

## **Workstream: The socio-technical complexity of leveraging Artificial Intelligence (AI)**

Our AI research program focuses on advancing the field of AI by pursuing five key goals.

First, we aim to expand the philosophical underpinnings of AI by leveraging the philosophy of knowledge and different methods on how the Eastern and Western philosophies of consciousness and knowledge can be used to make further advancements in AI, including super alignment and GAI. In doing so, we take account of the complex environment in which the development and adoption of these technologies are embedded, such as human psychology, cultural diversity, economic incentive structures, and institutional frameworks.

Second, we endeavor to unpack the broad spectrum of tools subsumed under the umbrella term AI to explore their specific strength and potential, weaknesses, and limitations. In doing so, we recognize the value of neural network-based learning, symbolic knowledge representation, and logical reasoning. We are curious to explore the potential of a 3<sup>rd</sup> wave of AI that leverages the potential of both research streams.

Third, we explore how to optimally leverage and integrate AI with other available and emerging technologies (e.g., exascale, edge and quantum computing, Distributed Ledger Technology), sensing technologies (e.g., localized surface plasmonic resonance (LSPR), capacitive or nano biosensors), new materials (e.g., perovskite solar technology, photosynthetic organisms in building materials), or biotech advancements (e.g., gene-editing and molecular-guided therapy, microdosing of drug administration through new micropumps and actuation mechanisms). We are confident that AI can boost the latent potentials of these innovations in a plurality of ways and open a wide range of creative combinations designed to reach specific goals within the unique ecosystem.

Fourth, we recognize that bias is one of the major drawbacks of modern AI methodologies. With AI spreading into many critical areas, such as healthcare and benefits, policing and criminal justice, education, and recruiting, it is crucially important to identify the sources of bias and develop methods to improve fairness in algorithmic decision-making.

Lastly, we aim to study the environmental and social impact of large-scale operation of advanced machine learning models such as Chat GPT and develop frameworks to evaluate the overall cost of running such methodologies. By pursuing these goals, we aim to make significant contributions to AI while ensuring that AI is developed and used responsibly and ethically.