CARNEGIE
ENDOWMENT FOR
INTERNATIONAL PEACE

ESOC
Empirical Studies of Conflict

NOVEMBER 2022

A CERN Model for Studying the Information Environment

Alicia Wanless and Jacob N. Shapiro

A CERN Model for Studying the Information Environment

Alicia Wanless and Jacob N. Shapiro

Contents

Introduction

After the Second World War, European science was suffering. Scientists were leaving Europe in pursuit of safety and work opportunities, among other reasons. To stem the exodus and unite the community around a vision of science for peace, in 1949, a transatlantic group of scholars proposed the creation of a world-class physics research facility in Europe.[1] The grand vision was for this center to unlock the mysteries of the universe. Their white paper laid the foundation for the European Center for Nuclear Research (CERN), which today supports fundamental research in physics across an international community of more than 10,000 scientists from twenty-three member states and more than seventy other nations.[2] Together, researchers at CERN built cutting-edge instruments to observe dozens of subatomic particles for the first time. And along the way they invented the World Wide Web, which was originally conceived as a tool to empower CERN's distributed teams.

Such large-scale collaboration is once again needed to connect scholars, policymakers, and practitioners internationally and to accelerate research, this time to unlock the mysteries of the information environment. Democracies around the world are grappling with how to safeguard democratic values against online abuse, the proliferation of illiberal and xenophobic narratives, malign interference, and a host of other challenges related to a rapidly evolving information environment. What are the conditions within the information environment that can foster democratic societies and encourage active citizen participation? Sadly, the evidence needed to guide policymaking and social action in this domain is sorely lacking.

Researchers, governments, and civil society must come together to help. This paper explores how CERN can serve as a model for developing the Institute for Research on the Information Environment (IRIE).[3] By connecting disciplines and providing shared engineering resources and capacity-building across the world's democracies, IRIE will scale up applied research to enable evidence-based policymaking and implementation.

Where CERN "exists to understand the mystery of nature for the benefit of humankind," IRIE will exist to understand the mystery of the information environment for the benefit of democracies and their citizens.[4] While the laws of physics change slowly, the conditions within the information environment are little understood and changing rapidly through the addition of new technology. Additionally, studying the information environment will require analysis of personal and at times sensitive data of internet users, increasing the need for international collaboration. As CERN did by inventing new collaboration tools, IRIE will leverage those same technologies to unlock the collective genius of researchers and practitioners from a variety of fields to strengthen democracy, alongside interested citizens who contribute their own data and expertise. While there are obvious differences between the field of physics and those emerging to study the information environment, aspects of CERN's development can guide the creation of IRIE, an institution that can uniquely address the growing needs of researchers.

Why the CERN model?

CERN is the world's biggest particle physics laboratory.[5] In its own words, CERN probes "the fundamental structure of particles that make up everything around us . . . using the world's largest and most complex scientific instruments."[6] In so doing, CERN enables fundamental research, advances technology, brings nations together, and builds capacity for the field. CERN is run by a consortium of twenty-three member states and works with researchers from countries around the world. CERN has also been a catalyst for innovation, encouraging collaboration between academic researchers and the private sector to develop infrastructure.

Speed and scale are the impetus for turning to CERN as a model. Current approaches to studying the information environment and its effects on democracy are not keeping pace with rapidly evolving threats. Most funding for research on the information environment is issued on a project basis to researchers from academia and civil society.[7] These grants fund data collection and tool creation for specific projects, but they rarely support professional staff or core scientific infrastructure.[8] As a result, each group builds its own data collection and processing pipelines, many of which are quite similar, mostly using graduate students and postdoctoral researchers inexperienced at basic engineering tasks. As a result, the system

as a whole is wildly inefficient. The right scientific processes have simply not been put together to understand the environment where humans form their understanding of the world.

It does not have to be this way. While the engineering resources required to build shared infrastructure for studying the information environment might be beyond the reach of individual labs, they are much smaller than what governments routinely spend on other scientific instruments.

What this state of affairs means in practice, coming back to our analogy, is that research on the information environment is in a similar position to that of physics in 1932—with one lab splitting the first atom using an early particle accelerator, another discovering the neutron, and yet another across an ocean finding the positron—when what is really needed is the equivalent of a Large Hadron Collider to begin unlocking the mysteries of the universe.[9] While the technology didn't yet exist to build the instruments that would begin unlocking particle physics in the 1930s, it does today for researching and analyzing the information environment.

In other words, it's a matter of political will.

What Is the Information Environment?

The information environment is the space where human cognition, technology, and content converge. It is the space where information is processed by humans and, increasingly, machines in accordance with the dominant social norms to make sense of the world. This space also includes the technology used to help process and access this information, such as printing presses, cell phones, the web, and augmented reality, as well as related content such as books, videos, status updates, and online articles. These outputs move through the information environment via multiple channels, such as television, radio, gaming platforms, and social media, but also in person. The information environment is an adaptive system growing in complexity with the emergence of new social norms and technologies. And we barely understand how it all works.

The importance of the information environment is clear. Pollution within it arguably stymied public health measures during a global pandemic and could create challenges for adolescent mental health.[10] But research about issues such as the relationship between screen time and youth mental health often finds more correlation than causation.[11] Little is known about how to measure the effects of threats such as disinformation, nor is there reliable evidence on the impact of the most important interventions to counter them, including those for preventing violent extremism.[12] Beyond specific problems and threats, there is a lack of a baseline for what makes a healthy information environment that organically fosters democracy—much less a basic articulation of what constitutes as healthy.

In the absence of evidence on what interventions might help ensure the integrity of the information environment, policymakers are at risk of turning to old methods of control that are a poor fit for this new space. Right now, interventions to stop unwanted behavior are favored over measures that might better foster vibrant, fact-based political conversation. Indeed, in the absence of clearly articulated democratic principles for governing the modern information environment, the only model for countries to follow is the censorious one pushed by authoritarian states.[13] Liberal democratic societies must find a path forward before the temptation for blunt solutions precipitates a slide into autocracy. In this, CERN offers a second nugget of inspiration.

Built for Peace; Build for Democracy

CERN emerged from a postwar impetus for peace-building initiatives and was framed around a pursuit of science for peace. Today the mission for IRIE would be to save democracy. A 2021 study by the Swedish International Institute for Democracy and Electoral Assistance found three times the number of countries sliding into authoritarianism than becoming more democratic over the last five years. At the same time, the gap between the democracy people want and the democracy they feel they have is big and growing—even in democracies.[14]

Although the causes of the global democratic recession are multiple and complex, it is clear that significant challenges stem from the manipulation and pollution of the spaces where public deliberation occurs and where people form their shared realities. Because democracy is inherently political, it is critical that IRIE support nonpartisan, independent science in defining the field and building governance structures for the information environment.

CERN's international governance model provides inspiration here too. Taking a multinational and interdisciplinary approach protects against the organization's abuse because it ensures no single member can exploit the center if they backslide into authoritarianism. A multinational approach also enables greater diversity as well. Just like the CERN works with researchers from Brazil, India, and South Africa, IRIE too should support and engage a broad array of international researchers, policymakers, and practitioners committed to democratic principles in their study of the information environment. By bringing together a range of nations, IRIE can create data-sharing protocols that are in line with a range of research ethics frameworks and data privacy laws.

In its multistakeholder approach, IRIE has yet another opportunity to support democracy through active engagement with citizens in its work. Democracy is as much about values as it is citizens. Indeed, informed participation by citizens of their own free will is the basis of

democracy's legitimacy. With declining rates of trust not only in key democratic institutions but also in many of the technology companies that dominate the modern information environment, it is not enough for IRIE to simply study the relationship between information and democracy.[15] To go deeper, IRIE will need to not just engage media, communication, and democracy scholars and practitioners but also cognitive psychologists and behavioral scientists to studying the entire system that is the information environment, including the people within it, their relationship to information, and how a variety of information types from various sources impact their understanding.[16]

IRIE must also engage citizens in this research and related policymaking, exploring how shared infrastructure could lead to co-creation between citizens and scientists. At the most basic level, this could include engaging citizens in sharing their data for research purposes, ensuring they understand its use, the results, and also their options for safeguarding data throughout the research lifecycle and beyond. Here lessons can be learned from initiatives such as Mozilla Rally.[17] Such collaboration might also entail finding approaches to enable more meaningful democratic participation with citizens. Such an approach has the added benefit of addressing science elitism while also rebuilding trust.

Fostering a Field

While CERN is an inspiration, there are important differences between particle physics and the field emerging to research the information environment. The former had already identified a range of fundamental principles by the time CERN was being conceptualized. The latter still struggles with basic observation and indeed shared definitions and assumptions. This apparent obstacle highlights the opportunity for IRIE's field-building component to speed research.

A Path to Consilience

Knowledge of the information environment is undeveloped today, especially compared to physics in the 1940s. Researchers from a variety of fields including sociology, political science, psychology, computer science, law, communications, and more are applying distinct methods to aspects of the information environment. The differences between how these fields describe similar problems and frame research questions underscore the need for a center of gravity to foster consilience and systemic thinking.

Research on the information environment, by nature of the current funding model and lack of consilience, is a patchwork. Aspects of the information environment are narrowed down into manageable bits—such as a case study of a single influence operation, measuring

the efficacy of a specific intervention such as fact checking, or the polarizing effects of using a specific technology or mode of communication. But to unlock the mysteries of the information environment, all these things must be understood in the context of the wider systems in which they occur. For example, when online disinformation is studied in isolation from the larger system in which people's social media choices are less polarized than the television they watch, it can seem more alarming than it really is and perhaps garner more policy attention than it should compared to other threats such as online harassment and abuse.[18] Likewise, in looking at only one channel of distribution, such as a social media platform, researchers can miss other crucial points of amplification and intervention, such as podcasts or pundits operating on traditional media. Moreover, an emphasis on studying threats or unsavory phenomena detracts from research that might unlock the conditions in the information environment that might foster democratic processes and citizen engagement.

IRIE can provide a center of gravity for developing the field by working to generate consilience between various disciplines, creating resources that help put information ecosystems into context, and developing shared frameworks. For example, a fellowship program that brings researchers, policymakers, and practitioners from different fields together to build evidence-based answers to specific policy-relevant questions could catalyze shared approaches across disciplines that now study aspects of the information environment in isolation from one another.

Taking a systemic approach to understanding how the modern information environment is shaping society, and how in turn social norms shape the information environment, is key. Defining the field in this manner moves the discourse beyond research about specific influence campaigns, which are inherently partisan, or specific platforms, which have concentrated financial interests. Studying how the information environment functions as a whole can help IRIE rise above debates about who are arbiters of truth in any particular instance.

Enabling Scientists to Focus on Science, Not Engineering and Infrastructure

CERN built shared resources for scholars from around the European continent, enabling them to spend less time fabricating and designing and more time on science.

Many kinds of research on the information environment are data intensive. Data engineering and infrastructure are expensive, and resources within research communities are scarce. Those studying the information environment struggle to hire technically savvy people, often competing for talent with tech companies.[19] Data pipelines designed by graduate students and postdoctoral scholars are seldom designed with repurposing in mind, and thus researchers often re-create others' work. By developing shared engineering infrastructure and providing access to engineers, IRIE can not only enable both quantitative and qualitative research but also enable real-time or near-real-time research. These shared resources could include representative samples of activity, products that transform and

aggregate data to ease analysis, easy-to-use interfaces that make data-driven tasks (such as visualizing co-occurrence networks of unusual phrases found in screenshots) accessible for qualitative researchers, and research practices developed by ethicists who can help substantive scholars better consider fundamental moral dilemmas.

The importance of doing all of this is highlighted by the growing mismatch between the skills researchers have, which are focused on working with text, and those they need with the emerging prominence of video in the information environment. With billions of worldwide monthly active users, Facebook, YouTube, Instagram, and TikTok all include or exclusively host video content, to say nothing of the videos posted directly to news media websites.[20] Compared to text, analyzing video requires significantly more technical resources for storage and computation to complement manual analysis. With current limitations, many researchers export transcriptions of what is said in videos and analyze just the text, a practice that misses out on the inherently visual nature of the medium and only works in English and around a dozen other languages.[21]

If IRIE starts now, this engineering and infrastructure support can be in place to take advantage of access to data from very large online platforms that the European Union's Digital Services Act (DSA) and Code of Practice on Disinformation, which will begin to open up in 2024.[22] While game-changing, data access alone is insufficient to speed up knowledge development. Researchers need increased engineering resources and infrastructure to make sense of these data in context. Through shared engineering resources, IRIE can enable researchers to take advantage of the opportunities presented by the DSA.

Capacity Building

CERN was created with education in mind and continues to play an important role in training scientists and engineers, as well as students and educators at all levels.[23] A similar commitment to training on the information environment and democracy should go one step further to support policymaking. This could take the form of multiple types of training, from traditional academic fellowships supporting researcher use of tools developed by IRIE to more complex programs that bring different types of stakeholders—from academia, civil society, government, and industry—together to collaborate in answering specific policy problems. The key is to create a feedback loop between research, policy, and social action such that outputs from IRIE help inform what conditions within the information environment are beneficial for democracy and identify the interventions to best support that while mitigating threats to it.

A Community-Based Phased Approach

From the first public call for the creation of CERN in late 1949 by French physicist Louis de Broglie at the European Cultural Conference to breaking ground in Geneva in the spring of 1954 to launching its first accelerator the Synchrocyclotron in 1957, CERN took time to become a large-scale physical reality.[24] The development of IRIE can move much faster but should follow a similar community-based approach as laid out by CERN.

The banner for the CERN was taken up at the 1950 UNESCO General Conference, which tabled a resolution to "assist and encourage the formation of regional research laboratories in order to increase international scientific collaboration." In early 1952, eleven countries created CERN's provisional council, comprised of leaders who each tackled different aspects of building a world-class research facility.[25] Cornelis Bakker prepared plans for the CERN's first machine. Niels Bohr led on theory. Odd Dahl was tasked with envisioning even bigger infrastructure that would attract top scientific talent. Lew Kowarski spearheaded organizational development "from financial procedures to buildings and workshops."[26] Governments who wanted to host the new center were invited to submit proposals, with Geneva ultimately winning out. In the meantime, theoretical work continued in Copenhagen.[27] By the close of June 1953, the CERN Convention was finally approved with twelve countries agreeing to financial contributions based on recent net national income.[28] Practically speaking, CERN offers a clear path for institution building.

This paper is among the first of several calls to build a CERN for the information environment.[29] To be more than a call, it will need champions at the national level among several countries. Ideally, a forum of democracies would take up the banner, such as the G7 or the Organization for Economic Cooperation and Development (OECD). Like-minded champions are needed, champions who see the gap in evidence-based policymaking and the need to identify which interventions to the information environment can strengthen democracy and which can counter threats within it. These champions must also be willing to pool resources together to find answers more quickly. Given the pace of government change, this could take years. In the meantime, the wider research and policy communities can do much to make IRIE a reality.

Figure 1: Timeline for the Development IRIE Following the CERN Model

	A vision for IRIE emerges	IRIE Council created, engaging G7 states	First IRIE 'accelerators' launched	IRIE is a world-class multinational, multistake-holder research facility
IRIE	2022	EARLY 2023	LATE 2023	2025 AND BEYOND
CERN	1949	1952	1954	TODAY
	First public calls for CERN made	Eleven countries create CERN provisional council	CERN breaks ground in Geneva	CERN is one of the world's largest and most respected scientific research centers

Some of this has been already happening. The European Digital Media Observatory Working Group on Platform-to-Researcher Data Access has begun developing detailed guidelines for researcher data access. The University of Michigan is developing the Social Media Archive, accepting data deposits, and making more data available for research purposes.[30] Princeton University and the Carnegie Endowment for International Peace have begun developing IRIE's shared engineering infrastructure.[31] But the work needs to be done in a more structured and collaborative way. There is a long road ahead, and many communities will be needed to make an IRIE in CERN's image.

Moving down that road means getting to reasonable answers on a range of questions, such as the following, while recognizing that global consensus may be an unattainable goal.

1. **How should the information environment be framed and studied?** The information environment is often conflated with just digital media, but it is much more complex. Work needs to be done to frame the information environment and build consilience across the fields currently studying aspects of it. One option is the concept of information ecology, which borrows from physical ecology to provide a framework for systematizing research such that knowledge builds over time to generate a shared understanding of the information environment and how it works.[32] Ideally, this question would be addressed by an interdisciplinary committee of researchers.

2. **What kinds of data can be used to study the information environment?** While greater access to social media data has long been a request from the research community, many other data sets already exist that can help put the information environment into context.[33] The Media and Democracy Data Cooperative collects political content in text from a variety of sources.[34] Both the World Bank and the OECD have collected data related to education levels, access to technology, literacy rates, and more. Other organizations have ranked countries on access to information and press freedoms.[35] Polling companies have long tried to make sense of people's perspectives, tracking things like trust in institutions over time.[36] Firms already track broadcast media content and research media consumption habits.[37] These data sets can be combined to offer snapshots of individual information ecosystems to help researchers put their studies into a wider context. Work must be put into identifying those data sets and developing a framework for aggregating them.

3. **How will legal regimes governing aspects of the information environment affect IRIE?** Federal and regional governments are passing data protection laws that may impact IRIE's ability to access and use data (even public data) collected and stored by online platforms.[38] Additionally, social media companies are increasingly seeking clarity from regulators regarding when and how they can share data with researchers while avoiding consumer protection violations.[39] It will be essential that lawmakers and regulators from democratic nations be involved as the data sharing and storage policies for IRIE are finalized. This will also be helpful in encouraging better data portability features within social media platforms and more voluntary data transparency from companies. Questions that will need answering include:

How do data protection laws interact with the sharing of social media data? When these laws vary across borders, how can they be reconciled to ensure complete data sets? How can data portability and interoperability laws support consistent standards? How can data coming from different sources be cleaned up and made useable across different data sets? How do researchers in one country study online harm in other parts of the world while respecting the data rights of all individuals?

4. **How can citizens be better engaged in research on the information environment?** Given the role citizens play in democracies, democratizing how the information environment is studied also requires citizen engagement. Researchers at Northeastern University are engaging citizens to help study online behavior.[40] Can the process of sharing user-generated data lead to further engagement in research projects and related outputs? What collaborative opportunities might help shed light on ways technology might foster democracy and citizen participation?

5. **How can greater diversity of research be fostered?** Existing research is almost entirely from and about Europe and North America. How can IRIE help build meaningful research capacity across a variety of countries and cultural contexts? How can IRIE best speed shared learning and sustain connections between researchers outside of prominent Western institutions? What are the obstacles to achieving this? Engaging with initiatives like the Centre for International Governance Innovation's Global Platform Governance Network and the International Conference on Information & Communication Technologies and Development to convene researchers is surely part of the answer.[41] Regionally, research centers such as the Doublethink Lab in Taiwan are building up capacity on specific topics such as influence operations in the wider Indo-Pacific region and can be conduits to specialized research communities.[42]

6. **What is the relationship of data storage to the research lifecycle?** What is the ongoing relationship between citizens whose data is studied at IRIE and its storage throughout the research lifecycle, including after a project is completed? How long will data be stored and what protections will be put in place for its archival use? What limits do national or regional data protection laws put in place around data retention, and how do these affect the viability of different kinds of research? What recourse do citizens have to control their data? How will data collected by IRIE be secured?

7. **What are the most pressing research questions that need to be answered and how can they be studied?** Given limited resources, consensus must be built on which research questions and measurement experiments are the most crucial to democratic societies. Efforts such as the International Observatory on Information and Democracy, which hope to present an overview of scientific research, might take on such a guiding role. Likewise, the European Media and Information Fund has been identifying research questions and funding projects focused specifically on disinformation.[43]

8. **What ethical principles should govern this research?** The information environment is poorly understood, and there is no consensus on principles for researching it. The more that is understood about the information environment and how it can be influenced, the greater the chances that research about it can be abused. For these reasons, an ethical framing is needed to underpin IRIE, one that must incorporate a variety of national guidelines that govern research involving human subjects, including personal data.[44]

9. **What kinds of large-scale engineering infrastructure can speed up that research?** Many tools can be built, but identifying what would be most helpful and needed both now and over time will help prioritize development. Princeton and Carnegie have already begun conducting scoping studies to help inform IRIE.[45] Some necessary tools include enhanced image and video processing capacity. Here, initiatives such as the Media Verification Team in Greece and vera.ai in the EU have begun work.[46] Others might help those analyzing narratives in real-time to better understand their spread and impact, for example, in crisis situations. Beyond new tools, though, there are a host of existing tools housed at a variety of institutions, which, if affiliated with IRIE by either moving to a shared facility or becoming part of an official network, could help create a suite of resources in relatively short order.

10. **How can a feedback loop be best created such that research informs policymaking?** Scaling up research will help fill the evidence gap, but that evidence must be translated into policy and social action to be implemented by governments, regulators, civil society, and platforms. Academics have limited incentives to write on the policy implications of their work and often lack the background to do so. Policymakers and practitioners are rarely steeped in all academic literature coming out about the information environment, particularly given that challenges around definitions in the field make it harder to track. What mechanism should a new institute put in place to help bridge this research-policy divide while remaining firmly apolitical? What resources are required to undertake translation between the two communities? The Centre for the Study of Democratic Institutions at the University of British Columbia has been exploring these gaps in the context of connecting researchers and transparency reporting, and the Knight Foundation's Knight Research Network has been bringing many stakeholders together to bridge the research-policy divide in the United States.[47] The European Commission's Joint Research Centre has also been developing frameworks for "evidence users" (policymakers) and "evidence providers" (researchers) to better support each other in using science to inform policymaking.[48]

11. **How can capacity be developed to help researchers make use of IRIE and new opportunities to data access offered by the DSA?** What kinds of training and support programs are needed to ensure researchers can use the tools developed by IRIE? What sort of systemic shifts are needed? Are existing funding models fit for purpose? Here again, the European Digital Media Observatory is developing

training programs for researchers, and the Center for Democracy & Technology has done excellent work tracking and analyzing legal frameworks in the context of supporting research.[49]

12. **What is the impact of this research on the physical environment?** Does centralizing infrastructure in one place decrease the carbon footprint of conducting research on the information environment? What other measures can be taken to reduce IRIE's carbon footprint?[50]

Existing communities of experts have been working on many of these questions for some time. These communities need to be brought together to work toward a shared aim of ensuring the integrity of the information environment to foster democracy.

Next Steps

The first step is finding a champion in the form of at least one democracy to support the overall coordination effort. This can be the most challenging type of work for which to find funding, as the value of coordination tends to be underestimated. Yet the key to the success of IRIE is building a multistakeholder community to make the vision a reality. The coordination phase should include creating an IRIE council that would act as a secretariat to oversee the development and operations of the institute. Initially, this council might comprise leaders tackling the questions listed above. As IRIE emerges as a world-class research facility with support from member countries, the council's composition might shift to incorporate a mix of rotating representation of the multistakeholder community it serves—for example, an equal number of members representing academia, civil society, and the group of member countries. Similarly, an equivalent representation from the tech sector might be included as observers to engage, but not control or influence, IRIE operations. The ongoing selection of the council could itself be an exercise in democracy, with processes for selection and engagement being developed with participation from stakeholder communities. The member states might opt for a structured rotation set on a schedule. The aim is to create an inclusive and functional council to which IRIE staff will report and to foster independence of research and governance (in a self-organizational approach similar to CERN's). As with CERN's leadership, the IRIE council might be supported by working groups that, for example, help prioritize among research questions and new infrastructure developments.

In terms of engaging governments, as a starting point, member states might include the G7, the European Union, or individual countries like New Zealand. IRIE should pursue a globally representative membership with a focus on enabling research in less

well-resourced countries. But a criterion of government participation should be a country's commitment to liberal democracy. As with CERN, member countries would commit to annual contributions that could be scaled according to each state's latest GDP. For IRIE to work, however, it cannot be caught in an endless renewal loop, which renders an initiative susceptible to political shifts. Research on the information environment does involve longer-term efforts that require stability, such as understanding the long-term impacts of interventions like media literacy.

Government alone should not be expected to support IRIE, which will pursue diverse funding streams. The tech sector could support IRIE through an endowment, thus ensuring the independence of researchers. Other approaches might include pay-to-play models whereby researchers contribute a portion of their research grant funding to have access to IRIE infrastructure but also raising funds through training programs, endowed and named chairs, and unrestricted gifts from donors. Indeed, nongovernmental donors have already stepped into the fray. Generous support from the William and Flora Hewlett Foundation, Craig Newmark Philanthropies, the John S. and James L. Knight Foundation, and Microsoft enabled the background research and outreach that have contributed to this paper and the development of IRIE's first infrastructure. IRIE could also provide paid-for support to solve a specific policy issue or test interventions at the behest of donors, so long as those projects meet the ethical standards to be set by the IRIE Council. This could include identifying ways to test the impact of interventions made as part of the EU Code of Practice on Disinformation.

Figure 2: IRIE Council

GREATER DIVERSITY · DATA ACCESS · INFRASTRUCTURE · IRIE COUNCIL · ETHICS · POLICY LOOP

GOVERNMENT

INDUSTRY
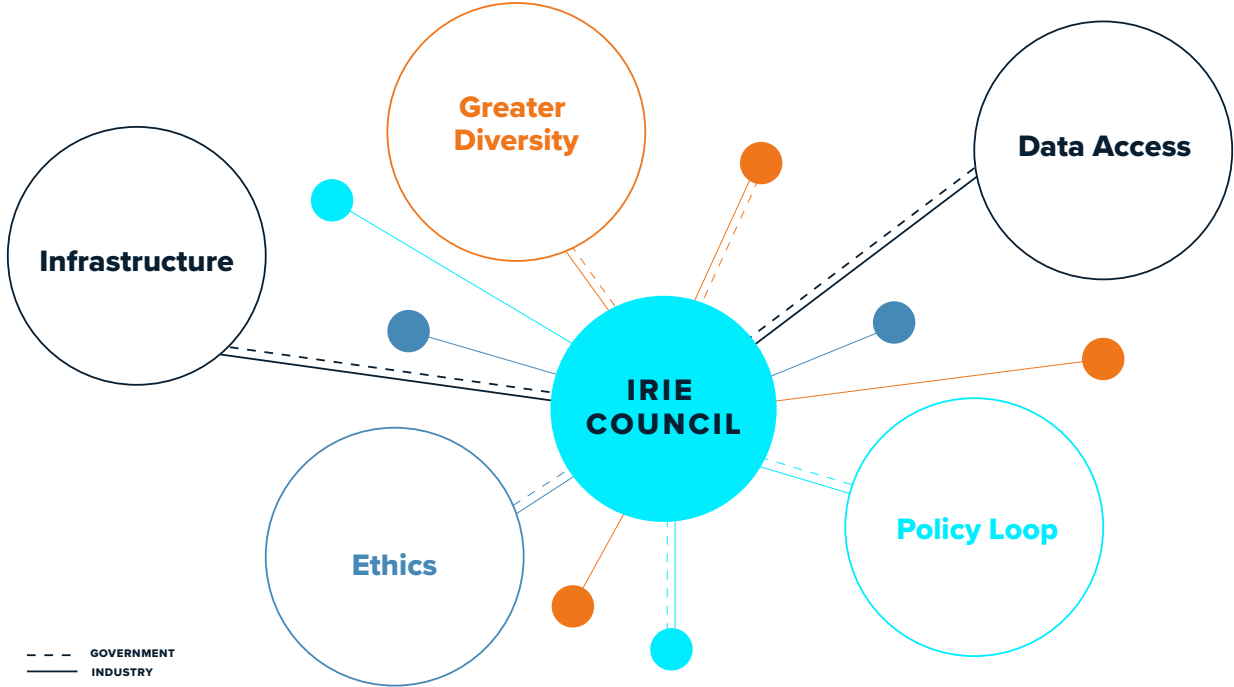
As with CERN, ultimately IRIE will be housed in a physical location. Ideally, this will be in a democracy that has privacy and data protection laws that enable IRIE's work while protecting citizens and their data. As with CERN, IRIE will continue to work with affiliated institutions around the world, supporting and scaling up research as part of a community. Structurally, IRIE will consist of several interrelated departments that oversee aspects of operations including the theoretical framing for studying the information environment, maintaining and developing data and infrastructure, translating research into policy, and capacity building for the wider community of stakeholders including researchers and policymakers. And while this is a significant effort, work has already started, in the form of planned research infrastructure with the aim to launch tools next fall. In other words, with the right support, the creation of IRIE as a physical space need not take years, as it did with the CERN, but months.

Conclusion

For nearly seventy years, CERN has been a center of gravity for physics and a model for how to support large-scale research collaboration across numerous different countries. Given the challenges facing democracy today related to the information environment, a similar level of effort is required for research on the information environment. We sit at a crossroads with democracy in peril, our very options for intervening to threats like disinformation threatening to exacerbate the situation. Wouldn't it be better if we were able to understand how the information environment works and what conditions within it fostered democracy? IRIE can help, but we need to act fast.

About the Authors

Alicia Wanless is the director of the Partnership for Countering Influence Operations at the Carnegie Endowment for International Peace, which aims to foster evidence-based policymaking to counter threats within the information environment. Alicia also leads a multistakeholder network in partnership with the G7 Rapid Response Network to support efforts in Ukraine.

Jacob N. Shapiro is professor of politics and international affairs at Princeton University. His research covers conflict, economic development, and security policy. He directs the Empirical Studies of Conflict Project, a multi-university consortium that studies politically motivated violence in countries around the world.

Notes

1   "Where Did It All Begin?," CERN, https://home.cern/about/who-we-are/our-history#:~:text=CERN's%20origins%20can%20be%20traced%20to%20the%201940s&text=Their%20vision%20was%20both%20to,unity%20in%20post%2Dwar%20Europe.

2   We analyzed a number of possible model institutions as part of this exploration. For further information see: Kamya Yadav and Jen Rosiere Reynolds, "Financial Models of Large-Scale Scientific Instruments and Organizations," Institute for Research on the Information Environment, June 11, 2022, https://informationenvironment.org/wp-content/uploads/2022/09/RA2-Financial-Models-of-Large-Scale-Scientific-Instruments-and-Organizations.pdf; Kristen DeCaires Gall and Diego A. Martin, "Peer Review and Access Models for Large-Scale Scientific Instruments," Institute for Research on the Information Environment, July 10, 2022, https://informationenvironment.org/wp-content/uploads/2022/09/RA1-Peer-Review-and-Access-Models-for-Large-Scale-Scientific-Instruments.pdf; and Jacob N. Shapiro, Natalie Thompson, and Alicia Wanless, "Research Collaboration on Influence Operations Between Industry and Academia: A Way Forward," Carnegie Endowment for International Peace, December 3, 2020, https://carnegieendowment.org/2020/12/03/research-collaboration-on-influence-operations-between-industry-and-academia-way-forward-pub-83332.

3   "Institute for Research on the Information Environment," Carnegie Endowment for International Peace, 2022, https://carnegieendowment.org/irie.

4   "CERN Answers Queries From Social Media," CERN, https://home.cern/resources/faqs/cern-answers-queries-social-media.

5   Ceri Perkins, "CERN: Everything You Need to Know," *Science Focus*, June 22, 2022, https://www.sciencefocus.com/science/cern.

6   "Home," CERN, https://home.cern/about.

7   Victoria Smith and Natalie Thompson, "Survey on Countering Influence Operations Highlights Steep Challenges, Great Opportunities," Carnegie Endowment for International Peace, December 7, 2020, https://carnegieendowment.org/2020/12/07/survey-on-countering-influence-operations-highlights-steep-challenges-great-opportunities-pub-83370; and Kamya Yadav, "Countering Influence Operations: A Review of Policy Proposals Since 2016," Carnegie Endowment for International Peace, November 30, 2020, https://carnegieendowment.org/2020/11/30/countering-influence-operations-review-of-policy-proposals-since-2016-pub-83333.

8   It is not that money is unavailable for work on the information environment; the question is around the efficiency of current approaches. In 2021 the U.S. Department of Defense gave nearly $1 billion to a company called Peraton to "counter misinformation from U.S. adversaries," a fraction of which could support significant investments in research to understand the information environment. By comparison, the National Science Foundation has an annual operating budget of $8.8 billion, supporting all types of scientific research well beyond the information environment alone. See John Hewitt Jones, "DOD Awards $1B Contract to Peraton to Counter Misinformation," *FedScoop*, August 11, 2021, https://www.fedscoop.com/dod-awards-1b-contract-to-peraton-to-counter-misinformation; and "At a Glance," National Science Foundation, accessed October 18, 2022, https://www.nsf.gov/about/glance.jsp.

9   Mike Poole, John Dainton, and Swapan Chattopadhyay, "Cockcroft's Subatomic Legacy: Splitting the Atom," *CERN Courier*, November 20, 2007, https://cerncourier.com/a/cockcrofts-subatomic-legacy-splitting-the-atom/#:~:text=In%20April%201932%20John%20Cockcroft,while%20working%20on%20cosmic%20rays.

10   Lotte Pummerer, Robert Böhm, Lau Lilleholt, Kevin Winter, Ingo Zettler, and Kai Sassenberg, "Conspiracy Theories and Their Societal Effects During the COVID-19 Pandemic," *Social Psychological and Personality Science* 13, no. 1 (2022): 49–59, https://doi.org/10.1177/19485506211000217; Holly Shannon, Katie Bush, Paul J. Villeneuve, Kim GC Hellemans, and Synthia Guimond, "Problematic Social Media Use in Adolescents and Young Adults: Systematic Review and Meta-analysis," *JMIR Mental Health* 9, no. 4 (2022): e33450, https://doi.org/10.2196/33450; Elia Abi-Jaoude, Karline Treurnicht Naylor, and Antonio Pignatiello, "Smartphones, Social Media Use and Youth Mental Health," *Canadian Medical Association Journal* 192, no. 6 (2020): E136–E141, https://doi.org/10.1503/cmaj.190434; and Elizabeth J. Ivie, Adam Pettitt, Louis J. Moses, and Nicholas B. Allen, "A Meta-analysis of the Association Between Adolescent Social Media Use and Depressive Symptoms," *Journal of Affective Disorders* 275 (2020): 165–174, https://doi.org/10.1016/j.jad.2020.06.014.

11   Candice L. Odgers and Michaeline R. Jensen, "Annual Research Review: Adolescent Mental Health in the Digital Age: Facts, Fears, and Future Directions," *Journal of Child Psychology and Psychiatry* 61, no. 3 (2020): 336–348, https://doi.org/10.1111/jcpp.13190.

12   Jon Bateman, Elonnai Hickok, Laura Courchesne, Isra Thange, and Jacob N. Shapiro, "Measuring the Effects of Influence Operations: Key Findings and Gaps From Empirical Research," Carnegie Endowment for International Peace, June 28, 2021, https://carnegieendowment.org/2021/06/28/measuring-effects-of-influence-operations-key-findings-and-gaps-from-empirical-research-pub-84824; Laura Courchesne, Julia Ilhardt, and Jacob N. Shapiro, "Review of Social Science Research on the Impact of Countermeasures Against Influence Operations," *Harvard Kennedy School Misinformation Review*, September 13, 2021, https://doi.org/10.37016/mr-2020-79; and Irina Jugl Friedrich Lösel, Doris Bender, and Sonja King, "Psychosocial Prevention Programs Against Radicalization and Extremism: A Meta-analysis of Outcome Evaluations," *European Journal Of Psychology Applied To Legal Context* 13, no. 1 (2020): 37–46, https://doi.org/10.5093/ejpalc2021a6.

13   Jacob N. Shapiro and Alicia Wanless, "Why Are Authoritarians Framing International Approaches to Disinformation?," Lawfare, December 28, 2021, https://www.lawfareblog.com/why-are-authoritarians-framing-international-approaches-disinformation; and Alicia Wanless, "One Strategy Democracies Should Use to Counter Disinformation," Carnegie Endowment for International Peace, March 28, 2022, https://carnegieendowment.org/2022/03/28/one-strategy-democracies-should-use-to-counter-disinformation-pub-86734.

14   "The Global State of Democracy Report 2021," International Institute for Democracy and Electoral Assistance, 2021, https://www.idea.int/gsod; and Latana and the Alliance of Democracies, "Democracy Perception Index 2021," https://latana.com/democracy-perception-index.

15   "2022 Edelman Trust Barometer: The Cycle of Distrust," Edelman, 2022, https://www.edelman.com/trust/2022-trust-barometer; and Brian Deagon, ""Why Users Stick With Social Media Despite a Serious Lack of Trust," *Investor Business Daily*, August 12, 2021, https://www.investors.com/news/technology/social-media-companies-flop-in-trust-survey.

16   Peter Gluckman and Kirstiann Allen, "Understanding Wellbeing in the Context of Rapid Digital and Associated Transformations: Implications for Research, Policy and Measurement," The International Network for Government Science Advice, August 2018, https://www.ingsa.org/wp-content/uploads/2018/10/INGSA-Digital-Wellbeing-Sept18.pdf.

17   "Its Your Data. Use It for Change.," Mozilla Rally, https://rally.mozilla.org.

18   Daniel Muise, Homa Hosseinmardi, Baird Howland, Markus Mobius, David Rothschild, and Duncan J. Watts, "Quantifying Partisan News Diets in Web and TV Audiences," *Science Advances* 8, no. 28 (2022): https://www.science.org/doi/10.1126/sciadv.abn0083.

19   Victoria Smith and Jen Rosiere Reynolds, "Existing Initiatives' In-House Technical Capabilities," Institute for Research on the Information Environment, June 7, 2022, https://informationen-vironment.org/wp-content/uploads/2022/09/RA3-Existing-Initiatives-In-House-Technical-Capabilities.pdf.

20   S. Dixon, "Most Popular Social Networks Worldwide as of January 2022, Ranked by Number of Monthly Active Users," Statista, July 26, 2022, https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users.

21   "Use Automatic Captioning," YouTube Help, Google, https://support.google.com/youtube/answer/6373554#zippy=%2Cautomatic-captions-on-long-form-videos-and-shorts.

22   "Report of the European Digital Media Observatory's Working Group on Platform-to-Researcher Data Access," European Digital Media Observatory, May 31, 2022, https://edmoprod.wpengine.com/wp-content/uploads/2022/02/Report-of-the-European-Digital-Media-Observatorys-Working-Group-on-Platform-to-Researcher-Data-Access-2022.pdf.

23   "Train, Educate and Engage," CERN, https://home.web.cern.ch/about/what-we-do/train-educate-engage.

24   "Origins," CERN, https://timeline.web.cern.ch/origins; "Breaking Ground," CERN, https://timeline.web.cern.ch/breaking-ground; "CERN's First Accelerator — the Synchrocyclotron — Starts Up," CERN, https://timeline.web.cern.ch/cerns-first-accelerator-synchrocyclotron-starts.

25   "Origins," CERN.

26   "Early Beginnings," CERN, https://timeline.web.cern.ch/early-beginnings.

27   "Where to Build?," CERN, https://timeline.web.cern.ch/where-build.

28   "The Convention Is Complete," CERN, https://timeline.web.cern.ch/convention-complete.

29   Stephan Lewandowski, Laura Smillie, David Garcia, Ralph Hertwig, Jim Weatherall, Stefanie Egidy, Ronald E. Robertson, Cailin O'Connor, Anastasia Kozyreva, Philipp Lorenz-Spreen, Yannic Blaschke, and Mark Leiser, "Technology and Democracy: Understanding the Influence of Online Technologies on Political Behaviour and Decision-making," EUR 30422 EN, Publications Office of the European Union, 2020, https://publications.jrc.ec.europa.eu/repository/handle/JRC122023; Anastasia Kapetas, "The Case for a 'Disinformation CERN,'" The Strategist, Australian Political Science Institute, May 18, 2021, https://www.aspistrategist.org.au/the-case-for-a-disinformation-cern;Alicia Wanless, and Jacob N. Shapiro, "Canada Needs Academic, Industry Collaboration to Fix the Internet, Not Rushed Legislation," Hill Times, May 20, 2021, https://www.hilltimes.com/2021/05/20/canada-needs-academic-industry-collaboration-to-fix-the-internet-not-rushed-legislation/298322.

30   "Democratizing Social Media Data for Research Impact," Inter-university Consortium for Political and Social Research, University of Michigan, June 15, 2022, https://www.icpsr.umich.edu/web/about/cms/4081.

31   "Carnegie Endowment and Princeton Launch Groundbreaking Initiative to Counter Threats to the Information Environment," Carnegie Endowment for International Peace press release, June 22, 2022, https://carnegieendowment.org/2022/06/22/carnegie-endowment-and-princeton-launch-groundbreaking-initiative-to-counter-threats-to-information-environment-pub-87372.

32 Whitney Phillips and Ryan M. Milner, *You Are Here: A Field Guide for Navigating Polarized Speech, Conspiracy Theories, and Our Polluted Media Landscape* (Boston: MIT Press, 2021), https://direct.mit.edu/books/book/5041/You-Are-HereA-Field-Guide-for-Navigating-Polarized.

33 Kamya Yadav, "Countering Influence Operations: A Review of Policy Proposals Since 2016"; and Victoria Smith and Natalie Thompson, "Survey on Countering Influence Operations Highlights Steep Challenges, Great Opportunities."

34 "Home," Media and Democracy Cooperative, https://mddatacoop.org.

35 "Freedom in the World," Freedom House, 2021, https://freedomhouse.org/report/freedom-world; "RSF's 2022 World Press Freedom Index: A New Era of Polarisation," Reporters Without Borders, 2022, https://rsf.org/en/rsf-s-2022-world-press-freedom-index-new-era-polarisation?year=2022&data_type=general.

36 "2022 Edelman Trust Barometer: The Cycle of Distrust," Edelman.

37 For example, see "Find It Here," TVEyes, https://www.tveyes.com.

38 See, for example, American Data Privacy and Protection Act of 2022, H.R. 8152, 117th Congress (2022), accessible at https://www.congress.gov/bill/117th-congress/house-bill/8152/text#toc-H4B489C75371741CBAA5F38622BF082DE. See especially definition 23.

39 Facebook faced an investigation by the Federal Trade Commission in the United States, but new rules were never issues. See Nicholas Confessore, "Cambridge Analytica and Facebook: The Scandal and the Fallout So Far," *New York Times*, April 4, 2018, https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html; and "FTC Imposes $5 Billion Penalty and Sweeping New Privacy Restrictions on Facebook," Federal Trade Commission press release, July 24, 2019, https://www.ftc.gov/news-events/news/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions-facebook.

40 Molly Callahan, "Can We Better Understand Online Behavior? These Researchers Will Dig Deep to Find Out," News@Northeastern, October 7, 2021, https://news.northeastern.edu/2021/10/07/exploring-online-behavior.

41 Global Platform Governance Network, Centre for International Governance Innovation, https://www.cigionline.org/activities/global-platform-governance-network; and "ICTD 2022," International Conference on Information and Communication Technologies and Development, https://ictd.org.

42 "About Doublethink Lab," DoubleThink Lab, https://doublethinklab.org.

43 Prefiguration Group of the International Observatory on Information and Democracy (IOID), Forum on Information & Democracy, https://informationdemocracy.org/working-groups/ioid; and "Debunking Disinformation in Digital Media," European Media and Information Fund, https://gulbenkian.pt/emifund.

44 Michael Zimmer, "Addressing Conceptual Gaps in Big Data Research Ethics: An Application of Contextual Integrity," *Social Media + Society* 4, no. 2 (2018); and Peter Gluckman, "Using Evidence to Inform Social Policy: The Role of Citizen-based Analytics," Office of the Prime Minister of New Zealand's Chief Science Advisor, June 19, 2017, https://dpmc.govt.nz/sites/default/files/2021-10/pmcsa-17-06-19-Citizen-based-analytics.pdf.

45 Nejla Asimovic and Jane Esberg, "Social Media Data in Conflict Research," Institute for Research on the Information Environment, May 2022, https://informationenvironment.org/wp-content/uploads/2022/09/RP2-Social-Media-Data-in-Conflict-Research.pdf; Nilima Pisharody and Jen Rosiere Reynolds, "Current Academic Research on the Information Environment," Institute for Research on the Information Environment, June 17, 2022, https://informationenvironment.org/wp-content/uploads/2022/10/RP1-Current-Academic-Research-on-the-Information-Environment-1.pdf; Nils B. Weidmann, Margaret E. Roberts, Zachary Steinert-Threlkeld, and Sebastian Hellmeier, "Scoping the Institute for Research on the Information Environment," Institute for Research on the Information Environment, June 17, 2022, https://informationenvironment.org/wp-content/uploads/2022/09/RP4-Scoping-the-Institute-for-Research-on-the-Information-Environment.pdf; Darren L. Linvill and Patrick L. Warre, "A Survey of

Public-Oriented Organizations Analyzing Social-Media Disinformation," Institute for Research on the Information Environment, 2022, https://informationenvironment.org/wp-content/uploads/2022/09/RP3-A-Survey-of-Public-Oriented-Organizations-Analyzing-Social-Media-Disinformation.pdf; Samantha Bradshaw and Bridget Barrett, "Civil Society Organizations' Data, Access, and Tooling Needs for Social Media Research," Institute for Research on the Information Environment, June 1, 2022, https://informationenvironment.org/wp-content/uploads/2022/09/RP5-Civil-Society-Organizations-Data-Access-and-Tooling-Needs-for-Social-Media-Research.pdf; Cody Buntain, "Accelerating Research With Multi-National, Multi-Platform Image Archives," Institute for Research on the Information Environment, June 16, 2022, https://informationenvironment.org/wp-content/uploads/2022/09/RP7-Accelerating-Research-with-Multi-National-Multi-Platform-Image-Archives.pdf; Kamya Yadav and Alicia Wanless, "Social Listening Companies and Access to Sensitive Data," Institute for Research on the Information Environment, May 26, 2022, https://informationenvironment.org/wp-content/uploads/2022/09/PE2-Social-Listening-Companies-and-Access-to-Sensitive-Data.pdf; and Victoria Smith and Jen Rosiere Reynolds, "Existing Initiatives' In-House Technical Capabilities."

46   Giorgos Kordopatis-Zilos, "How Similar Are Two Videos?," MeVer, August 22, 2019 https://mever.gr/2019/08/22/how-similar-are-two-videos; and vera.ai, European Union, www.veraai.eu.

47   "Centre for the Study of Democratic Institutions,"Centre for the Study of Democratic Institutions, University of British Columbia, https://democracy.ubc.ca; "Knight Research Network," Knight Foundation, https://knightfoundation.org/democracy-in-the-digital-age; and Heidi Tworek, "Platform Transparency Requires Research—But Who Will Do It?," Center for International Governance Innovation, September 22, 2022, https://www.cigionline.org/articles/platform-transparency-requires-research-but-who-will-do-it.

48   Florian Schwendinger, Lee Tops, and Viktoria Kovacs, "Competences for Policymaking," EUR 31115 EN, Publications Office of the European Union, 2022, https://op.europa.eu/en/publication-detail/-/publication/f6ba21e3-ee15-11ec-a534-01aa75ed71a1/language-en.

49   "Who We Are," Center for Democracy & Technology, https://cdt.org/who-we-are.

50   Steven Gonzalez Monserrate, "The Cloud Is Material: On the Environmental Impacts of Computation and Data Storage," *MIT Case Studies in Social and Ethical Responsibilities of Computing* winter 2022, https://doi.org/10.21428/2c646de5.031d4553.

Carnegie Endowment for International Peace

The Carnegie Endowment for International Peace is a unique global network of policy research centers in Russia, China, Europe, the Middle East, India, and the United States. Our mission, dating back more than a century, is to advance peace through analysis and development of fresh policy ideas and direct engagement and collaboration with decisionmakers in government, business, and civil society. Working together, our centers bring the inestimable benefit of multiple national viewpoints to bilateral, regional, and global issues.

Partnership for Countering Influence Operations

The Partnership for Countering Influence Operations (PCIO) is a multistakeholder initiative focused on fostering evidence-based policymaking to address issues within the information environment at the Carnegie Endowment for International Peace. Since launching in January 2020, PCIO has built an active community across governments, industry, academia, and civil society.

Princeton University

Princeton University is a vibrant community of scholarship and learning that stands in the nation's service and in the service of humanity. Chartered in 1746, and known as the College of New Jersey until 1896, it was British North America's fourth college. Princeton is an independent, coeducational, nondenominational institution that provides undergraduate and graduate instruction in the humanities, social sciences, natural sciences, and engineering.

Empirical Studies of Conflict Project

The Empirical Studies of Conflict Project (ESOC) is a multi-university consortium launched in 2009 to support research on insurgency, civil war, and other politically motivated violence, worldwide. Based at Princeton University, ESOC empowers scholarship and to help address pressing security threats ranging from civil war to misinformation campaigns.