Data Sharing for Research: Executive Summary

A Compendium of Case Studies, Analysis, and Recommendations

FPF's Compendium of Case Studies, Analysis, and Recommendations is the result of a thorough analysis of a diverse sample of data-sharing partnerships between companies and academic researchers. As part of a process that included nearly two dozen live interviews, FPF has produced a series of case studies that examine the different forms that data-sharing partnerships may take, how the parties think about the benefits of such a partnership, and what is done to mitigate risk, including risks related to the use of personal data.

The resulting report demonstrates a broad consensus regarding the potential benefits of industry/academic data-sharing partnerships, including the acceleration of socially beneficial research, enhanced reproducibility of research breakthroughs, and broader access to valuable data sets. At the same time, companies and academic researchers show that they understand and take steps to mitigate risks - particularly ethical and data protection risks. Increasingly, stakeholders are identifying risks arising from re-identification threats or data breaches while acting to mitigate those risks through the use of Data Sharing Agreements (DSAs) and Privacy Enhancing Technologies (PETs).

FPF's analysis of corporate-academic data-sharing partnerships provides practical, evidence-based recommendations for companies and researchers who want to share data in an ethical, privacy-protective way. These case studies demonstrate that corporate-academic data-sharing partnerships offer compelling benefits to companies, research, and society. Risks exist, but effective mitigation strategies can reduce the likelihood of harm to individuals, communities, and society. For many organizations, data-sharing partnerships are transitioning from being considered an experimental business activity to an expected business competency. This trend is most pronounced among established firms; it is an opportunity for researchers to access new data for scientific discovery.

Case Study Themes and Conclusions

From our case studies, FPF was able to identify ten major themes that appear across many, if not all, types of data-sharing partnerships we analyzed. For instance, we have observed that successful data-sharing partnerships use Data-Sharing Agreements that require both the company and researchers to take steps to protect privacy. Some of the challenges of data sharing include technical knowledge and infrastructure gaps between companies and researchers and the continuing need for ethics and privacy review for industry-based research. Promising practices for data sharing include the use of Privacy Enhancing Technologies and company-created, public-facing data-sharing menus to facilitate new partnerships. While data sharing has significant costs and inherent risks, the risks can be managed, and the benefits to researchers, companies, and society make data sharing worth the effort.

Building on these themes and the case studies, FPF has prepared a list of key recommendations that may help guide other organizations and researchers looking to establish a similar program or partnership. We lay out these recommendations in depth below.

For anyone interested in more information on data sharing for research, FPF offers the Ethics and Data in Research Working Group, which analyzes US legislation impacting research and data, discusses ethical and technological research challenges, and develops best practices for privacy protection, risk reduction, and data sharing in research. Learn more and join the Working Group. Explore The Playbook: Data Sharing for Research and the companion infographic for more information on data-sharing partnerships and best practices. For inquiries about this report, please contact Shea Swauger, Senior Researcher for Data Sharing and Ethics, at sswauger@fpf.org.

Recommendations for companies that share data for research or are considering a data-sharing program:

Transparency and Openness

- Create a public page that lists what data the company is willing to share, describe the data as much as possible, and update the list regularly. This could be done unilaterally or as part of a consortium of companies seeking to share more data for research.
- Be transparent about requirements that academic partners must meet and publicly post a typical Data Sharing Agreement, if possible.
- Create a public form for researchers to ask questions, request data, or initiate a partnership. This signals the kind of information the company needs about the researcher and the proposed research.

Privacy and Security

- Use Privacy Enhancing Technologies (PETs) to bolster data privacy, but select PETs judiciously so as not to exclude researchers from less-resourced institutions.
- When using PETs is impractical, reduce the data's sensitivity (through minimization, aggregation, and other techniques) to a level that enables sharing while maintaining privacy.
- Ensure cybersecurity and privacy teams co-design privacy safeguards when sharing data.
- Include metadata as part of internal privacy reviews before sharing.

Governance

- Assign multiple people with expertise in data science, statistics, research, policy, and regulatory compliance to manage data-sharing activities. This role is likely to be an extension of existing responsibilities, although, in some instances, a dedicated team might be feasible.

- Connect the general counsel, marketing/communications, and core software engineers to the data-sharing team for effective coordination.
- Adapt the Data Sharing Agreement to the amount of money and personnel available; more adaptable DSAs require more people and time.

Control

- Choose the appropriate data-sharing partnership type: Open data, closed trusted partnerships, and data intermediaries all require investment, personnel, and institutional support but can vary in duration and intensity.
- Companies should consider implementing a spectrum of data-sharing models (open data to closed trusted partnerships), which would likely lead to more collaborations and greater social impact.
- Ensure researchers have full authorial control over the publication venue and all content not directly relevant to the data.
- When appropriate, reserve the right to review data before publication to assess privacy risks, accuracy, or any analytical limitations of the data.

Recommendations for researchers interested in using data held by a company for research:

Initiative

- If a company doesn't have a data-sharing menu or public page describing its data-sharing activities, contact them to inquire about a potential partnership.
- Provide lots of communication on the front end of the partnership and plan check-ins at key points of the research lifecycle through publication.

Internal Partnerships

- When starting a data-sharing partnership, involve the university general counsel early on and check to see if the university has a standard Data Sharing Agreement or an example agreement used for a previous project.
- Contact the university's Research Integrity Office and Information Technology Office before any data is shared to ensure the university can support the project technologically and ensure regulatory compliance.
- Contact the university library for additional support for research, data management, and privacy.

Privacy and Security

- Get training on Privacy Enhancing Technologies and contact all relevant technical support offices to ensure the university can support data sharing using PETs.
- Include data transfer and privacy-related technical infrastructure in all research funding proposals or projects.

Communication

- Coordinate with the company about requirements for publishing, data sharing, data retention or deletion, citation, and promotion of the research to maximize the audience and impact.
- Maintain academic independence of any research produced from the data-sharing partnership, but, when appropriate, allow companies to review data to assess privacy risks, accuracy, or any analytical limitations of the data before publishing.