PUBLIC INTEREST RESEARCH ALLIANCE

PIRA Principles & Operational Guides

November 2019

ASPEN TECH POLICY HUB
THE ASPEN INSTITUTE

CITRIS AND THE BANATAO INSTITUTE

CITRIS POLICY LAB

The PIRA was incubated at the Aspen Institute Tech Policy Hub and is hosted at the CITRIS Policy Lab at the University of California, Berkeley

About PIRA

PUBLIC INTEREST RESEARCH ALLIANCE

The Public Interest Research Alliance (PIRA) is a proposed multistakeholder, non-binding coalition committed to the establishment of shared principles and operational guides for the appropriate collection, storage, and use of platform data for public interest research.

Questions? Comments? Want to join? Contact
Dr. Brandie Nonnecke (nonnecke@berkeley.edu),
Director, CITRIS Policy Lab, UC Berkeley

Contents

01 //

PREAMBLE

01 // PREAMBLE

Giving researchers the ability to collect and analyze data through online platforms (e.g., Facebook, Twitter, YouTube) leads to innumerable positive impacts, including identifying and rectifying human rights abuses, enabling greater understanding of the role of bots in influencing public opinion, and supporting transparency into toxic and discriminatory online behavior. Yet platforms are increasingly constraining researchers' access to their data and tools due to increased concern over security and privacy of data, as well as fear of exploitation by malevolent actors. There is a growing tension between the need to contain and remove nefarious content, such as hate speech, terrorist content, and human trafficking recruitments, and the desire to make data available for research and investigations.

The Public Interest Research Alliance (PIRA) is a proposed multistakeholder, non-binding coalition dedicated to establishing shared principles and operational guides for appropriate data access, sharing, ownership, security, and privacy standards for public interest research across varied collaboration models and research domains. PIRA is inclusive of stakeholders from academia, civil society, government, industry, intergovernmental organizations, and the public. By tapping into the expertise and priorities of

different stakeholders, the multistakeholder model can help bring to light complex and evolving challenges and solutions related to protecting the privacy and security of data collected through platforms, while also ensuring the data are accessible in ways that can enable valuable public interest research.

Building upon the foundational work of Social Science One, the Institute for Secure Sharing of Online Data (ISSOD) within the SMaPP Lab at NYU, the Human Rights Center (HRC) at UC Berkeley, the Internet Observatory at Stanford, and the Global Network Initiative, PIRA will serve as a trusted entity to develop overarching principles and guides to inform appropriate data access, sharing, ownership, security, and privacy standards for industry-researcher collaborations. At its initial meeting, PIRA will bring together stakeholders to develop the principles and operational guides. Four draft principles created as a starting point are listed below, along with guidelines to enable implementation:

Multistakeholder Collaboration

PIRA ensures the appropriateness and effectiveness of the PIRA Principles by engaging members from academia, civil society, government, industry, intergovernmental organizations, and the public. Members engage in a collaborative approach to the ongoing development, implementation, and oversight of the principles and operational guides.

Governance, Accountability, & Transparency

Members must adhere to a set of shared governance mechanisms that support accountability and transparency in data access and use for public interest research.

Responsibility

PIRA members must implement strategies to ensure responsible research practices, including implementation of robust evaluation processes to ensure that data access and use do not infringe upon the rights of data subjects.

Data Privacy, Security, & Integrity

Members must put in place appropriate governance and technical strategies to support privacy, security, and integrity in data access and use for public interest research.

More in-depth descriptions for the draft PIRA Principles and operational guides are included below.

02 //

PRINCIPLES & OPERATIONAL GUIDES

02 //

PRINCIPLES & OPERATIONAL GUIDES

2. 1 | Multi-Stakeholder Collaboration

We propose that PIRA should ensure the appropriateness and effectiveness of the PIRA Principles (hereinafter referred to as "Principles") and operational guides by engaging stakeholders from academia, civil society, government, industry, intergovernmental organizations, and the public as members. Through participation in an annual meeting, members will engage in a collaborative approach to the development, implementation, and oversight of the Principles and operational guides.

Operational Guide: PIRA will be guided by a multistakeholder governance structure to better ensure accountability in oversight and implementation of the Principles and operational guides over time. PIRA will be composed of three bodies: the advisory board, the multistakeholder working groups, and members.

Advisory Board: The Advisory Board will consist of at least one member from all relevant stakeholder groups: academia, civil society, government, industry, intergovernmental organizations, and the public. Participation on the Advisory Board will be for a two-year term, with opportunity for extension up to one additional year. Appointees will be voted in by a majority vote from all members. Appointments will be staggered to ensure year-to-year consistency in leadership. In addition to guiding the operation of PIRA, the Advisory Board will review the Multistakeholder Working Group annual assessments of (1) the performance of PIRA, (2) the appropriateness of the PIRA Principles and operational guides and suggested revisions, and (3) members' compliance with the Principles and operational guides.

Multistakeholder Working Group: The main operating body of PIRA will be the Multistakeholder Working Group, which will be composed of four sub-groups that each contain at least one member from all stakeholder groups: academia, civil society, government, industry, intergovernmental organizations, and the public. Participation in the Multistakeholder Work-

ing Groups will be for a two-year term, with opportunity for extension up to one additional year. Any member organization will have the opportunity to join at least one sub-group. Appointments will be staggered to ensure year-to-year consistency in leadership. The Multistakeholder Working Sub-Groups will include:

1. The Evaluation Sub-Group, tasked with evaluating the performance of the PIRA organization, which includes writing an evaluation report, providing it to the Advisory Board for review, and publishing it on the PIRA website. The Evaluation Sub-Group will meet twice per year (every six months), once at the annual meeting and again six months later to review the PIRA organization;
2. The Operationalization Sub-Group, tasked with working with the advisory board and members to draft the operating Principles and operational guides, evaluate the appropriateness of the Principles and operational guides and suggest changes/revisions, report findings to the Advisory Board and members, and publish the Principles and operational guides on the PIRA website. The Operationalization Sub-Group will meet twice per year (every six months), once at the annual meeting and again six months later to review the Principles and operational guides;
3. The Advisory Sub-Group, tasked with advising members on implementation of the Principles and operational guides. The Advisory Sub-Group will meet monthly to review participating member inquiries about how best to comply with the Principles and draft guidance in response. The Advisory Sub-Group can collaborate with the Operationalization Sub-Group to ensure correct interpretation of the Principles and operational guides; and
4. The Compliance Sub-Group, tasked with working with members to develop annual assessment reports on platform and research members' compliance with the Principles and operational guides. The Compliance Sub-Group will provide the annual assessment reports to independent assessors (e.g., Osborne Clark and SSP Blue), who will evaluate members' compliance and issue final reports. The Compliance Sub-Group will meet every other month to work with members to develop their annual assessments. The independent assessors will be vetted by the Advisory Board.

Members: Members include stakeholders from academia, civil society, government, industry, intergovernmental organizations and the public. Members sign on to assist in drafting, refining, and implementing the Principles and operational guides.

Requirements:

- PIRA members have the opportunity to nominate representatives to the Advisory Board and the Multistakeholder Working Group. Each founding member organization will be allowed to appoint a maximum of two individuals, one each to the Advisory Board and to a Multistakeholder Working Sub-Group at the initiation of PIRA. Afterwards, a voting process will be implemented to elect new members to the Advisory Board and Multistakeholder Working Sub-Groups.
- PIRA members are required to actively participate in PIRA meetings. There will be an annual PIRA meeting to review the Principles and operational guides and discuss emerging issues and opportunities.
- PIRA members are expected to document their efforts to implement the Principles and operational guides into their internal policies and practices to support public interest research collaborations.
- PIRA members are expected to collaborate with the Compliance Sub-Group to develop a self-assessment to determine whether their internal policies and practices for public interest research collaborations adhere to the Principles and operational guides.

Stakeholders

PIRA should include representatives from:

- Academia
- Civil society
- Government
- Industry
- Intergovernmental Organizations
- Public

We are seeking founding member organizations or individuals, including at least one representative from each stakeholder category listed above. PIRA is expected to be launched in spring 2020.

2. 2 | Governance, Accountability, & Transparency

Members must adhere to a set of shared governance mechanisms that support accountability and transparency in data access and use for public interest research.

Operational Guide: Members must adhere to the PIRA governance process to ensure collaborative engagement of all members and to support accountability and transparency in data access and use for public interest research.

Governance: PIRA requires that members adhere to collectively defined shared governance mechanisms that define the roles and responsibilities of members. At the initial meeting, founding members will define governance mechanisms that at a minimum include:

- Equal representation and influence of relevant stakeholder groups in PIRA operational bodies, meetings, and outputs.
- Defined voting mechanisms, including the potential application of collective decision-making through consensus or majority vote.
- Defined commitments of members to the PIRA operational bodies.

As noted earlier, platform and research members must provide an annual assessment of their adherence (or not) to the Principles and operational guides. Reporting is critical to the function of PIRA as it enables lesson sharing and further development of the Principles and operational guides. These requirements are further outlined below.

Accountability: PIRA seeks to support accountability in data stewardship decisions by having platform members document and publicly communicate the process and criteria they use to make decisions about what data are to be made available for public interest research, any restrictions on data access and use, and who will be selected to receive data and other outputs.

To better ensure accountability to all stakeholders, difficult data stewardship decisions should be reviewed by a trusted intermediary. The Advisory Sub-Group can serve as a trusted intermediary between the data controller (e.g., the platform), researcher, and data subject by taking into account the researchers' goals for accessing and analyzing data and legal and social considerations necessary to ensure that the data access and use do not infringe on data subjects' rights. Members of the Advisory Sub-Group will be recruited to serve on review panels. Members will be required to disclose potential conflicts of interest. If a significant conflict of interest is identified, that individual will not be allowed to serve on the review panel. If needed, additional members for the Advisory Sub-Group will be recruited to ensure an adequately sized review panel (3-5 individuals).

Transparency: In order to ensure transparency, members will be asked to provide an annual self-assessment of their implementation of the Principles and operational guides to the Compliance Sub-Group. The Compliance Sub-Group will work with members to develop annual assessment reports on members' compliance.

The Compliance Sub-Group will provide the annual assessment reports to independent assessors (e.g., Osborne Clark and SSP Blue), who will eval-

uate members' compliance and issue final reports. These reports will be presented to the PIRA Advisory Board for review and approval before being published in an annual report on the PIRA website. While legal and proprietary restrictions may prohibit some members' ability to fully communicate their implementation of the PIRA principles, members are expected to exercise best efforts in documenting and communicating their implementation of the principles and operational guides (or lack thereof). Members should also publish their own transparency reports indicating their adherence (or lack thereof) to the principles and operational guides.

2. 3 | Responsibility

PIRA members must implement strategies to enable platforms and researchers to support responsible research practices, including implementation of robust evaluation processes to ensure that data access and use do not infringe upon the rights of data subjects.

2.3.1 | Platforms

Platforms must put in place rigorous precautions to ensure that data access and use do not infringe upon the rights of data subjects and, to the extent possible, data are made available in appropriate formats for public interest research.

Operational Guide: The following should be integrated into platforms' core policies, procedures, and practices:

Adherence: Participating platforms should ensure the Principles and operational guides are implemented into their core decision-making on data access and sharing to ensure support for public interest research over time. Participating platform companies should ensure that senior decision-makers—such as boardmembers, senior officers, and others responsible for key decisions that impact data access, sharing, ownership, security, and privacy for public interest research—are fully informed about these Principles and operational guides. These senior decision-makers must implement strategies to ensure their policies and practices align with changing data attributes, research needs, and data subjects' rights.

Sharing: Data should be shared in secure and safe ways that enable researchers to carry out public interest research. Descriptions for possible storage models and tiered access controls, data security, and privacy are included below in the "Data Privacy, Security, & Integrity" section.

Data Ethics Review: All research projects should undergo a data ethics

review process before research commences. This can include review by a panel of experts on the human rights impacts of the types of data that would be shared for the research. The Advisory Sub-Group may be able to serve in this capacity if the members have sufficient expertise.

Replicability: Data used for public interest research should be stored in a way that enables replicability of studies. Replication is a key component to ensure the validity of results and to test the generalizability of the findings to other domains.

2.3.2 | Researchers

Researchers must be held to a high standard of responsibility, and must implement rigorous precautions to ensure that data access and use do not infringe upon the rights of data subjects.

Operational Guide: The following should be integrated into researchers' core policies and practices:

Ethical Standards: Research should adhere to the ethical standards of the researcher's profession. For example, academic researchers should go through a peer review process to determine the scientific merit and human subjects review to mitigate harm to data subjects. Other researchers, such as journalists and lawyers, should adhere to the ethical research processes of their respective profession. These processes should be clearly communicated to the data controller and public and presented in their research findings to better ensure accountability and transparency.

Peer Review: When possible, researchers should engage in a peer review process to evaluate the merit of their research study, including a data ethics review by a panel of experts who have expertise in the security and privacy impacts of the types of data being shared for the research. The Advisory Sub-Group can provide peer review and data ethics review if members have the relevant expertise. If the Advisory Sub-Group lacks expertise, the researcher can assemble an external data ethics review panel. The makeup of the external data ethics review panel, including potential conflicts of interest, must be documented and communicated to the Advisory Sub-Group. The determination of the external data ethics review panel must be communicated to the data controller and the Advisory Sub-Group before data can be accessed and research can commence.

Human Subjects Review: For academic researchers, research projects should undergo and receive certification from an accredited human subjects institutional review board (IRB) to ensure compliance before data can be accessed and research can commence. The IRB review process ensures

protection of the rights and welfare of human subjects. If a researcher operates in a country or institution without an accredited internal IRB process, approval should be sought from an external IRB, such as the Western IRB (see wirb.com).

Independence: Platforms (or other data controllers) should not have final approval over whether research is published; however, they will be allowed to review the results before publication. Safeguards, such as review by the Advisory Sub-Group, should be put in place to ensure the selection of projects and funding sources do not unduly influence the research and outcomes.

Replicability: Researchers must clearly document their research methods and, when possible, store data in ways that enable others to replicate and validate their findings.

2. 4 | Data Privacy, Security, & Integrity

Members must put in place appropriate governance and technical strategies to support privacy, security, and integrity in data access and use.

Operational Guide: The following should be integrated into members' core policies, procedures, and practices:

2.4.1 | Data Ethics Review

An independent data ethics review should be pursued when possible. A panel of experts who have expertise in the human rights impacts of the types of data being analyzed should be consulted before data are shared and analyzed. See operational guides on implementation of data ethics review mechanisms under the "Peer Review" section, above.

2.4.2 | Privacy, Security, & Integrity Strategies

A number of technical strategies should be explored by PIRA members to ensure privacy, security, and integrity in data access and use in public interest research:

Storage Models and Tiered Access Controls:

- Secure Enclave - A secure enclave is an isolated execution environment that uses software and hardware features to enable secure computation on sensitive data (Oasis Labs, 2018).
- Data Trust - Based in trust law, a data trust establishes a legal mechanism by which an entity can "steward, maintain and manage how data is used and shared" (Wylie & McDonald, 2019).

- Data Collaborative - A data collaborative combines data from different sectors to enable shared access in ways that create public value (Verhulst, Young, & Srinivasan, 2018).

Differential Privacy: Differential privacy is a formal mathematical framework for quantifying and managing privacy risk. It protects the privacy of individuals in a dataset by guaranteeing that no information specific to any individual data subject is revealed. Numerous companies, such as Facebook, Google, Apple, and Uber, already implement differential privacy (Wood et al., 2018).

Secure Multiparty Computation: Secure multiparty computation enables parties to perform computations on data that both parties want to keep private (e.g., matching hashes for terrorist video content that has been removed on multiple platforms) (Archer et al., 2018).

Chain of Command: Cryptographic Audit Trail: Generation of a unique identifier tied to an individual piece of data that can be used to verify its provenance, such as using cryptographic hashing or blockchain (Hoy, 2018).

03 //

KEY TERMS & DEFINITIONS

03 //

KEY TERMS & DEFINITIONS

Best Efforts	Participating members will implement reasonable steps to fully adhere to and implement the PIRA Principles and operational guides.
Data Controller	Article 4 of the EU GDPR defines a data controller as a “natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law.”
Data Processor	Article 4 of the EU GDPR defines a data processor as a “natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller.”
Data Stewardship	The management and oversight over data assets.
Data Subject	Article 4(1) of the EU GDPR defines data subject as an “ identifiable natural person” who “can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.”
Multistakeholder Governance Model	The Multistakeholder Governance Model brings together stakeholders with a shared problem or goal to discuss and develop shared strategies.

PIRA Multistakeholder Working Group	Composed of at least one representative from all stakeholder groups (i.e., academia, business, civil society, governmental and intergovernmental organizations, and the public), the PIRA Multistakeholder Working Group is the main operating body of PIRA and is composed of four sub-groups: (1) the Evaluation Sub-Group, (2) the Operationalization Sub-Group, (3) the Advisory Sub-Group, and (4) The Compliance Sub-Group. The Multistakeholder Working Group serves to guide development and implementation of the PIRA Principles and operational guide and works with members to develop an annual report to assess members' compliance with the principles and operational guides.
Platform	A website that facilitates exchanges between two or more interdependent groups, for example, social media content producers and consumers.
Personal Data	Article 4(1) of the EU GDPR defines personal data as "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person." Article 9(1) of the EU GDPR prohibits the "processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation."
Public Interest Research	Research conducted to investigate an issue of public interest concern. Public interest research can be conducted by researchers within academia, civil society, government, industry, intergovernmental organizations, legal professions, and journalism.

ACKNOWLEDGMENTS

We would like to acknowledge with gratitude the distinguished experts who took the time to talk with us and share their knowledge and advice:

- Ed Bice, CEO & Board Member, Meedan
- Betsy Cooper, Founding Director, Aspen Tech Policy Hub
- Andrew Gruen, Manager of Strategic Initiatives, Stakeholder Engagement for Non-Commercial Data Sharing, Facebook
- Simon Hania, Data Protection Officer, Uber
- Alexa Koenig, Executive Director, Human Rights Center, UC Berkeley
- Chaya Nayak, Manager of Strategic Initiatives, Head of Non-Commercial Data Sharing, Facebook
- Nate Persily, James B. McClatchy Professor of Law, Stanford Law School and Co-Chair, Social Science One
- Uttara Sivaram, Global Privacy & Security, Uber
- Joshua Tucker, Professor of Politics and Director of the Social Media and Political Participation (SMaPP) Lab, New York University

ABOUT

The CITRIS Policy Lab is a sub-organization of the Center for Information Technology Research in the Interest of Society and the Banatao Institute (CITRIS), headquartered on the UC Berkeley campus.

Founded in 2001, CITRIS leverages expertise on the campuses of UC Berkeley, UC Davis, UC Merced, and UC Santa Cruz to develop technology applications with societal and economic benefits. The CITRIS Policy Lab was established in 2018 to support interdisciplinary technology policy research analyzing technology capabilities and their implications for society. Through its collaboration with public and private sector stakeholders, the CITRIS Policy Lab addresses core questions regarding the role of formal and informal regulation in promoting innovation and amplifying its positive effects on society.

CITRISPolicyLab.org

The Aspen Tech Policy Hub is a West Coast policy incubator, training a new generation of tech policy entrepreneurs. Modeled after tech incubators like Y Combinator, the Aspen Tech Policy Hub trains tech experts in the policy process through an in-residence fellowship program in the Bay Area, and encourages them to develop outside-the-box solutions to society's problems.

AspenTechPolicyHub.org