Model Data Use Agreements: A Practical Guide

Amy O’Hara

amy.ohara@georgetown.edu

October 26, 2020
Scope

Objectives – data sharing that is efficient, secure, equitable, and transparent

Not just data sharing – data exchange and data access

Outcomes – knowledge production, more data access, building trust
Data Use Agreements

• Documents that describe what data are being shared, for what purpose, for how long, detailing access restrictions or security protocols

• Related to Memoranda of Understanding, Data Use Licenses, Data Exchange Letters, Non-Disclosure Agreements

• Parties include data provider and data requestor
What is Required

• Clear vision, objectives
• Understood vocabulary
• Coordination across sectors and organizations
• Collaboration across roles and disciplines
Examples: Roles and Disciplines

- Some are already working together
- Need to have better information transmission
The Goal
Framework: Five Safes

- Safe Projects
- Safe People
- Safe Settings
- Safe Data
- Safe Outputs
Embedding the Five Safes into DUAs

• Intended data uses
• Terms for data access and handling
• Describe project outputs and expected publications
• Acceptable data uses, linkages, and topics of analysis
• Roles and responsibilities for the provider and researchers
• Metadata
• Reporting or disposition requirements
Preparation

• What data are available?
  – Is quality known? Is there documentation? Have other researchers used the data? What elements are captured or retained?

• How long will it take?
  – Negotiating and then getting the data. Is there an established data request process?

• Adherence to laws, regulations, and policies permitting data use
  – Are there templates or existing agreement to follow? What data sharing has been permitted before?
Needs and Wants of Both Parties

- Researchers
  - Preparation
  - Negotiation
  - Responsibilities
  - Transparency

- Providers
  - Incentives/needs
  - Documentation
  - Extraction
  - Sharing/hosting
Consideration for Data Subjects

- What are the potential benefits, costs, and risks for the data subjects?
- How will these be communicated to them?

---

<table>
<thead>
<tr>
<th>Active engagement</th>
<th>Passive engagement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-creation</td>
<td>Reporting</td>
</tr>
<tr>
<td>Reviewing</td>
<td>Commenting</td>
</tr>
<tr>
<td>Project conception</td>
<td></td>
</tr>
<tr>
<td>Accessing and/or collecting data</td>
<td></td>
</tr>
<tr>
<td>Cleaning, linking, and de-identifying data</td>
<td></td>
</tr>
<tr>
<td>Analyzing data</td>
<td></td>
</tr>
<tr>
<td>Publishing and sharing results</td>
<td></td>
</tr>
<tr>
<td>Post-project treatment</td>
<td></td>
</tr>
</tbody>
</table>
In Advance

• Read guides

• Call a privacy professional

• Consult a lawyer

• Connect the authorizers and the data people

• Do your homework
Develop the Agreement

Title
Data Use Agreement for [Data/System] Access between Party 1 and Party 2

Parties and Purpose
This Agreement is between Party 1 [Office, Agency, Department, Institution] and Party 2 [Office, Agency, Department, Institution]. Party 1 and Party 2 are entering into an Agreement that will allow the exchange of data and clarification of data access and use. Party 1 will provide data collected to Party 2 for the purposes of [specify].

Authority
Party 1 is a(n) [specify] organization whose mission is [specify]. The authority to enter into this Agreement is [xxx]. This authority permits the release of [data] [law/code] permits disclosure of [data] for [specify] functions. Party 2 is an organization whose mission is [specify].

Terms and Conditions
Description of planned data use by Party 2, consistent with Purpose above:
- Treatment of data anomalies, including technical assistance from Party 1 as needed
- Terms for storage, treatment of original data, handling of Port Information, and data linkage protocols
- Conditions for storing modified data (including integrated, recorded derived data) during and after the project
- Terms for storage of researcher generated files (including retention extent permitted by law, the original data received from Party 1, Party 2 for [specify] period)

Data Elements
The following data will be provided under this Agreement:
[Specify list of data elements from named program/systems, noting which populations, and/or geographies are sought.]

Approved Research Uses
[Describe project objectives, intended data use, expected linkages.]

Limitations on Liability
In no event shall either party be liable to the other party under this Agreement or to any third party for special, consequential, incidental, punitive, or indirect damages, irrespective of whether such claims or damages are founded in contract, tort, warranty, operation of law, or otherwise or whether claims for such liability arise out of the performance or non-performance by such party hereunder.

Monitoring and Breach Notification
In the event of an actual or suspected security breach involving its information system(s), Party 2 will immediately notify Party 1 of the breach or suspected breach and will comply with applicable breach notification laws. The parties agree to cooperate in any investigation and remedy of any such breach, including, without limitation, complying with any concerning unauthorized access or disclosure.

Remedies in Event of Breach
The parties recognize that irreparable harm may result in the event of a breach. Agreement. In the event of such a breach, the non-breaching party may be entitled to restrain the other from any continued violation. This section shall survive termination of the Agreement. In the event that a breach is identified and it is determined that the breaching party is responsible for the breach, the breaching party shall be liable for the reasonable costs incurred by the non-breaching party to meet all federal and state legal and regulatory disclosure and notification requirements, including, but not limited to, costs for investigation, attorneys’ fees, risk management, cost required individual or public notification, fines, and mitigation activities.

Public Information
To promote organizational transparency, and in support of data discovery for current and future researchers, Party 2 may publish non-sensitive data documentation to public-facing websites. This documentation may include a project abstract, description, or summary of results. Use of Name
Neither party will use the other party’s name, logos, trademarks, or other marks without that party’s written consent.

Community Stakeholders
The parties agree to engage community stakeholders in the course of this research project. No confidential data will be released or discussed with third parties, but the parties may agree to disclose de-identified aggregate reports to support their initiatives and engage community stakeholders.

Costs
This project shall not result in the transfer of funds from one party to another. Party 1 agrees to provide technical assistance to Party 2 to develop and deliver the initial data extract. If the parties determine that additional staff or supports are necessary at any stage of this research project, Party 2 agrees to seek funding to support those needs.
Identifiers

• Will user get direct identifiers

• Who can handle identifiers

• Protocols (e.g., honest broker, trusted third party)

• Maintenance/disposition

• Certifying de-identification
Security

• Requirements based on data, laws, and regulations

• Institutional requirements

• Monitoring and control

• Liability
Output Review

• Monitoring scope

• Pre-publication review

• Notification prior to release

• Capturing code, data, results, IP
Publishing Results

• Notice or approval

• Branding

• Citation
If They Say “No”

- Legal
- Policy
- Cultural
- Financial

Data breach
Conflicting results
Political backlash
Bad publicity
Burden

Project goals
One of the most common reasons data providers say no does not relate to a legal or political issue, but to a practical one. Staff members are pressed with work for core agency operations and reluctant to prepare datasets and documentation for external users. Given fixed local government infrastructure, there is likely no way to eliminate the extra work for the staff, but you can articulate benefits to demonstrate the value of data sharing that justify the costs.

- Offer additional analysis useful to the agency’s work. Agency staff might lack the time or expertise for analyzing own data for internal purposes. In return for data sharing, you can offer to fulfill simple agency requests or return enhanced files to the agency. For example, in exchange for a file of the locations of homeless prevention payments by the District of Columbia, Greater-Urban DC provided the local homeless coordination agency geocoded files and simple maps so they could explore the programs’ impact on neighborhoods.

- Offer access to relevant indicators derived from another office’s data. Agencies are often more willing to share data with an outside organization than with other agencies. For example, knowing the number of births or the number of new housing units planned by census tract helps school districts forecast enrollment, but planning and health departments might feel more comfortable sharing the data with a third party rather than directly, either for legal or political reasons. Even if the data sharing among agencies is possible, the school analysts could prefer to receive aggregate, clean data rather than doing the cleaning and analysis themselves.

- Save the agency time by answering community inquiries. Local agencies might already time fulfilling community data requests one by one. If you provide publicly available updated summary statistics in charts and maps, agency staff can refer data seekers to your website and spend less time answering inquiries. The Baltimore Neighborhood Indicators Alliance publishes its aggregate indicators on the city’s open data portal, making the data available for residents and city officials alike.

- Reassure the data owners that you have the skills to use the data. This includes letting them know that your qualified staff will not need assistance in basic data use perhaps through examples of other data analysis you have done.

- Offer to share back documentation of the files. Often local administrative data will not come with documentation of the contents or quality, so you can share back data dictionaries or diagnostics with the agencies. Agencies can use the documentation for sharing the data with other organizations in the future and might be interested in improving their data quality.
Following Through

- Getting it signed
- Agreement management
  - Compliance with the agreement
  - Modifying the agreement
- Do the work and explain results to stakeholder
- Build the trust