Reuben College
UNIVERSITY OF OXFORD

Generation AI

A Global Challenges Programme Initiative

Reuben College's Global Challenges Programme aims to act as an incubator of interdisciplinary ideas and practical action on major global challenges. Structured as a series of concentrated 'challenges', this proposal outlines thematic initiatives intended to tackle concerns about the impact of advanced AI systems on human development.

Background

Advanced AI systems have shown extraordinary capabilities, enabling new ways of learning and working in multiple domains.

Automated analysis: Since 2022, 149 foundation models have been publicly released, many surpassing human performance on a handful of benchmarks, such as image classification, basic reading comprehension, visual reasoning, and natural language.

Content production: AI systems have a prodigious ability to produce believable and usable digital content. One recent paper reports that as much as 57% of the content on the internet has been machine generated or machine translated.

Broad spectrum input and output: AI systems can accommodate an expanding range of input types from various sensors. Examples include text, voice, images, motion, audio, genetic sequences, code bases, brain scans, and eye tracking.

Complex task execution: Another recent phenomenon is the testing of foundation models on so called agent-based benchmarks. These measure performance on tasks that require on-the-fly parsing of complex tasks into simpler steps in many different kinds of environments.

Implications for children, parents, and adolescents

In future years, it is possible that the majority of digital information will be generated, edited, or translated by an AI. Digital communication between persons will be mediated by AIs, as co-pilots, virtual assistants, and personal agents. As AIs become more powerful an increasing fraction of the functionality of digital environments will be dominated by machines, including content creation and commerce. Our virtual and physical environments may become "AI rich" yet this may not necessarily enrich our lives.

For **children**, digital technologies are already embedded in daily life. According to Ofcom, one-quarter of 5-7 year old children own a smartphone and one-third report using social media independently.

Adolescence marks an exploration of identity, social concerns, refinement of moral reasoning, and critical thinking. As of 2020 nearly all adolescents (95%) in the United States have at least one mobile device of their own, and 89% own a smartphone.

Parents and caregivers face a potentially vexing dilemma in evaluating current use of digital technologies - let alone the potential power of multi-modal embedded AI agent systems. Limitations to the use of technology are perceived as socially limiting to children and adolescents, as well as ultimately self-defeating.

Challenges

On May 10 2024, the Global Challenges Programme brought together world-class researchers, innovators, and policy makers at a workshop *Generation AI: Opportunities and Challenges for AI in human development*. At this gathering, we identified a series of specific challenges preventing progress.

Save the date for 14 September 2024, where we will officially kick off the programme and address the challenges noted below.

Research Challenge

Conventional academic practices for the conduct, review, and publishing of research are unsuited to urgent calls for action on complex social issues. New methods are sought to generate representative, accessible, accurate, and timely information available to a wide range of stakeholders.

This challenge will fund work to accomplish three objectives:

1. Increase the quantity and quality of 'mechanistic' studies on technology use and child/adolescent development with the aim of individualised risk prediction.
2. Use Open Research principles within the field which overcome data access, incentives, and limiting scholarly practices.
3. Aggregate and summarise best in class research in a form that can be used by stakeholders outside the academy, including journalists and policymakers.

Ethical Design Challenge

AI systems should be designed in line with ethical and child rights considerations. Yet, translating these principles into action requires a unique set of skills not found in either academic or commercial settings.

Key developmental milestones in childhood and adolescence are relevant to product design as a feature: including language learning and reading, cognitive and affective control, attention, identity, self-esteem, and agency. In order to do this, AI systems must go beyond optimising for *instrumental* measures of accuracy, task completion, efficiency, and user engagement. They should consider *relational metrics* for example promoting parent/caregiver-child bond and peer-to-peer engagement.

In consultation with the interests of the funders of this challenge, we will apply a rigorous design and innovation methodology for particular human developmental pathways, skills, or topics.

- For example, we can imagine a challenge to create a research-aligned tool for parental and caregiver advice pre-tested with input from a range of users.
- Another example could be how we might consider a demonstration tool that enables children and young people to input on what AI literacy means for them.

Resources Challenge

Goals, objectives, and incentives of AI systems reflect the values of those who fund and create the systems themselves. A *public interest technology ecosystem* may be needed to organise skills and resources necessary to field viable, equitable, inclusive and scalable technologies that embed child and adolescent development in the design process.

This challenge will produce a series of white-papers on the creation of a public interest technology ecosystem and case studies on companies and organisations that exemplify this approach.

Collective Response Challenge

Building positive habits when using new AI systems goes beyond individual responsibility and requires a collective response. New modes of engagement as members of a community, in virtual and physical worlds, have the potential to be a powerful means of setting new social norms around use of technology.

This challenge will conduct a small-grants RFP for community and faith-based organisations. Each applicant organisation will aim to set new collective norms around technology use - in line with best research practice.

Global Policy Challenge

Many countries are advancing legislative and policy agendas aimed at protecting children and adolescents from online harms, including those posed by AI. As advanced technologies are deployed globally, there is an urgent need to ensure equitable protection irrespective of the location or the device ownership of the user.

This challenge will support the work of the Digital Futures Lab and 5Rights Foundation to lead on a global advocacy agenda in three areas: design of service, child online protection, and children & young people's rights.

Structure, key experts, and organisations

To address these challenges in an integrated manner, this consortium will be led by Andrew Serazin and Katie Glover from the Global Challenges Programme at Reuben College, University of Oxford. Individual challenges will be led by experts from the following members of the Generation AI working group:

- Lord Tarassenko, University of Oxford, Reuben College
- Jun Zhao, University of Oxford, Computer Science
- Andrew Przybylski, University of Oxford, Oxford Internet Institute
- Gaia Scerif, University of Oxford, Experimental Psychology
- Elizabeth Wonnacott, University of Oxford, Department of Education
- Amy Orben, MRC Cognition and Brain Sciences Unit, University of Cambridge
- Sarah-Jayne Blakemore, University of Cambridge
- Baroness Kidron, 5Rights Foundation
- Zelda Yanovich, Fam Studios