

DATA SHARING AGREEMENTS:

MOVING BEYOND COMPLIANCE TO ENSURE EFFECTIVE GOVERNANCE

JULY 2020



D4
AD

 **NCHEMS**



ABOUT THIS BRIEF

Data Sharing Agreements: Moving Beyond Compliance to Ensure Effective Governance explores important considerations for entering and maintaining strong data sharing partnerships. This brief offers best practices and advice for organizations in creating the agreements, navigating regulation, and developing governance structures.

The author, Patrick Lane, is vice president for policy analysis and research at the Western Interstate Commission for Higher Education (WICHE), an organization that has worked to facilitate cross-state data sharing of education and employment data in addition to other efforts to improve access to and success in postsecondary education. Information contained in this brief is drawn from lessons learned during the Multistate Longitudinal Data Exchange Effort as well as general research. The views presented in this brief are those of the author.

ABOUT DATA FOR THE AMERICAN DREAM

Data for the American Dream (D4AD) is a consortium bringing together Schmidt Futures, Lumina Foundation, Walmart Foundation, and the Walton Family Foundation. D4AD currently funds pilot initiatives in three states (Colorado, Michigan, and New Jersey) that will help provide low-income, lower-skilled, underemployed, and unemployed workers access current and actionable data, enlisting local case managers from public and private agencies to counsel job seekers, help them access needed services, and reach the most underserved populations.

The National Center for Higher Education Management Systems (NCHEMS) is the implementation partner of D4AD. NCHEMS is a private nonprofit organization whose mission is to improve strategic decision making in postsecondary education for states, systems, institutions, and workforce development organizations in the United States and abroad.



INTRODUCTION

Data for the American Dream is an ambitious initiative to take advantage of information about education and training opportunities, employment outcomes, and labor markets to give low-income, unemployed, and underemployed individuals better career opportunities. The theory behind this effort is that these individuals will be better able to make life-changing decisions about which education and training opportunities to pursue or what career pathway to follow if they have better data and information about job openings available and the outcomes of related education and training programs.

Sharing data between agencies, or across states, can be an important strategy for evaluating how different credentials benefit their holders in the workforce and is therefore a foundational strategy for states involved with Data for the American Dream. This brief examines in closer detail how data sharing agreements between entities are not only necessary for ensuring compliance with relevant statutes, but also for establishing effective governance relationships. These governance arrangements will ultimately help determine how successful—or not—the data sharing efforts are, as the way that partners interact with one another has important implications on varied areas, including privacy and data security, sustainability, research effectiveness, and more. The brief draws on lessons learned from a project called the Multistate Longitudinal Data Exchange (MLDE) managed by the Western Interstate Commission for Higher Education (WICHE) and is one in a series of pieces examining issues related to improving data infrastructure to better serve low-income, unemployed, and underemployed individuals.

BACKGROUND: WHAT DATA ARE INVOLVED?

Data for the American Dream has funded three projects to improve the information available to low-income, underemployed, and unemployed individuals as they consider potentially life changing decisions about education and training options and career pathways. These efforts will draw on myriad data sources, but one key foundation will be using information about the employment outcomes of education and training programs to present information to the target population about how different options may pay off.

These data sources are regularly combined by state agencies and have been successfully shared across state lines by multiple efforts.¹ The projects funded by Data for the American Dream may also involve other data, such as information about job openings, but this brief primarily focuses on the legal and governance arrangements developed to support sharing of data from the employment and the education and training sectors.

COMPLIANCE REQUIREMENTS FOR SHARING DATA

Linking education and training programs with employment data—primarily wages—inevitably involves the successful establishment of data-sharing agreements. Given the sensitivity of the data, one of the primary purposes of data sharing agreements is compliance with federal and state laws and regulations related to data privacy and security. Broader governance purposes are discussed in greater detail below. It goes without saying that legal counsel with expertise in the relevant statutes must be involved in drafting, reviewing, and approving these agreements.

FERPA Required Elements.

Written agreements are required before education data holders can disclose personally identifiable information (PII) to another entity under what is known as the “audit and evaluation” exemption, which is typically the section of the statute used when data are combined to show information about program outcomes.²

The U.S. Department of Education has made substantial materials and resources available to assist data holders in complying with Family Educational Rights and Privacy Act (FERPA). The Privacy Technical Assistance Center (PTAC) (provided by the Department) helps education entities comply with FERPA and regularly provides resources—as well as direct technical assistance—related to data sharing.³ The Department has also directly issued specific guidance.⁴

According to this guidance, these agreements must include the following:⁵

- Designation of Authorized Representative. The entity sharing data must formally name the receiving entity as such.
- Information about data being disclosed (and the purpose behind the disclosure). The agreements should be very specific about the information that is included and detail the purpose and justification for the disclosure, including a description of how the disclosure will help audit or evaluate a relevant education program.
- Requirements for data destruction. The agreements should specify the methodology to be employed in destroying data at the conclusion of the agreement as well as a definitive timeline.
- Policies and procedures to prevent further redisclosure. This is perhaps the most important component from a privacy and security perspective, but also the most vague.

The final requirement engenders significant discussion. The standard applied is that the entity disclosing data must ensure (through the written agreement) that the entity receiving the data uses “reasonable methods” to protect data from any further redisclosure.⁶ Given the lack of specificity, a promising strategy is to rely on other widely adopted and accepted criteria, such as those developed by the National Institute for Standards and Technology (NIST) Cybersecurity Framework.⁷

Elements Required by the Statute Regulating the Use of Unemployment Insurance Data

Using earnings data from the UI data system requires compliance with additional statutes—particularly the requirements listed under 20 C.F.R. 603. Similar to FERPA, the law allows disclosure of PII (including linking to education records) under a strict set of requirements. Essentially, state agencies are allowed to share earnings information with other state officials, including from other states, as they carry out their “official duties.”⁸ Further, data may be shared with “agents or contractors” of public officials, potentially enabling third parties to facilitate the exchange of data between two states when necessary.⁹

Data sharing agreements involving data from the UI data system are required to have the following:

- Limitations on use. The agreement must identify a specific purpose for the exchange of data and show how its use complies with the law.
- Data safety and security. Similar to FERPA, laws governing data from the UI system require that those data be kept secure, but do not detail specific criteria or processes that must be employed.
- Requirements for employee education. Staff at the receiving entity that will access data must be fully educated on the limitations of use and procedures for handling the data (and the entity must affirm that this has taken place).
- Requirements for data destruction. Similar to FERPA, the agreements must detail how and when data will be destroyed.¹⁰

Data Sharing Agreements: Conclusions.

With numerous examples of safe, secure, and compliant data sharing taking place in virtually every state these days, developing cross-state data sharing agreements is potentially more straightforward than it has been in the past. In addition to the federal examples discussed above, some states have applicable legal standards, and laws in effect in other countries are increasingly relevant.¹¹ Formal legal advice is obviously a requirement to ensure compliance. These documents should also formalize the structures and relationships that will govern data sharing efforts.



EFFECTIVE GOVERNANCE ARRANGEMENTS

The governance of complex data sharing relationships may be developed and enshrined in multiple different written agreements. The data sharing agreement developed by WICHE for the Multistate Longitudinal Data Exchange devotes substantial attention to governance matters that are generally beyond the scope of FERPA and laws governing the disclosure of UI data. Effective governance arrangements go far beyond compliance with federal and state laws because they are key for sustaining effective data sharing relationships to improve data infrastructure in ways that help low-income, unemployed, and underemployed individuals find beneficial education and career pathways. This section discusses the need for a formal body to govern the exchange, then details some of the key tasks to be overseen by this entity.

The governance regime between multiple states and agencies should be established in a written agreement (which can be the same data sharing agreement discussed above, or a separate agreement). The governance issues addressed include a range of topics and build the trust necessary for agencies to sustain partnerships with one another, protect student and employee data, and carry out effective research and evaluation. These governance components can be established and attached to data sharing agreements as appendices or exhibits, with the main body of those agreements focusing on legal compliance. This allows additional flexibility for adaptation of structures without renegotiating formal data sharing agreements, which can be time-consuming and resource intensive.

Establish Governance Bodies. A key initial step is to establish decision-making bodies that will manage the processes of exchanging data, ensure adherence to agreements, and make decisions affecting all participating entities. This governance structure is necessary well before data sharing agreements have been completed (in fact, the lead governing body will likely be responsible for developing the required agreements), but its structure should be formalized in the agreement. Key considerations are to ensure adequate representation by important characteristics of the types of entities party to the agreement, such as geography and sector, allow for participant turnover, and provide a forum for addressing unexpected issues. Depending on the number of participating entities, this can either be fairly straightforward—with all parties being represented in cases where there are few participants—or more complicated, where representative governance is required due to larger numbers.

In the case of the Multistate Longitudinal Data Exchange, the work is directed by a governance committee with a representative from each state. A subcommittee also oversees the development of the technical infrastructure.

Determine Decision-Making Processes. As noted above, the decision-making processes for complex data sharing arrangements must be carefully crafted. Given that participation is essentially voluntary, it is obviously desirable to achieve consensus on all major decisions. Processes for determining things like which data elements to include, what minimum cell sizes to apply, and when and how to allow external access to data resources can affect all participants, but decisions can be implemented in ways that allow for differences to achieve some form of consensus. For example, if some members prefer to allow external researchers to complete a research project using the data, but one member does not, the decision-making process can allow for such projects to go forward without the objecting member's data.

Define Roles and Responsibilities of Participants. Numerous different entities may participate in data sharing relationships. The governance arrangements must clearly identify the roles and responsibilities of each. As an example, in a straightforward education and employment sharing arrangement between two states, workforce and education agencies will have different responsibilities for data provision and analysis. Further, participants may have responsibilities for responding to questions and requests from other participants. Governance arrangements should clarify what is expected of participants and what represents a reasonable burden of participation.

Enforce Data Stewardship. The governance regime must also address the concept of data stewardship and how it compares to data ownership. A promising practice in data sharing relationships is that the data recipient does not, in any true sense of the word, take over ownership of the data it receives. Instead, the providing entity maintains control even though it has provided data to its partners. Should a participating agency elect to leave the data sharing relationship, the governance regime should easily allow it to do so, while requiring other participants to delete data they received from that partner.

Establish Criteria for Participation (and Dissolution of Relationships). The governance arrangement should also allow for partners to join the effort and set up processes in the event a participant elects to leave the arrangement. Criteria will likely include some base level of data availability (though it is likely that not all partners will be able to contribute all elements in their domain equally); a commitment to carry out certain actions within an agreed upon period of time; participation in common activities to maintain and develop common data infrastructure; and, where needed, an agreement to provide resources to support the exchange.

Additionally, governance arrangements must consider the potential for one or more partners to leave the relationship. Such actions raise important questions about whether other partners (or the exchange itself) may keep data provided by that partner and, if so, for how long. If participants decide that when one entity leaves an exchange, the other partners must delete the data provided by that entity, the architecture and infrastructure must be able to accommodate this, which leads to important design and development considerations.



Develop Data Breach Response Plan. Unfortunately, in today's world, having a data breach plan is a requirement for ethical data use. As effective breach and adverse event planning is considered standard practice for sensitive data, systems that share individual level data must have detailed plans in place for how they will coordinate responses in such events. The addition of multiple stakeholders through data sharing agreements significantly increases the complexity of such plans and requires careful thought and planning.

Determine Rules for Data Use and Dissemination. Although describing exactly how data can (and cannot) be used will be detailed in the data sharing agreement, questions will also arise about interpretation of data, use of particular elements, and allowable research and evaluation questions. The governance arrangement must have a process for surfacing and addressing these questions as they arise.

A key issue that may arise is in the differing interpretation or use of a particular element. While the governance regime does not need to enforce rigid definitions across all participants, all participating should understand how data are being used. A relatively common example of this is how states and others use quarterly earnings data to report annual wages. There are multiple different ways of calculating this element that could lead to different entities reporting different annual wages for the same individuals. As long as there is transparency in how information is produced and an understanding among all parties as to how differences may arise, it may not be necessary to conform to exact definitions, which can be a significant challenge in large collaborations.

Additionally, attention must be paid to ensuring all parties understand how and when their data are being publicly disseminated. The exact nature of this can range from requiring notification prior to dissemination, to submitting for approval prior to release. In most relationships, all parties will be involved in both sides of this transaction (i.e. publishing reports built on data provided by others, as well as providing data that will be published by others).

Consider Rules for Data Use by External Researchers. The types of administrative data being considered for these types of data sharing relationships can answer important policy and practice questions for the data-owning agencies that participate, but they could potentially be important tools for external researchers, policy research organizations, and others looking to answer key questions about education, training, and employment. A governance regime must establish rules and criteria to evaluate and accept or reject such external requests, recognizing that providing individual-level data—even when deidentified—can require new data sharing agreements among partners and the external researcher.

Define Architectural and Technical Requirements Through Governance Processes and Data Use Needs.

There are many different technical designs and infrastructure arrangements that would support the basic requirement of sharing administrative data between entities. But each design choice may have implications for data use and for the relationships between participants, so the governance arrangement must define the technical design rather than the other way around. Consider the example above where one participant is able to leave the arrangement and have its data deleted from the system. In such a case, the system (and other participants) must be able to remove those data without sacrificing the utility of the infrastructure for the rest of the participants.

The way data flow in a system, including the degree to which they are centralized or federated, is another choice that should be driven by the decisions made by participants and their needs.

Further, the particular data uses envisioned may impact the design and implementation of the process to match two identities from different data sources (a process known as “identity resolution”). Generally speaking, transactional uses such as allocating benefits or intervening with a particular person, require the highest level of accuracy in data matching. Research uses, on the other hand, can tolerate some level of error in matching without greatly affecting overall research findings. The degree to which “fuzzy matches,” where near matches such as “Pat Lane” and “Patrick Lane,” are considered equivalent is an important consideration prior to the development of infrastructure.

With the size of these systems and the limited number of identifying data elements available, it is virtually impossible to inspect each match to determine its accuracy, but participants must understand the potential level of error within a system for both false positives (i.e. determining two people are the same person when they are not) and false negatives (i.e. determining two people are different individuals when they are the same person), and how it affects their envisioned usage.

Address Data Sharing Liability. Governance arrangements must also clearly address the issue of liability in the event of an unauthorized data disclosure. These components should be included in the data sharing agreements between participants. Using the types of sensitive data envisioned in the Data for the American Dream effort must necessarily involve discussions of liability, especially in a data sharing context. When there are multiple states involved in sharing data, the resulting agreement should establish which entities are responsible for liability incurred as the result of adverse events like data breaches. Such provisions will need to identify a process for determining responsibility for the breach, how required actions will be undertaken (such as notification of individuals whose data may have been accessed by unauthorized parties), and who will bear financial responsibility depending on how responsibility is assigned.



CONCLUSION

Formal and legally compliant data sharing agreements are the foundation of any effort to link together datasets from multiple agencies, including those from multiple states. Complying with federal and state statutes is a methodical process of ensuring certain requirements are met, but it is not sufficient for a successful data sharing effort. A strong and effective governance regime that oversees the work is as essential as a data sharing agreement. The structure of the governance arrangements should allow for collaboration of all parties throughout the different components of data sharing, ranging from initial data preparation to their ultimate destruction. The governance arrangements should also be formalized in ways that help ensure the sustainability of the collaboration.

As state agencies and other partners work to develop the necessary data infrastructure to accomplish their project goals and help further the mission of Data for the American Dream, data governance—whether the effort includes cross-state data or not—must be at the forefront of each team's efforts.

ENDNOTES

- 1 See www.wiche.edu/mlde and <https://coleridgeinitiative.org/> for further information on cross-state data sharing initiatives.
- 2 20 U.S.C. § 1232g. FERPA includes other exemptions, including one for entities conducting studies or carrying out research, but that work is focused on work related to predictive testing.
- 3 Privacy Technical Assistance Center, "Written Agreement Checklist." (2015.) Retrieved from https://studentprivacy.ed.gov/sites/default/files/resource_document/file/Written_Agreement_Checklist_0.pdf February 14, 2020.
- 4 U.S. Department of Education, "The Family Educational Rights and Privacy Act: Guidance for Reasonable Methods and Written Agreements," (n.d.) Retrieved from https://www2.ed.gov/policy/gen/guid/fpco/pdf/reasonablemtd_agreement.pdf February 14, 2020.
- 5 Information in the list drawn from U.S. Department of Education, "The Family Educational Rights and Privacy Act: Guidance for Reasonable Methods and Written Agreements," (n.d.)
- 6 Ibid.
- 7 See <https://www.nist.gov/cyberframework> for further information.
- 8 20 U.S.C. § 603.5.
- 9 Ibid.
- 10 See 20 U.S.C. § 603.9 & 603.10 for further information.
- 11 The General Data Protection Regulation (GDPR) adopted by the European Union has important potential ramifications for education data systems. The significant legal questions and nuance are beyond the scope of this paper, but competent legal counsel should be able to ensure compliance without affecting the overall functionality of shared data systems.

D4
AD

WWW.D4AD.COM

