accelerator

# We Need Faster, More Efficient Research

The online information environment is pervasive in human life, shaping outcomes across fields ranging from mental health, to disaster response, to political violence. Unfortunately, research is not keeping pace with the scale of challenges or speed of changes, leaving policymakers without reliable evidence to guide critical decisions. Researchers are spending significant time on duplicative engineering work instead of focusing on social science because we have not built the kinds of shared infrastructure needed to efficiently study the information environment. Beyond perennial challenges with data access, existing tools are too technically complex for most scholars to easily use, and no one is maintaining systems that could support the field as a whole but are too expensive and time-consuming for any one research lab.

The Accelerator will speed scientific progress by providing shared technical resources and supporting access for a broad group of researchers to ultimately enable evidence-based policymaking.

## Surveying the Field

Incubated by the Carnegie Endowment for International Peace and Princeton University, our interdisciplinary team assessed the potential for shared infrastructure to radically improve the pace and depth of scientific progress on the information environment by:

- Commissioning studies on specific research challenges with partners from 13 universities and four NGOs in the US and Europe;
- Analyzing 3,923 academic papers in a stratified random sample of work across dozens of journals in five academic fields;
- Interviewing more than 240 researchers worldwide to understand current practices and discuss what kinds of infrastructure would speed their process; and
- Examining the structural and financial models of 141 organizations, including academic groups, major scientific instruments, non-profits, and social listening companies.

We identified three core activities that would dramatically speed research: (1) developing tools to

speed common research processes; (2) collecting data on typical activity, which requires substantial ongoing maintenance; and (3) fostering a global, cross-disciplinary, and multisector community by supporting fellows and serving as a training hub.

## Shared Infrastructure to Advance Research

There is a clear need for a large-scale, multi-national institution to support research on the information environment, as described in this white paper. The Accelerator is developing initial infrastructure in partnership with community members and will contribute to this overall vision by providing global resources to address issues faced by researchers examining the information environment.

In February 2023, we identified six initial tools that could support a wide range of research styles, developed mitigation measures for each tool's technical risks, and conducted expert-led reviews of privacy, ethical, and legal considerations for each tool. As of Fall 2023, we have started prototyping two of these tools and are working to establish a multi-university consortium to develop and maintain them. The first two tools are:

**Typical activity datasets:** baseline datasets for multiple platforms, designed to enable researchers to characterize typical activity, build various baseline measures on it (e.g., of toxicity or trends in usage patterns or topics), and match activity with content in broadcast media. These baselines will eventually be enriched with a number of features to allow users to choose or generate the most relevant baseline for their work.

**Image Processing Pipeline:** a computational pipeline that will ingest visual media, convert it to easily used dense embeddings, merge these embeddings with related metadata (e.g., associated text or author information), and provide functionality for 1) searching for images shared in online spaces and the authors who post them, and 2) comparing user-provided data to Accelerator-collected baselines.

## From Data to Knowledge

To achieve the Accelerator's goals, it is establishing partnerships with initiatives focused throughout the research process. These collaborations will not only bring innovative social science or technology development to the project but will ultimately also make more effective use of data access, thereby strengthening the evidence base and advancing community and policy.

## Creating Change Together

Effectively managing online commons is one of the transcendent challenges of our era. The field would produce broader and deeper research if understanding the information environment wasn't so expensive and difficult. With such research in hand, decision-makers would be more likely to implement policies based on sound scientific evidence, as we expect them to do in other policy areas ranging from education to commerce. Together, we can overcome the legal, organizational, and technical challenges that have so far impeded research and evidence-based policymaking on the information environment.

---

*The Accelerator has been generously funded by Princeton University, Microsoft, and the John S. and James L. Knight Foundation. Craig Newmark Philanthropies and the William and Flora Hewlett Foundation also supported the exploratory studies.*