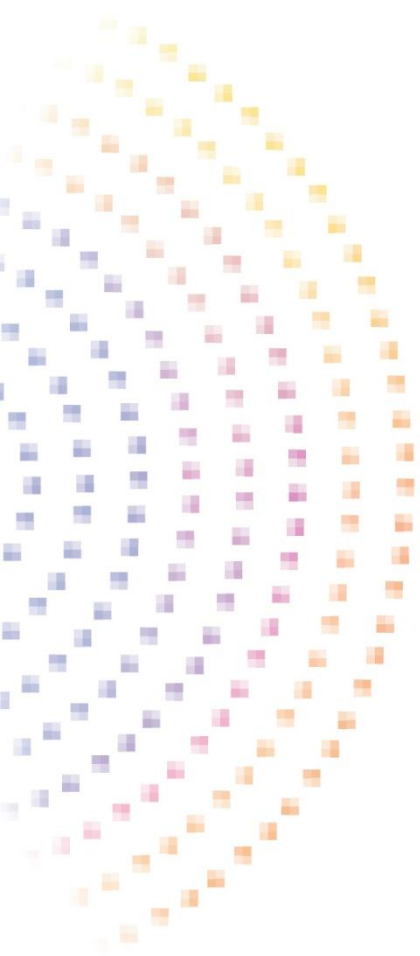# The Christchurch Call Initiative on Algorithmic Outcomes:
Ethics and Governance

The Christchurch Call Initiative on Algorithmic Outcomes: Ethics and Governance

Background

In September 2022 the New Zealand and United States Governments, Microsoft, and X launched the Christchurch Call Initiative on Algorithmic Outcomes (CCIAO).

Working with [OpenMined](), the CCIAO contributors proposed to develop and test a new system that would allow a trusted third party to answer critical questions about algorithmic outcomes, and trust and safety systems, using data from technology platforms, while using privacy-enhancing-technologies to ensure the underlying data and the privacy of users are protected.

This tool could enable accredited researchers in the Call Community to answer important questions about how recommendation systems impact users and what online factors play a role in radicalisation to violence. It could also help us assess and improve online positive interventions that help to reduce the risks of harm and violence. This could include study in specific areas such as gender-based online harm and its links to terrorist and violent extremist content and groups.

As it scales, the CCIAO tool could enable wider outcomes, including helping to unlock insights from new sources. It could help the Call Community study the risks and possibilities of foundation artificial intelligence (AI) models relating to terrorist and violent extremist content, facilitate regulatory compliance, and allow the multistakeholder Call Community to better understand the online environment in a holistic way.

Achievements in Phase One

In Phase One of the project, the CCIAO tool has been built, and put through multiple, rigorous proof-of-function tests. These tests have showed that the CCIAO tool can help mitigate the privacy, security, legal, intellectual property, and reputational challenges associated with enabling third party research on sensitive data. It is the first such system that has been tested and verified for use by multiple platforms and, importantly, is the only one we are aware of purposed for the public good.

The CCIAO tool has gone through an end-to-end multi-stage proof of function test at each of the three online service provider participants, using both synthetic and real data. It successfully completed this first phase of testing in August 2023.

This testing has proved the tool can operate in multiple environments and provide a consistent and secure process for answering defined questions from external parties. A small group of external researchers has been testing the system's functionality and has managed to submit several successful requests. This important work will be presented at the workshop on Saturday, 11 November and positions us for the next phase of the project.

Next Phases and long-term vision

The tool is now ready to scale up – involving more organisations and bringing in more researchers. This requires development of detailed governance structures and an ethics framework for the tool.

This is important to achieve before scaling to ensure we are able to exclude potential malicious uses and other risks.

In Phase Two, the CCIAO proposes to:

- Establish a governance framework that supports multistakeholder outreach and fulfilment of the Call Commitments and principles;

- Build an agreed ethics framework that supports researcher access, enables better understanding of human/AI interactions, and research for public good;

- Investigate funded research partnerships with specialist experts, data custodians, and interested communities of practice;

- Investigate methods for assessing AI risk relating to terrorist and violent extremist content;

- Investigate a regulatory use case, as part of the European Union Digital Services Act;

- Continue to grow the number of contributors and the supporting infrastructure, with the aim of creating a global network of AI transparency tools for the public good.

The proposed long-term vision is for a broad network of CCIAO and similar technologies to be available across multiple institutions/entities, with a responsible governance framework that allows accredited researchers globally to analyse data in a privacy-protected way, across platforms. The network will be governed by an ethics and accredited user framework to enable safe, effective, access consistent with the shared objectives of the contributors and the Call Community.

Elements of a Governance Structure and Ethics Framework

A governance structure and ethics framework is important to ensure a means of governance and oversight of the initiative and future strategy; to enable CCIAO to grow within an agreed set of rules, ensuring core principles and ethics guidelines are upheld and the plan for Phase Two is effectively executed.

The Call will work with the Community to ensure governance and strategic oversight are appropriately managed and well-structured. This will include:

- Core principles to be applied to the CCIAO project, ensuring it is used for agreed public good purposes, that it protects the privacy of users and proprietary information held by platforms, and that it has an impact and helps to fulfil the Call commitments including consistency with human rights, transparency, and multistakeholder engagement;

- A governance board composed of the contributing organisations;

- Mechanisms to engage the Call Community and external experts on the fulfilment of the principles;

- Systems for bringing onboard new contributors, evaluating and adapting the network as it scales, and working with Call Community members to scope involvement;

- Exploring funding options for the next phases, ensuring well-documented, structured management of grants and contributions.

- Managing researcher accreditation and participation consistent with governance principles and the ethics framework, ensuring a diversity of expertise and experience.

- Identification and mitigation of risks to the CCIAO including e.g. from malicious misuse, unauthorised access, and other potential problems.

- Maintaining an adaptable, streamlined, scalable and effective project that meets the needs of the Call community and CCIAO contributors.

Next Steps

The CCIAO participants will hold a workshop for Call Community members in the margins of the 2023 Leaders' Summit. This will provide an important opportunity for describe progress to date and show how the CCIAO works in practice. It will also provide an initial opportunity to describe the process from here on the governance structure and ethics framework, and the future steps for the project.

A more detailed paper will be circulated to the Call Community, drawing on initial feedback from the workshop, in the coming weeks.