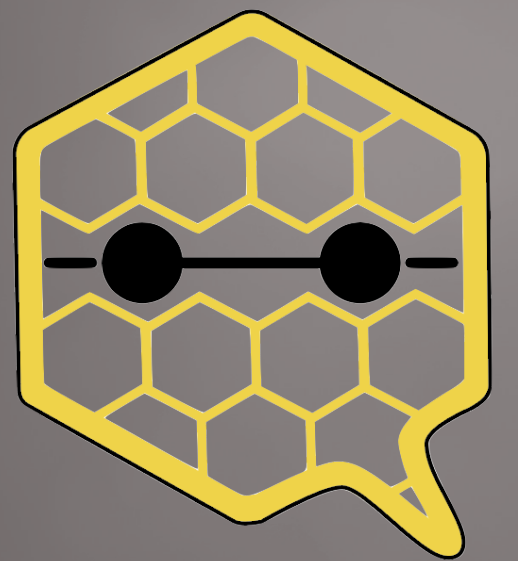


Human Interactions, Values, and Experiences with Conversational AI: A Research Agenda



Many Voices, Shared Intelligence.

The HIVE CAI Research Lab, led by Dr. Malak Sadek at Cambridge's Centre for Human-Inspired AI, is an interdisciplinary research group built around one question:

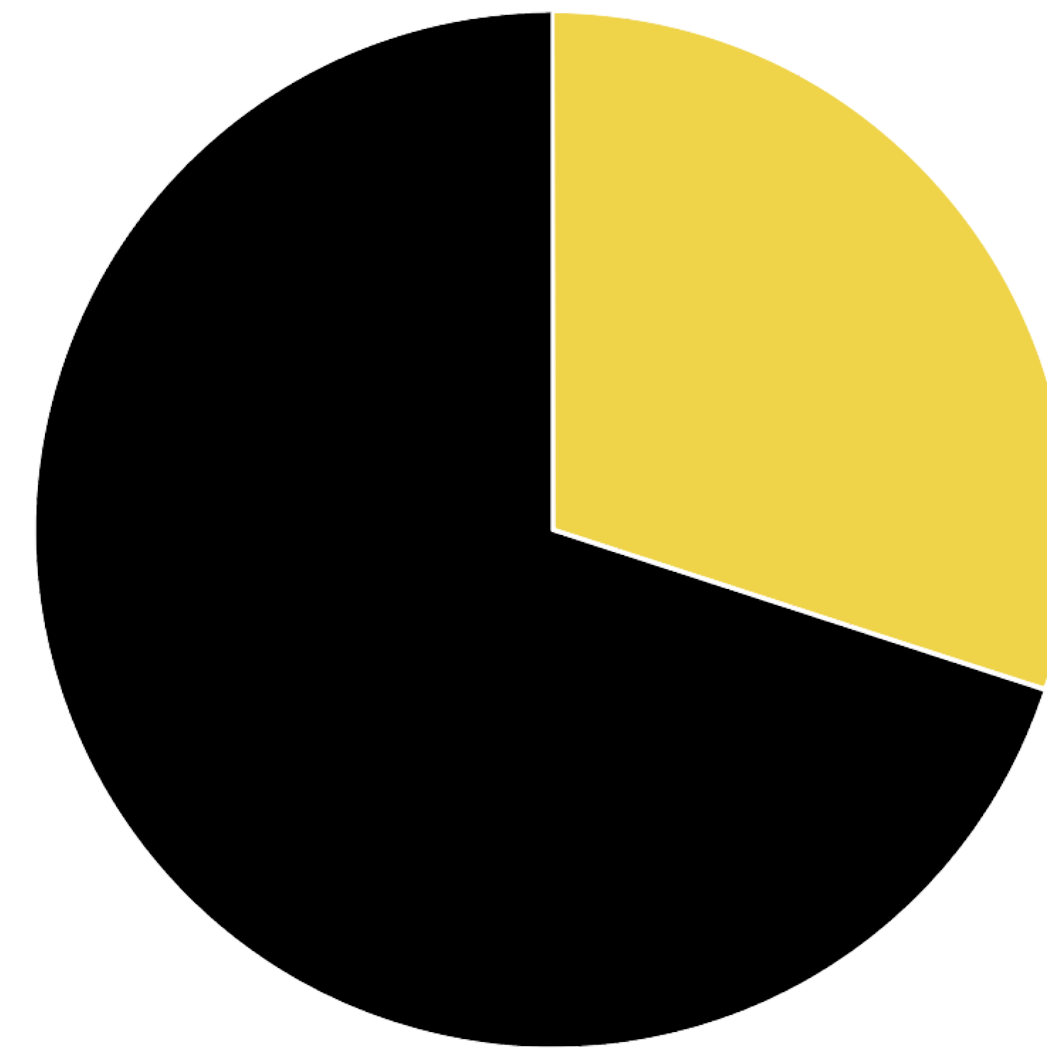
How do we make conversational AI systems that genuinely represent and support the diverse human values of the people who build and use them?



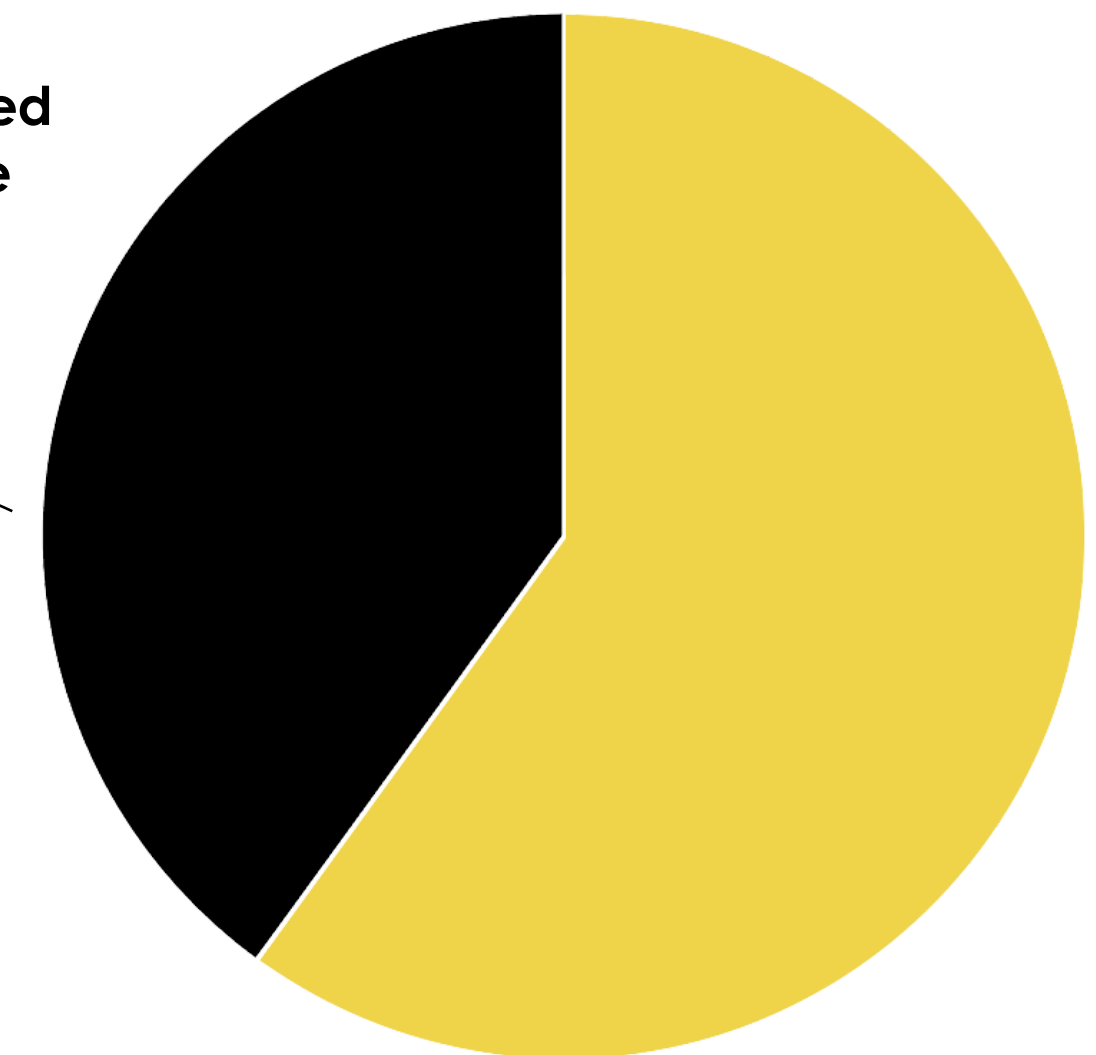
The Gap Between Values and Practice

1

Conversational and agentic AI have moved from pilot to production faster than most organisations can govern them. McKinsey's 2026 AI Trust Maturity Survey of roughly 500 organisations found that **only around 30% of organisations have reached a mature stage in the dimensions that matter most for embedding values: strategy, governance, and control over increasingly autonomous systems.** With average organisational responsible AI maturity scoring only 2.3/4 in 2026, nearly **60% of organisations cited knowledge and training gaps, not tooling, as the leading barrier to putting responsible AI into practice.**

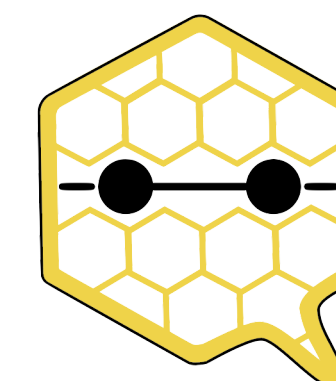


Only ~150/500 organisations achieved maturity in AI strategy, governance and control.



~300 organisations blamed knowledge and training gaps as the biggest barrier to responsible AI practices.

Regulation is racing to catch up rather than leading the way: obligations under the EU AI Act for high-risk systems, covering many conversational and decision-support tools, were due to bind deployers from August 2026 and are only now being contested and potentially delayed under the EU's Digital Omnibus negotiations.



HIVE^{CAI}

Human Interactions, Values, & Experiences with Conversational AI

Our Vision: Closing the Values Gap

2

It has never been clearer that organisations can write down their AI principles far more easily than they can operationalise them, and the people most affected — the practitioners building these systems, and the users on the other end of the conversation — are rarely given a real say in what those values mean in practice.

This is the gap that the HIVE CAI Research Lab works to address. We work on both sides of that gap at once through two research streams.

Living with Conversational AI Values

How people perceive, interpret, experience, and are influenced by conversational AI value expressions.

Example Projects:

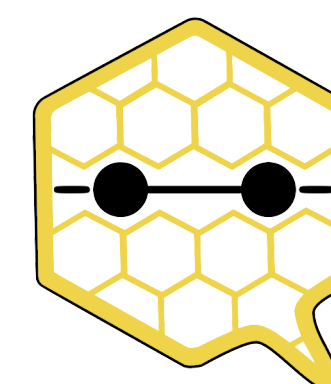
- Conversational AI literacy & user agency — helping people understand, evaluate, and steer conversational AI behaviours.
- Impacts on users — understanding how diverse users perceive, and are affected by the values embedded in conversational AI systems.

Shaping Conversational AI Values

How human values get embedded, configured, and operationalised in conversational AI systems themselves, turning stated principles into concrete, testable design decisions.

Example Projects:

- Practitioners' work with values — understanding how AI practitioners conceptualise and operationalise values in their day-to-day work, and where that process breaks down.
- Value visualisation & customisation — toolkits and interfaces for designing and building more value-sensitive chatbots.
- Co-design for human-centred applications — designing with stakeholders to build and evaluate human-centred conversational AI systems.



HIVE^{CAI}

Human Interactions, Values, & Experiences
with **Conversational AI**

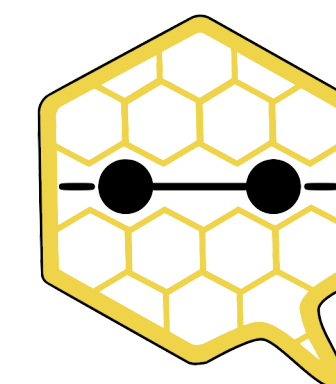
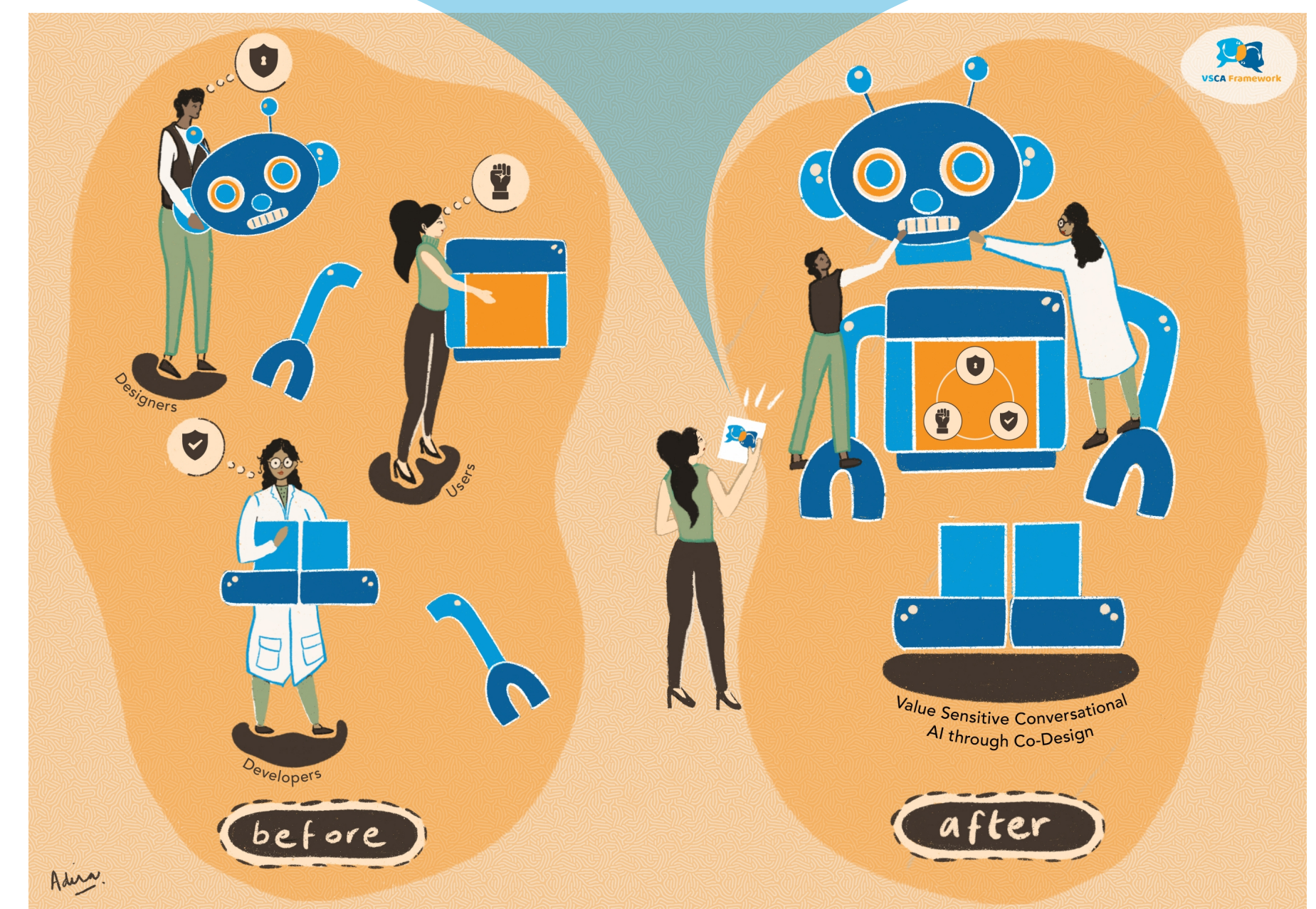
Evidence in Practice: The VSCA Framework & Toolkit

The Value-Sensitive Conversational Agent (VSCA) Framework gives design teams a practical, human-centred process for building conversational AI that reflects the values of everyone with a stake in them, without requiring teams to rebuild their existing workflows.

The toolkit has been field-tested with conversation designers, developers, and product managers, in organisations ranging from one-person start-ups to teams at Amazon. Across deployments, users rated prototypes built with VSCA as **50% better aligned with the creators' intended values** and **60% better aligned with their own personal values**, compared to a control prototype, and both technical and non-technical professionals were able to put its outputs directly into their workflows.

The work is published as "The Value-Sensitive Conversational Agent Co-Design Framework" in the International Journal of Human Computer Interaction.

The lab's wider portfolio also includes "KNIT: Computational Boundary Objects for Real-Time Convergence in Interdisciplinary Teams", produced with Imperial College London's Collective Innovation Lab, which received an **Honourable Mention** from academic peer reviewers.



HIVE^{CAI}

Human Interactions, Values, & Experiences
with Conversational AI

Looking Ahead: Where We're Going Next

As agentic AI adoption accelerates and organisations' own responsible AI maturity struggles to keep pace, we are conducting **cutting-edge research** into **higher-autonomy, higher-stakes agentic settings**: where governance and control lag the furthest behind.

We are also building **literacy and agency tools** that put some of that control back in users' hands, rather than treating trust purely as an organisational compliance problem.

And because knowledge and training gaps remain the top barrier organisations report, we are deepening **embedded fieldwork with practitioners** by studying how decisions regarding values actually get made, or lost, in day-to-day design decisions.

Dr. Malak Sadek directs the lab and is Deputy Director of the MPhil in Human-Inspired AI at CHIA. She holds a PhD in Design Engineering from Imperial College London, and has previously consulted on human-centred AI product design and worked with various multinationals, NGOs, academic institutions, and governmental organisations.

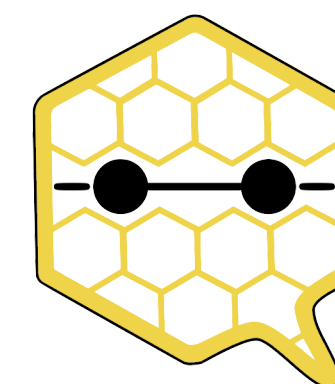
The lab's resident team spans postdoctoral researchers, research associates, and visiting researchers, and draws on a network of external collaborators.

How We're Getting There

This agenda is carried out in collaboration with an expanding network of collaborators spanning industry, civil society, and academia **across the UK, Europe, and the US**.

This breadth lets the lab move fluidly between academic research, industry practice, and public engagement.

Our work is funded by **The Alan Turing Institute**, the **Leverhulme Centre for the Future of Intelligence**, the **Institute for Humane Studies**, and **Cambridge Humanities Research**.



HIVE^{CAI}

Human Interactions, Values, & Experiences
with **Conversational AI**

Get Involved in HIVE CAI's Research and Work

5

We welcome **collaborators, sponsors, and prospective researchers** who want to work at the **intersection of conversational AI and human values**.

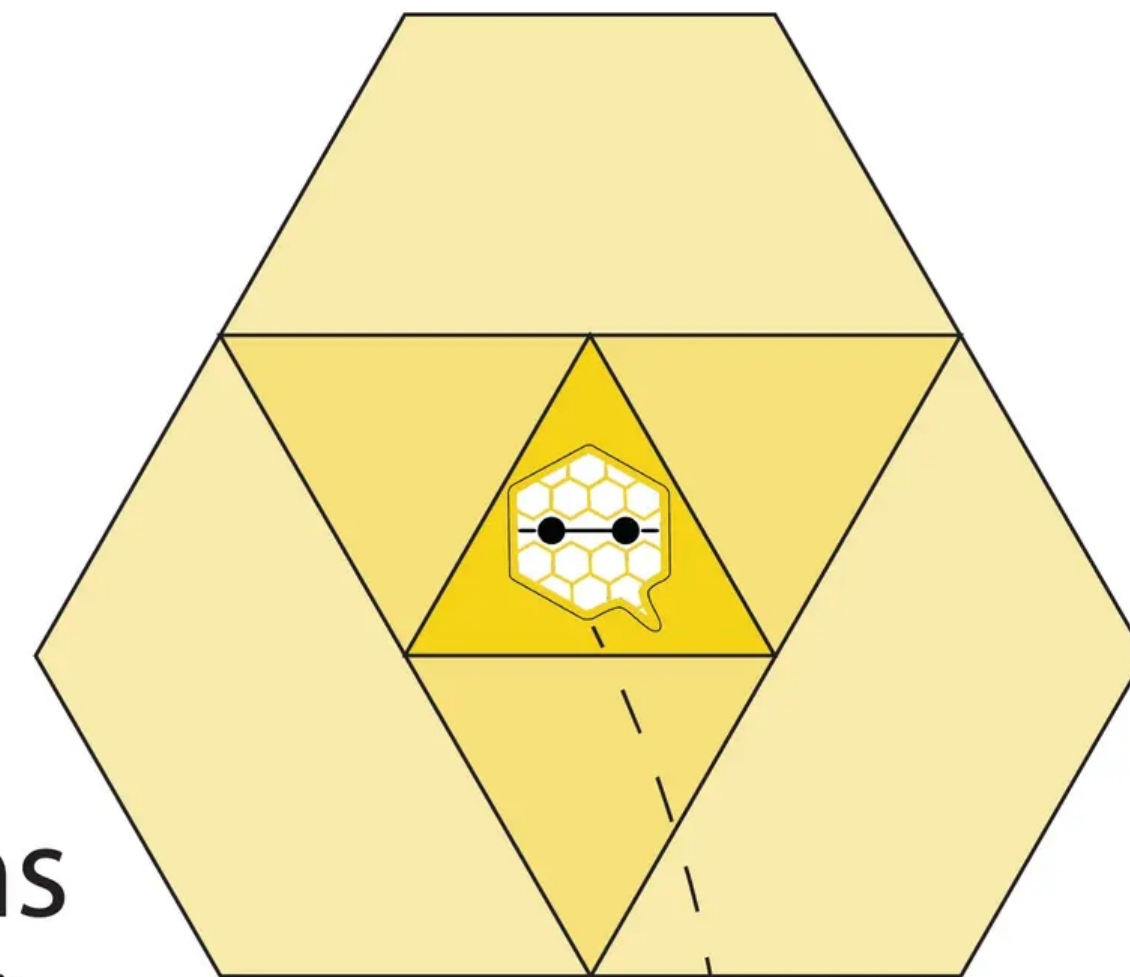
For research collaborations, joint grants, or shared initiatives, and for PhD, MPhil, visiting researcher, or postdoctoral opportunities, reach out to [**hive@chia.cam.ac.uk**](mailto:hive@chia.cam.ac.uk)

Humans

- Empathy & Dignity
- Ethics & Responsibility
- Human-Centred Research methods

Technology

- Cutting Edge Topics
- Novel Technologies & Applications
- Design & Development



Impact

- Top Tier Publication Venues
- Concrete Evidence of Relevance
- Tangible Outcomes

HIVE^{CAI} Lab

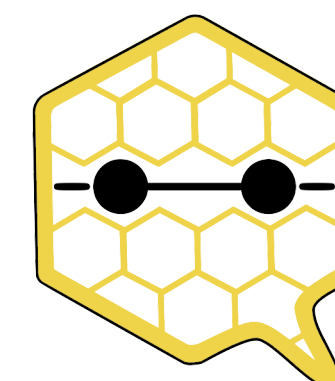
- Positive Research Culture
 - Interdisciplinarity
- Curiosity & Open Mindedness
- Honest and Comfortable Communication

Stay Connected with HIVE CAI

Website: <https://www.hive.chia.cam.ac.uk>

Google Scholar: <https://shorturl.at/8FPcg>

LinkedIn: <https://www.linkedin.com/company/hive-cai>



HIVE^{CAI}

Human Interactions, Values, & Experiences
with **Conversational AI**