnerships

Coming soon: 2-3 Keystone Projects



Webinar Series

Series of recurrent webinars presenting our funded projects and latest initiatives from the network.

Community Building



Key areas discussed at Townhalls included:

	AI and Sustainability		Financial Services
	Creative Industries		Manufacturing
	Defence and Security		AI and Social Science Research
	Health and social care		Large Language Models (LLMs)
	Future of work		Creative Roundtables

Responsible AI-powered organisations and economies



- Organisations and economies are rapidly evolving due to new ways of working in response to AI.
- Responsible AI practices need to **include systems-based approaches, beyond testing of individual AI algorithms**, to evaluate the wider implications of AI deployment into complex human/AI systems.
- This includes ensuring that innovators understand the overall consequences and measures, such as the **reskilling and upskilling** that will need to be developed. What are the responsible AI principles appropriate for sectors (e.g., healthcare and the justice system) and contexts (e.g., where there may be high consequences of failure)

Addressing harms, maximising benefits of AI



- How AI is framed and applied introduces **new opportunities but also trade-offs** for individuals, industries, and societies where the potential benefit may be outweighed by negative impacts on a wide range of issues.
- This includes issues such as **privacy, bias, accessibility, labour rights, social justice, and sustainability** (of people, organisations, and the environment). Many of these trade-offs have direct legal ramifications.
- This introduces the need for **deployment, validation, provenance, and auditing regimes** for AI, so decision makers can thoroughly understand and manage the limitations of AI systems to ensure they are safe, ethical, and beneficial, while simultaneously highlighting where further development is needed to extend their scope.
- Dimensions **of AI safety** (e.g., ensuring the system functions as intended with regards to ethics, policy, legal and technical aspects), **AI security** (e.g., ensuring the system is robust to malicious interference) and the countering of **AI misuse** offer many open questions to be addressed. What is needed to be in place so that AI works for the benefit of people and societies while harms are minimised?

AI Law, governance, and regulation in an international context









- There are few established **routes to control, transparency and redress for users** in relation to automation in digital public services, as well as crucial private sector interactions. For example, the rise of generative AI has posed significant (though not unique) challenges to the creative industries.
- Currently very different approaches to **AI governance are taken by the EU, China, the US, and elsewhere**, a global debate to which the UK must make a strong contribution or risk isolation.
- While longer-term AI safety issues have recently been highlighted by the UK government, **short to medium term risks** remain outstanding.
- What are the approaches that promote **trust, provide fairness, and accountability for users**, and provide certainty for international commerce?

Our Working Groups



WGs are made up of members that include academics, business representatives, government officials and other key stakeholders. WGs contribute to the AI UK Leadership Team's strategic plan A WG will address at least one of the following challenges:

SOCIETAL	Existing or emergent social or economic challenges that are either instigated by the use of AI or may be addressed through the use of AI.
DOMAIN	Domains where AI is increasingly being used.
RESEARCH	Emergent research themes that demand a concerted approach.

 Defence and Security	 Health and Care	 Law and Regulation	 Public Participation	 Creative Industries	 Future of Work
Chair: Muffy Calder	Chair: Prokar Dasgupta	Chair: Sana Khareghani & Rachel Coldicutt	Chair: Jack Stilgoe	Chair: Kate Devlin	Chair: Gina Neff

AI Fringe

AI for everyone

30 October – 3 November 2023

Perspectives from the AI Fringe



Key Discussion Points:

- Use and Abuse of Generative AI
- AI in Policing
- Contributions to a healthy information ecosystem
- Ensuring Sustainable AI development



Projects working with Partners

▶ **IMPACT ACCELERATOR PROJECTS**

Maximise impact from existing responsible AI research. Rapidly realise benefits for the economy, society, culture, policy, health, the environment and quality of life.

▶ **INTERNATIONAL PARTNERSHIPS**

Develop strategic collaborations with world-leading research organisations. Ensure society deploys and uses AI in a responsible way, beyond national boundaries.



Impact Accelerator



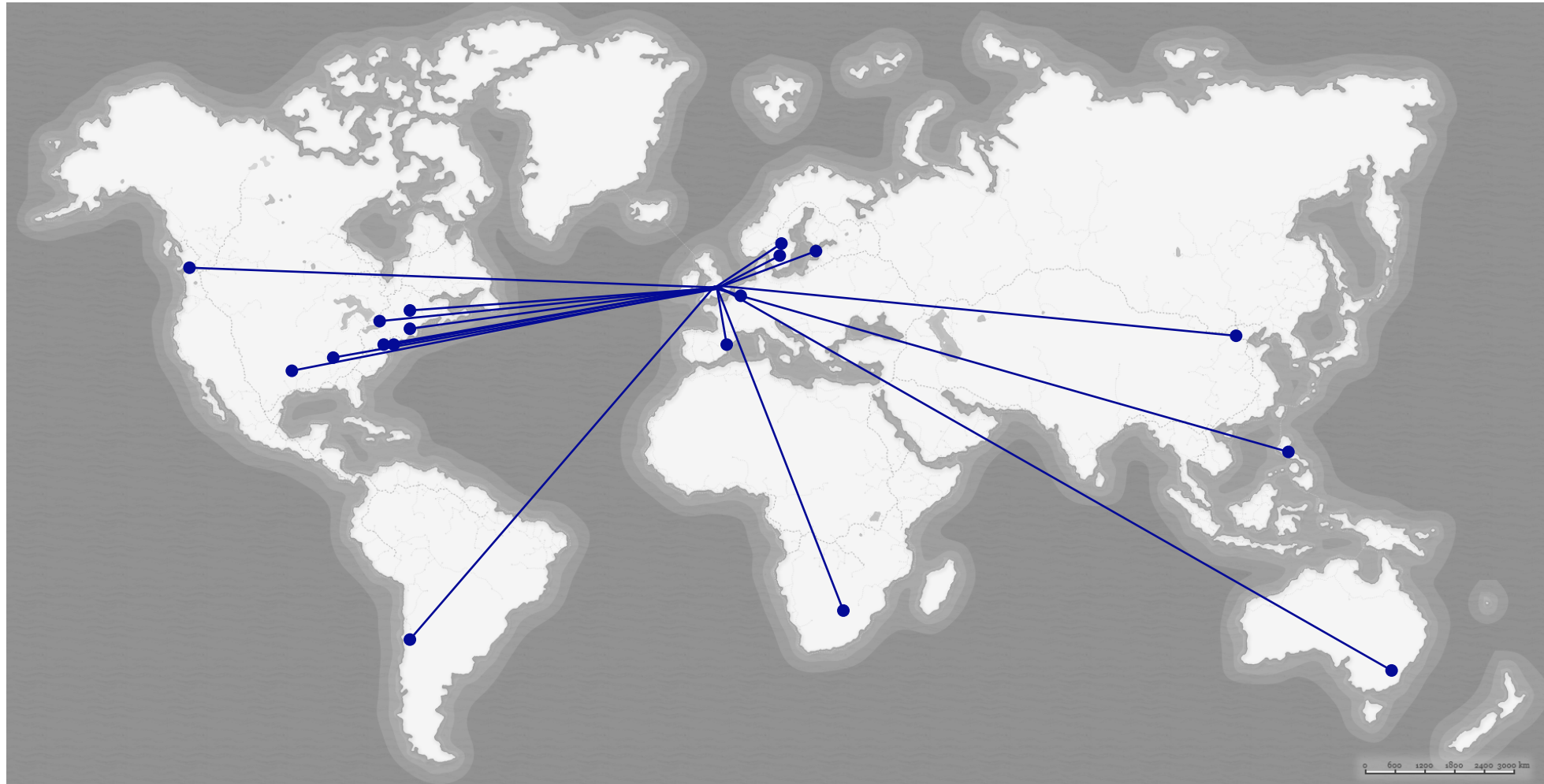
Equality by Design, Deliberation and Oversight	Karen Yeung
Automated Empathy – Globalising International Standards (AEGIS)	Andrew McStay
RAISE – Responsible generative AI for SMEs in UK and Africa	Bernd Stahl
Amplify: Empowering Underserved Voices in Spoken Language Interaction	Simon Robinson
RAKE (Responsible Innovation Advantage in Knowledge Exchange)	Marina Jirotko
AIPAS: AI Accountability in Policing and Security	Babak Akhgar
SAGE-RAI: Smart Assessment and Guided Education with Responsible AI	John Domingue

International Partnerships



Harnessing AI to enhance electoral oversight	Sam Power
The human role to guarantee an ethical AI for healthcare	Raquel Iniesta
Transparency Regulation Toolkits for Responsible Artificial Intelligence	John Downer
For the FATES of Africa: A co-developed pipeline for responsible AI in online learning	Nora McIntyre
TAS Hub and Good Systems Strategic Partnership	Helena Webb
Exploring Fairness and Bias of Multimodal NLP for Mental Health	Rafael Mestre
Disruption Mitigation for Responsible AI	Radu Calinescu
AI Regulation Assurance in Safety-Critical Systems	Jennifer Williams
Understanding Robot Autonomy in Public	Stuart Reeves
Responsible AI international community to reduce bias in AI music generation and analysis	Nick Bryan-Kinns
Responsible AI networks for industries and governments in developing countries	Nestor Castaneda

International Partnerships



Coming Soon: Keystone Projects



- Flexible, consortia-led research
- Focus on ensuring AI technologies are designed, deployed and used responsibly within societies
- Aimed at world-leading researchers from all disciplines
- Up to £3.5m available for 4 years
- Now concluding a 5-month process – **to be announced at CogX in LA**

How to work with our research team

We have **20 researchers** in the RAI core team. These include experts in deep learning, generative AI, social psychology, arts and humanities, human-computer/robot interaction, multi-agent systems, education, technology policy.

We have also funded **18 teams** across the UK working locally and internationally on RAI.

Pose a
challenge

Offer a platform

Attend our
partner network
events

Email: info@rai.ac.uk

OUR PARTNER NETWORK



What is it?



A catalyst for our Research.

Feed in challenges and translate research into real world impact.

Our Partner Network will:

...connect the AI ecosystem both nationally and international

...create and operationalise responsible AI research and innovation

...address skills gaps in the AI ecosystem

Why we want to engage with you



Partners bring:

...real-world perspectives on responsible and irresponsible applications of AI.

...case studies that present new challenges for users, organisations and society.

...opportunities to draw on early research and translate into the real world.

...promote, shape and share Responsible (Research and) Innovation practices.

...co-design of toolkits and methodologies for responsible and trustworthy AI.



How does it work?



Sign up to be an RAI Partner.
We want to engage with all non-academic stakeholders.

Three Partnership Tiers:

TIER 1

Connect & Participate

CONTRIBUTIONS

- Invest with us
- Liaison
- Provide case studies

BENEFITS

- Quarterly briefings
- Access to early drafts of white papers, thought leadership pieces.
- Speaking slot

TIER 2

Research Collaborations

CONTRIBUTIONS

- Provide case studies
- Provide datasets
- Work with our researchers

BENEFITS

- Direct involvement in the research
- Early insights into research outputs

TIER 3

Keep in Touch

CONTRIBUTIONS

- Promote RAI UK's brand & principles
- Invite us

BENEFITS

- Invitations to closed roundtables
- Logo on website
- Cross-promotion

Ways to get involved



Collaborate with our Researchers:

- ...present your challenges at our roundtables today – influence our sandpit workshops
- ...co-create research to address these challenges at our sandpits
- ...secondments (inward and outward) to share knowledge, experience and skills

Support the delivery of RAI UK's vision:

- ...co-fund research or outreach activities with us
- ...work with us to draft white papers/responses to consultations

How we can help you

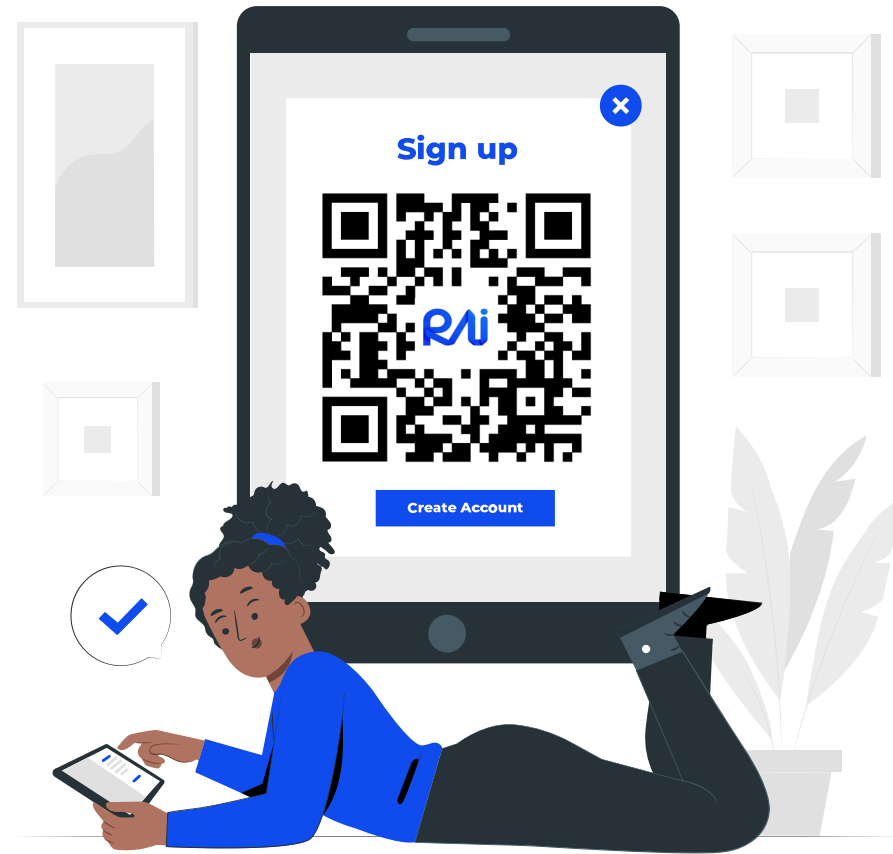


We will:

- ...develop and maintain a repository of frameworks (in progress)
- ...organise regular events to share latest research findings
- ...host events that bring together experts from different sectors
- ...facilitate connecting you with the academic community
- ...provide a talent pipeline



Sign up to the Partner Network





THANK YOU!

FOR WATCHING

