

MidWest Collaborative: Data for Evidence Based Policy

ABSTRACT

The Midwest has a vast network of regional activity throughout its agricultural and manufacturing base, as well as its supply chain, infrastructure and in its agile workforce, combined with hundreds of universities and colleges. Midwestern states also share regional challenges: including the dynamic, changing evolution of jobs within a manufacturing-heavy economy, the effects of trade agreements, an aging population, skilled worker shortages, migration out of rural communities, and the effects of the opioid crisis.

There is a new opportunity to address these challenges and a new focus on evidence based policy at the federal level. A group of universities and education/workforce agencies in the Midwest have worked together to design a system that enables individual states to answer critical questions that are relevant to society. Based on these activities, including training opportunities, policy forums, and conference presentations we believe that there are a large number of states in the region that want to be involved. This white paper shares the lessons we learned over the last three years and charts a course for formalizing our partnership as the MidWest Collaborative: Data for Evidence Based Policy.

OVERVIEW

State governments have the task of solving daunting problems on a multitude of fronts, and they need relevant and timely data to inform their policy decisions and to ensure public policies are implemented effectively. State and local government have invested both technical and human resources in building data systems within their states but there are substantial limitations in characterizing education to employment transitions.

The MidWest Collaborative is an opportunity for the region to build a data system, the associated legal framework, and human resources that will dramatically improve the ability to monitor the health of our region's interconnected education and workforce systems. This capacity is critical to states in the Midwest which have a strong network of regional activity throughout its agricultural and manufacturing base, as well as its supply chain, infrastructure and in its agile workforce, combined with hundreds of universities and colleges. Midwestern states also share regional challenges: including the dynamic, changing evolution of jobs within a manufacturing-heavy economy, the effects of trade agreements, an aging population, skilled worker shortages, migration out of rural communities, and the effects of the opioid crisis

We need systems that have the technical resources with modern data systems and the underlying computer capacity to mine large amounts of administrative data. The improvements in the technical capacity of Cloud data systems and the wider adoption of the Federal Risk and Authorization Management Program (FedRAMP) framework offer an enormous opportunity to integrate de-identified data resources from individual states to answer critical questions about the role of education and training in the Midwest and throughout the nation. These technical systems are improved by the use of open source and secure programming resources such as Python and Jupyter Notebooks that improve the capacity of administrative data systems to make detailed statistical studies of integrated data. Moreover, the use of the "Five Safes" framework (safe data,

safe projects, safe people, safe settings, and safe outputs), provides a deliberate methodology to use confidential microdata in ways that comply with federal and state law.

As this regional effort advances, connecting data across states, state and local agencies will need training in advanced analytics crosses state lines as well. Moreover, because government services have become so integrated, staff from multiple state agencies must participate collaborative in any efforts to improve the professional development of data science staff. In fact, much of what we know about the importance of data science and the role of Technical Training as a discipline reinforces this point of view. Staff from the agencies with direct responsibility for data science as well as the senior leadership of state and local government must work together to improve the skills and capabilities of all agency personnel to enable states to improve work.

The Opportunity

There is an opportunity in 2019 to link states together in ways that will improve the use of administrative data in the Midwest and serve as model for the entire nation. The Midwest region has been working hard to build a network of like-minded data science professionals and organizations that can partner together to make better use of data around education and workforce transitions. There have been meetings in Columbus, OH, Indianapolis, IN, Jefferson City Missouri, and Chicago, IL designed to brainstorm about the needs of these new systems. In addition, many staff from states in the region have participated in collaborative training programs on data science run through the Coleridge Initiative in Indiana, Illinois, Missouri, and Ohio.

A second reason there is a strong opportunity is that many states have already worked with entities like the Coleridge Initiative to test systems that link data together. Ohio, Indiana, Missouri, and Illinois have already conducted a number of studies using the ADRF platform on a diverse range of topics.

In Illinois, the Department of Corrections partnered with the Department of Employment Security to study the effects of industrial training. Another team of researchers proposed and answered questions about the employment of formerly incarcerated Illinois individuals, looking specifically at the employment across the border in Missouri. The Ohio offering of the Coleridge Initiative training integrated data from Ohio on housing, education, vocational rehabilitation, unemployment insurance, and higher education with data from Illinois and Missouri. Teams of researchers from around the region were able to propose and conduct studies that answered pressing policy questions. For example, one team worked to study the employment of higher education leavers outside of Ohio, thereby beginning to answer one of the most pressing questions that Governors and the Legislature always have, “How many of our college graduates are working outside of Ohio?” The Indiana and Missouri projects covered a diverse slate of topics involving university graduate migration, offender recidivism, community college graduation challenges, and the post-incarceration impact of vocational training. All project teams conducted exploratory analysis in their efforts to properly frame their statistical and machine learning-based modeling.

These opportunities are the result of hard work on the part of the state governments as well as the Coleridge partner institutions, which include a number of Midwestern state universities. Along with other universities such as Indiana University and The Ohio State University, these pioneers

in the states have trained over 350 staff from state and local government as well as select federal agencies. These participants are a powerful example of the opportunity that lies ahead, by improving the skills in these individuals the Coleridge Initiative and its partner institutions and agencies have laid the groundwork for major gains in improving public administration.

Targeted Investments

Philanthropic foundations, such as Schmidt Futures, the Overdeck Family Foundation, the Ewing Marion Kauffman Foundation and the Bill and Melinda Gates Foundation, have provided seed funding to support the initial development in key states. Securing these initial gains requires additional investments by government as well as funders. There are a number of opportunities to improve practice immediately that require direct support. These include the following investments:

Create a Collaborative Data Governance Structure. There is a strong basis for building a data collaborative in the states. Initially, state governance focuses on individual states, but with the addition of new states we require more deliberate procedures for data administration, data use, and privacy and security. While there is a strong tradition of individual state governance in all of the states, it will be essential to ensure that state systems can be mirrored in a manner consistent with state requirements to ensure rapid deployment of a single MidWest Collaborative data system.

Develop a core data model. The ADRF data model requires collaborative development. The work done to date, for the advanced analytics training courses laid a foundation for data systems development, but the states must pursue a common data model to ensure that the cost to states for joining the Midwest Collaborative remain low and barriers to entry are minimized.

Build Production Level Technical Capacity. The capacity to integrate state data exists in the ADRF as a FedRAMP compliant data system. However, the system needs to scale up to ensure that other states can join. Moreover, to make sure that the systems maintain privacy and confidentiality rules the system requires attention to maintenance of the legal frameworks.

Establish a Professional Development Curriculum. Coleridge has a powerful model for training. To build on this the training needs to evolve, developing modules to enable continuous improvement of individual skills, and offering refresher courses. The system uses a powerful curriculum model based on the Jupyter Notebooks that can expand flexibly to integrate a range of additional content.

Next Steps

Much work can be done in parallel if advisory groups are put in place quickly.

1. Build a common curriculum The success of the Coleridge Initiative is at least partly due to the momentum that has been established. The member entities should come together and develop a curriculum model for education and workforce transitions that can be adopted by participating universities

Timeline: 6 months

2. Ensure Technical Scalability. The collaborative needs to ensure that the ADRF is ready to scale in a manner that is consistent with their respective needs. The member entities should establish an advisory group that can guide the development of the ADRF.

Timeline: Immediate

3. Develop a Formal Data Model. The ADRF and Coleridge Initiative have a lot of experience with data from each state. But, there needs to be a concerted effort to build a formal data model that can be used by each state when it sends data to the MidWest Collaborative. There are existing models, such as the CEDS (common education data standards) but there needs to be a more flexible standard that can accommodate expanded data from the Workforce as well as education. The member entities should form an advisory group that can inform the development of the model.

Timeline: Immediate

4. Establish Governance Structure. The member entities need to formalize a structure for the MidWest Collaborative. Does it exist as a centralized intergovernmental body? Can it be sustained as a project of a collection of states? There are non-profits that aggregate the needs of states, such as the National Association of State Workforce Agencies, that might be an administrative home for this kind of organization. However, there are also options for creating a stand-alone non-profit or locating the Collaborative as part of a university research center.

Timeline: 6 months to a year

5. Establish Data Governance. Governance over data is a critical topic that must be developed early on in a project. There are typically three levels of governance in data systems, 1) the high level “Directors” or “Governors” level that sets the agenda, allocates resources, and monitors an executive team; 2) an “Executive Team” that includes data science leads for state government in education and workforce that supervises the day to day work, and 3) a “Data Management Team” that handles oversight of the technical aspects of a data system. There is also a need for an arms-length security team that can independently assess and ensure continued privacy and data security. In the case of a FedRAMP or NIST Moderate Baseline compliant system these staffing and governance questions are explicitly laid out. Collaborative approaches can be developed that ensure that states do not lose the right of governance over their state’s data, as evidenced by the 20-year functioning of the LEHD program.

Timeline: 3-6 months