FPF Corporate-Academic Data Stewardship Research Alliance
Best Practices Data Sharing for Research
(September 2019)

The following is a list of best practices companies should consider when sharing personal data with academic researchers. Some suggested practices may not be necessary or appropriate for every data sharing situation, but each should be considered for how (or if) it will apply to a given arrangement.

These suggested practices reflect the objectives and goals of the FPF Corporate Academic Data Stewardship Research Alliance, in particular enabling the responsible sharing of data in order to advance scientific research. Thus, they favor academic independence and freedom over tightly controlled research, and they encourage broad publication and disseminating of research results (while protecting privacy of individual research subjects). At the same time, they take a strong approach to protecting the privacy of individual data subjects.

This document is provided for informational purposes only. Nothing in this document is intended to provide, nor should be construed as, legal advice and is not a substitute for obtaining professional legal counsel from a qualified attorney on your specific matter.

1. Data sharing agreement
    - Sharing personal data with academic researchers should be subject to a written contract.
    - The contract should define the scope and purpose of the research, describe the types of data being provided, and include key provisions relating to privacy, security, IP, and data ownership. Contractual provisions should reflect most or all of the best practice principles included in this document.
    - Use of standardized contractual terms should be encouraged in order to streamline the negotiation process and provide greater clarity and comfort to both parties.

2. Due diligence and oversight
    - Data suppliers should engage in due diligence before sharing data with a data recipient.
    - Due diligence may review the nature and purpose of the research, the privacy and security safeguards in place, the data recipient's track record of responsible data stewardship, etc.
    - As appropriate, data suppliers should have ongoing oversight of the protection and management of the data.

3. Data minimization and de-identification
    - Only the data needed to accomplish the research purposes should be provided.
    - Data provided should be de-identified to the greatest extent compatible with the needs and purposes of the research.
    - Because de-identification encompasses a wide range of methods and strengths, the parties should have a clear and common understanding of the de-identification technique(s) to be used and the risk-reduction outcome intended to be achieved.
    - To the extent de-identification is a compliance requirement, the applicable legal standards must be met (e.g. the HIPAA de-identification rules for U.S. health data).

- The receiving party should be prohibited from any attempt to re-identify the data.

4. Data security and integrity
    - Appropriate controls and measures should be adopted to protect the security and integrity of the data.
    - Security controls should include a range of technical controls (such as encryption, firewalls, authentication, access controls that limit data access to those researchers with a clear need, logging and auditing), as well as physical controls and operational controls.
    - The security controls should be appropriate and proportionate to the nature and sensitivity of the data involved.
    - The security controls should meet any applicable legal requirements.
    - Independent security audits or certifications, including SOC 2 and/or ISO 27110, may be required as appropriate.
    - Security questionnaires may be used as appropriate. If there exists a commonly-used model security questionnaire in a given industry or sector, it should be considered (rather than a unique or proprietary questionnaire) in order to streamline the process.
    - A data recipient should provide prompt notification to the data provider in the event it becomes aware of, or has a reasonable suspicion of, any unauthorized access or disclosure affecting the data.

5. Vendor management
    - If shared data may be accessed by vendors or service providers operating on behalf of the data recipient, such vendors should be subject to written contract that includes clear privacy and security obligations. For example:
        - The vendor may only access and process the data for the purpose of providing the defined service on behalf of the data recipient.
        - The vendor must maintain appropriate data security measures and practices; and the vendor must notify the data recipient in the event of any known or suspected unauthorized access or disclosure involving the data.
        - The vendor must return or destroy any copy of the data when the services are completed. The vendor should be required to provide to the data recipient a certificate of deletion or destruction, which should be available to the data supplier upon request.
    - Data suppliers may wish to require data recipients to inform and/or obtain approval of the data supplier for any vendor that may have access to the data.
    - For highly sensitive data, it may be appropriate to require that any vendor personnel who access the data to have passed a background check.

6. Data retention and deletion
    - In the context of academic research, the need to retain data for long periods of time is common. For example, scientific research should be replicable and testable, which may require ongoing retention and access to the underlying data.
    - Nevertheless, the parties should agree on the retention periods or the criteria used to determine retention periods.

- If certain data the researcher receives from the data supplier is later determined to not be needed for the research, the researcher should return or destroy that data promptly.
- Further, the parties should define the steps that will be taken to archive, (further) de-identify, return, or destroy the data at the end of the retention period.

7. Ethical data use
    - Data shared for research should be used only for ethical purposes. Ethical data use is a broad concept that goes beyond the issues that a traditional institutional review board (IRB) might examine.
    - Certain data uses or research purposes that may be viewed as unethical may be prohibited by the data supplier.
    - The parties should develop a common understanding of the ethical issues and boundaries related to the research. Each stakeholder should assess and fulfill its own ethical responsibilities.
    - Provisions should be put in place to identify and prevent uses that may reflect bias or result in inappropriate discrimination based on factors such as race or ethnicity, national origin or immigration status, gender or gender identity, sexual orientation or preference, etc.

8. Independent review
    - As appropriate, the research should be independently reviewed to assess privacy protections, and additional safeguards to protect the rights and interests of the data subjects, and other ethical considerations.
    - Such independent assessment may be conducted by an IRB, a peer review process, an independent ethical review board, or some equivalent body.
    - Independent reviewers should have the appropriate knowledge and expertise to effectively understand and evaluate the risks and ethical issues raised by the research.

9. Publication expectations
    - Publication and peer review of research results should be encouraged.
    - Published papers and research results should avoid including information that could adversely affect privacy.
    - For instance, examples describing specific data subjects that might reveal an identifiable individual should be avoided.
    - Even the publication of "public" data that could make obscure information more prominent and could lead to embarrassment or other harms should be avoided.

10. Training and education
    - Data recipients should provide the data suppliers with assurances researchers and other individuals who will have access to personal data have received appropriate privacy and security training covering (1) the importance of protecting the privacy and security of the personal data and (2) any specific restrictions on the data set in question.
    - The parties may agree that the data recipient will utilize existing training materials developed by the data supplier, the data recipient, or a third-party provider, to the extent appropriate training materials are available.