

Moonshot in Artificial Intelligence: Trustworthy, Multicultural Generative AI Systems for Safe Physical Interaction with the Real World

Developed and supported by CLAIRE and euRobotics

7 November 2023

Advances in artificial intelligence (AI) over the last decade have begun to transform the way we live and work. AI is becoming a crucial technology in everything from engineering and medical breakthroughs to the advancement of science. The recent advances we have seen in generative AI models, such as Chat GPT and Bard, are not just triggering public debate but heralding a seachange in the progression of AI. They will be followed by new models and greater refinements of capability that can and will be deployed across all sectors in both digital and physical systems such as robots and virtual assistants.

Regrettably, these advances are being driven outside of Europe, under the control of a small number of large technology companies. There is now a growing concern that if this situation is unchecked, if Europe fails to respond in scale and time, it will lead to a lasting technological and economic dependence and a corresponding loss of global market share and strategic sovereignty. The importance of this can hardly be overstated and, as is widely observed, this is particularly problematic in areas such as healthcare, education and law enforcement, where European values are markedly different from those of the other countries, such as the United States or China, where this technology originates.

As we learned with the pandemic, challenging times need to be met with coordinated action that delivers a European response to strengthen our capability and market. Such action cannot be phrased in the normal terms of European funding mechanisms; instead it must create greater coherence through combined pan-European action. Therefore, in order to retake the initiative and ensure European sovereignty and prosperity we propose a moonshot: a large-scale, ambitious effort of substantial societal importance. Understandable to everyone, easily assessed as complete and usable as a means to focus effort, mobilise talent and pool the resources needed to succeed.

The moonshot goal is that, by the end of the decade, European citizens, industries and public organisations have trustworthy European alternatives to the generative AI systems created outside of Europe.

Success in this moonshot would

- **bring widespread and significant benefits to citizens, industry and society, in the form of alignment with shared values and of the global competitiveness of our economies;**
- **make major contributions to solving the grand challenges of our time, notably climate change, health and inequality;**
- **bring into existence AI systems that satisfy the seven trustworthiness criteria defined by the European Union;**
- **bring critical technology and infrastructure under European democratic control.**

The AI systems resulting from the moonshot will be trustworthy, multicultural, made in Europe – fit to be used safely and effectively in sectors such as healthcare and industrial production, as well as to address major societal challenges such as climate change. In this context, "trustworthy" refers to the seven trustworthiness criteria defined by the European Union, "generative AI systems" refers to

the class of AI systems, such as ChatGPT, Bard etc., that can be used broadly and with limited AI expertise to support a wide range of tasks and applications. **Making these AI systems or agents trustworthy, which they currently are not, is particularly important for applications where they interact with the real world, for example in the form of intelligent robotic systems.**

In the past, Europe has experienced major challenges to its technological sovereignty but has been able to successfully come together to meet them. After the Second World War, most particle physics research happened in the US, and Europe established CERN; at the end of the Apollo program, most space research and development happened in the USA and in the Soviet Union, so Europe established the European Space Agency (ESA), and when American companies dominated the aerospace industry, European countries jointly established Airbus. The examples of CERN, ESA and Airbus demonstrate that such large-scale efforts can succeed and achieve global impact.

The difference is that this time, there is more at stake, as AI is a fundamental technology that cuts across all areas of the economy and touches all areas of society. AI affects every citizen in Europe – it is, or will be used in every sector and by every citizen on a daily basis, if not now then in the near future. The provision of trustable AI is like the provision of electricity: We can live without it, but our lives will be measurably poorer as a result.

If Europe is to remain the egalitarian stronghold it seeks to be, with liberal democracy as its norm, then the challenge created by badly made or malevolent AI is too important to be addressed by the commercial interests of companies alone. **To orchestrate and execute the moonshot, we therefore propose to build on Europe's experience by creating a CERN for AI; an organisation and hub at the centre of a broad network of AI competence centres throughout the EU and associated countries.** The investment scale and timing must be sufficient to create a globally visible focal point that provides crucially needed large-scale compute and data infrastructure, experimentation and testing facilities. It must have the necessary on-site expertise to maintain, expand, and use these facilities efficiently. It can host large and carefully curated collections of data – including sensitive data, as well as create a critical magnet for global talent, with individuals located there for a fixed time, before returning to their home ecosystems to share their expertise and connections, thereby creating new value flows as the moonshot progresses. This is why the CERN model, including connections between the central facility and other centres of excellence, is the right approach.

We estimate the public funding required for this moonshot at roughly 100 billion Euros, to be invested between 2024 and 2029 by EU member states and associated countries, including the UK, and we advocate including interested non-European partners, such as Canada, Japan and possibly the United States. We note that 100 billion Euros over 6 years, or about 16.7 billion Euros per year, roughly corresponds to 10% of the EU budget, 12 times the budget of CERN, and 73% of the budget of NASA on an annual basis. For comparison, the inflation-adjusted cost of the Apollo programme was 240 billion Euros.

This is a wakeup call: Europe needs to engage in a race that we cannot afford to lose. The impact of AI when made trustworthy is significant, the strategic cost of missing a step too great to ignore. As Europeans we need to act now and act together to address the shortfall in capability and use the technology at our disposal to improve productivity and address our societal concerns. There is a narrow window of opportunity for us to create a globally significant force for AI in Europe and any further delay will be costly.

About CLAIRE

The Confederation of Laboratories for Artificial Intelligence Research in Europe (CLAIRE, claire-ai.org) is an international non-profit association created by the European AI community that seeks to strengthen European excellence in AI research and innovation, with a strong focus on human-centred AI. CLAIRE aims to establish a world-wide brand recognition for "trustworthy AI made in Europe" (at the level of CERN), and to position Europe in control of its own future. Launched in June 2018, with offices in The Hague, Saarbrücken, Prague, Rome, Brussels, Oslo, Paris and Zürich, and with over 460 members, covering jointly more than 24,000 employees in 37 countries, CLAIRE has the support of the vast majority of Europe's AI community, spanning academia and industry, research and innovation, as well as of nine European countries, and over 30 scientific associations across all of Europe. CLAIRE is a founding member of the AI, Data and Robotics Association (ADRA), and the CLAIRE vision for European excellence in AI is supported by the European Association for Artificial Intelligence (EurAI), by the Association for the Advancement of Artificial Intelligence (AAAI) and by the European Space Agency (ESA).

About euRobotics

euRobotics (eu-robotics.net) is an international non-profit association for all stakeholders in European robotics. Established in September 2012 and serving its founding purpose: to strengthen Europe's competitiveness and to ensure industrial leadership of manufacturers, providers and end-users of robotics technology-based systems and services, it forms the largest network of roboticists and business in Europe with about 250 institutional members, covering small and large companies, associations and institutions, universities, laboratories and RTOs. The objectives of euRobotics are to boost European robotics research, development and innovation, to foster a positive perception of robotics, to support the widest and most effective uptake of robotics technologies and services for professional and private use, and to ensure the excellence of the robotics science base in Europe is maintained. euRobotics is a founding member of the AI, Data and Robotics Association (ADRA) and was the private partner in the SPARC partnership with the EC during Horizon 2020.