Secure Shouldn't Mean Secret: A Call for Public Policy Schools to Share, Support, and Teach Data Stewardship

Maggie Reeves and Robert McMillan
Georgia Policy Labs, Andrew Young School of Policy Studies, Georgia State University

September 12, 2019

The authors thank Julia Lane, the anonymous reviewers, Alex Hathaway, and Sally Wallace for their thoughtful feedback on earlier drafts.

The public has long benefitted from researchers using individual-level administrative data (microdata) to answer questions on a gamut of issues related to the efficiency, effectiveness, and causality of programs and policies. However, these benefits have not been pervasive because few researchers have had access to microdata, and their tools, security practices, and technology have rarely been shared.[1] With a clear push to expand access to microdata for purposes of rigorous analysis (Abraham et al., 2017; ADRF Network Working Group Participants, 2018), public policy schools must grapple with imperfect options and decide how to support secure data facilities for their faculty and students. They also must take the lead to educate students as data stewards who can navigate the challenges of microdata access for public policy research.

This white paper outlines the essential components of any secure facility, the pros and cons of four types of secure microdata facilities used for public policy research, the benefits of sharing tools and resources, and the importance of training. It closes with a call on public policy schools to include data stewardship as part of the standard curriculum.

Framing

The benefits of sharing and linking agencies' microdata for applied policy research are clear. Conducting research with this detailed data allows government agencies, policymakers, and the research community to better understand the long-term impacts of programs, the consequences of policy changes, and the inputs that lead to success across a range of policy issues—without imposing additional data collection burdens on residents (Abraham et al., 2017; Culhane, Fantuzzo, Hill, & Burnett, 2018; Desai, Ritchie, and Welpton, 2016; Goerge, 2018; Lane, 2018; Liebman, 2018; United Nations Economic Commission for Europe, 2007).[2] Additionally, researchers and students in public policy schools aim to conduct research that matters, and the potential practical benefits from micro-level data analysis align well with this goal.

Despite the interest in data sharing, concerns about data use, security, and privacy (Goerge, 2018; Lane, 2018) often pause or end data sharing discussions. Many motivated researchers and government

[1] See the Commission on Evidence-Based Policymaking's 2017 report for a full history of these efforts.

[2] It is beyond the scope of this paper to do justice to the benefits of conducting this type of research and the risks and barriers that stakeholders face when sharing data. Please see the references throughout for a more thorough discussion.

agencies have spent inordinate amounts of time and money navigating these barriers. They have solved the legal, relational, and technological challenges by developing complex data sharing agreements, creating new technology, and building trust through pilot projects. These solutions often assuage government agencies' reluctance to share sensitive data (Foster, 2018; Lane, 2018), but they have been costly at the organizational and personal levels. For those partnerships that are successful, the benefits to the government agencies and communities they serve have been great. To date, these successes are isolated examples because solutions have not been widely shared or recycled.

We believe public policy schools can elevate and share best practices, models, and open-source tools for secure access to microdata.[3] These efforts are crucial for applied public policy research and the public good because these schools train the next generation of government leaders and researchers. However, there is no perfect method to navigate the intersection of data access and security for public policy research; each data partnership must navigate its unique circumstances. With this in mind, we now highlight the principles and core components that secure facilities must manage. We also outline the necessary components of data stewardship training for public policy students.

Successful Components

In outlining the features of a secure facility for microdata, we propose that the elements must attend to several principles: prioritize the public good (Abraham et al., 2017); balance protection of sensitive information with accessibility for researchers, increasing both privacy and the evidence base (Abraham et al., 2017); acknowledge solutions must have the flexibility needed to keep up with the pace of changing technology (Goerge, 2018); prepare the next generation of stakeholders to sustain solutions (Lane, 2018); and utilize tools and solutions that are replicable but not at the expense of meeting local demands (Lane, 2018). With these principles in mind, we add to the "five safes" model recently discussed in Desai, Richie, and Welpton's 2016 article to describe key components of microdata access and use—projects, settings, data, people, and outputs—and we briefly note commonly used tools. When public policy schools and researchers are utilizing microdata, we believe they should use these tools to ensure proper data stewardship.

Safe projects

This component considers the legal, moral, and ethical uses of data to ensure they are appropriate and further the public good. To this component, we add transparency to ensure the public can have the knowledge of and a voice in its usage (United Nations Economic Commission for Europe, 2007).

Tool	Purpose
Institutional Review Board approval of projects	Ensure beneficence, autonomy, and justice (Culhane et al., 2018)
Data sharing agreements (Culhane et al., 2018; Goerge, 2018) and specific project sign-off	Cover legal bases
Co-development of projects between researchers and government (Booker, Conway, and Schwartz, 2019)	Build trust and interest

[3] Several networks, such as the Actionable Intelligence for Social Policy, policy labs, NYU's Administrative Data Research Facility, and the Commission on Evidence-Based Policymaking, are trying to share best practices and tools and communicate lessons learned to help initiatives be successful, more prevalent, and sustainable.

Project selection criteria	Ensure mutual benefit and potential policy impact
Pre-analysis plans	Ensure transparency
Regular project meetings	Ensure government agencies are not surprised by findings

Safe settings

This component focuses on adequate controls on data access to limit unauthorized use. We add that the environment should also be accessible (Foster, 2018; United Nations Economic Commission for Europe, 2007).[4,5] A range of models exists, including restricted access facilities such as Federal Statistic Data Centers, central databases to share de-identified data on a project-by-project basis (Texas Education Research Center), individual researcher access on a single machine, and matching services (the Commission on Evidence-Based Policymaking's proposed National Secure Data Service). We highlight the pros and cons of four of these in the following section.

Tools	Purpose
Cloud instance or air-gapped machines	Separate dedicated networks or physically isolated machines with no network connectivity
Limited physical access and secure data rooms	Keep out untrained and unauthorized personnel
Separation of duties	Multi-party verification of information brought into or taken out of a safe environment
Hardened machines such as removal of internet access, geo-fencing of IPs, closed ports, and limited software installation	Prevent the entrance of malware, prohibit the transfer of files into or out of a machine, and simplified security administration
Multi-factor identification and user-based privileges	Add verification integrity and limit exposure to content by implementing named accounts
Firewalls	Threat monitoring and prevention for networked assets

Safe data

This component tries to remove the potential for the identification of individuals by reducing disclosure risks. Because the value of data has increased and hacking is now a daily norm, security must be increased to reduce the threat of harm due to data loss or inappropriate access. This component is changing rapidly as technology tools evolve.

Tools	Purpose
De-identification	Remove personally identifiable information
Encryption	Secure sensitive data elements needed for future research
Hashing	One-way encryption of sensitive data that is stored as a hashed value; typically used on SSNs or government IDs

[4] By accessible, we mean that the setting should not pose barriers so high that the obstacles outweigh the good intentions of researchers to use the data. For example, these access barriers could be geographical, time-based, or cost-prohibitive.

[5] Of course, the time and effort to establish data-sharing relationships are costly, but the additional costs of building secure data research facilities may not be scalable without significant investments.

Synthetic data sets	Artificially created data, usually through algorithms, used to test models without exposing sensitive data
Sample data	Limit exposure by limiting the amount of data available
Authorized users	Provide a gate to prevent non-trained personnel from having access to sensitive information

Safe people
This component ensures the people involved with microdata have the knowledge, skills, and incentives to use the data appropriately. We add that this component should also focus on training the next generation of data stewards.

Tools	Purpose
Background checks	Provide insight into a person's activities that are not available through regular hiring conversations
Signed security and use agreements	Ensure acknowledgment of risk and proper use; creates a liability if broken, which deters improper behavior
Active and passive training on security (physical, cyber), IRB, HIPPA, FERPA, and other legal frameworks or data sharing agreements as applicable	Explain the handling procedures that apply to specific datasets where some datasets may require extra precautions

Safe outputs
The final component is reducing the residual privacy risk in final outputs and publications. While many of these methods have traditionally been used to remove privacy risks, current technology and access to publicly-available data from private and government sources have increased the threat of re-identification. We add that these outputs should also contribute to the public good.

Tools	Purpose
Restrictions on data transfer through a separation of duties	Require collusion to extract sensitive data; improve organizational integrity
Cell size rules	Limit the ability to identify individuals
Review of outputs manually and algorithmically	Help ensure that non-sensitive data does not leave the environment; algorithmic review speeds the process
Audit logs of outputs, access, and data transfer	Assist in a forensic investigation in the event of a breach; provide a tool to assess abnormal behavior; can be fed into alert systems to notify administrators in the event of nefarious behavior

Safe Settings Models

Secure administrative research facilities take many shapes, and all navigate the challenges to use microdata for policy research. There is consensus on the importance of solving issues of trust, legality, security, financial sustainability, and reciprocity across these models, but there is no uniformity on which setup to use. We now present the pros and cons related to four microdata access models that are successfully used to conduct applied and academic policy research within a public policy school. This discussion is by no means exhaustive.[6] This discussion is intended to help public policy school deans consider ways they can strategically support their faculty and students in utilizing, building, or maintaining secure microdata facilities.

Policy Lab

Policy labs are partnerships among academics and government agencies where microdata is used to answer and solve pressing policy problems. These research practice partnerships established trust and mutually beneficial relationships, determined governance models, signed data sharing agreements, and built infrastructure to house data that is readily linkable. Many policy labs have not developed intending to be a secure microdata facility but, instead, have focused on being the research and thought partner for government agencies. Nonetheless, in the process of gaining trust and holding data for research purposes, most have navigated the same challenges related to data sharing and use.

As outlined in the ADFR Network's 2018 article, many data intermediaries follow a four-state maturation process: "1) establish data supply and research demand, 2) establish ad hoc policies and procedure for limited access and use, 3) support a virtual or physical location where data are managed or used for analysis, 4) establish core functions for scaling up data access and use." Some labs, such as the Georgia Policy Labs in Georgia State's Andrew Young School of Policy Studies, have gone through the first three stages in support of their mission but are in the process of scaling.

Opportunities. Policy labs present incredible opportunities for faculty and policy students to conduct applied policy research that benefits the public good. These labs have invested in relationships and solutions to many of the challenges that often thwart individual researchers from working with sensitive microdata. Once established, long-term data-sharing agreements with project-specific sign-off forms allow researchers affiliated with policy labs to move quickly when a project is initiated. Other tools employed by policy labs such as pre-analysis plans and ongoing communication protocols ensure that the research is focused on solving public policy problems—benefitting the public good. Policy labs use different staffing models, but all prioritize the relationships and communications with government partners. This prioritization supports co-learning environments, where both the academic researchers and government partners are learning from one another. These environments also have internal (staff/faculty) and external checks (government partners) on safe outputs before publication.[7]

Policy labs also afford significant training opportunities for students, faculty, and government partners. For example, in Georgia, the policy lab actively trains the students, faculty affiliates, and all staff on the security, privacy, and output requirements for all projects in addition to robust discussions about data linking, deidentification, and data cleaning. This intentional training in data stewardship serves to

[6] We have not found a comprehensive list of integrated data system efforts. Conversations with colleagues have established that counting and classifying these efforts is a moving target, especially as partnerships grow, morph, and disintegrate over time. Cataloging facilities is also a challenge, as they may be housed within universities, government agencies, or non-profits. The Federal Commission on Evidence-Based Policymaking has recommended a data-sharing service rather than a facility as the solution to linking and sharing federal microdata. We are supportive of this initiative but also aware of its limitations, especially for state and local microdata.

[7] The review by the data provider provides a distinct advantage, as data providers are inherently protective of the privacy of their constituents.

educate policy researchers on the “backend” of the data process. It illuminates the challenges, risks, and opportunities of this research, which helps all involved to be better data stewards. Additionally, some labs, including the Georgia Policy Labs, focus on upskilling government partners on analytical tools and methodology to enhance the productivity of all.

Challenges. Investment in a policy lab is not without its costs. Most policy labs or similar centers have received multimillion-dollar philanthropic investments to start up. These investments have supported the time it takes to build and begin impactful research, often taking 18 months to establish the team, data sharing agreements, and data infrastructure. Also, because policy labs often create integrated data systems as a tool rather than as a core function, they do not always have the expertise in-house to build an integrated data system from scratch.[8] A challenge for policy labs beyond initial investment costs is financial sustainability. Not established as consulting shops, these labs must balance the effort needed to conduct research in support of partners’ priorities while also navigating support from funders.

While there is a significant opportunity for policy labs to serve as secure microdata facilities, lab leadership must decide how to balance the opportunity for increased research productivity to support the public good with the risks of opening access to more researchers. These risks can be related to data security and access—for example, whether labs share data outside of their facilities—and trust, as many academics have not been trained to prioritize the relational aspects of working with government partners.

Individual Access

The most common form of microdata access within public policy schools is a one-on-one relationship between a faculty member and a government agency. These relationships may be based on a long history between the two or indicate that the researcher has worked through the pre-approval process to gain access to data. These relationships span from very casual to extremely formal and are accompanied by a similar range of guidance on the setting. For example, a faculty member accessing a federal data set may be given specific information technology guidance to ensure their computer, network, and access are configured to protect the data; this guidance may be paired with regular field audits to ensure compliance. On the other side of the spectrum, a data-sharing agreement may be signed with vague specifications and have few requirements for regular checks on access.

Opportunities. The dominant benefit for this setting is that individual researchers can work directly with microdata with few restrictions on location, inconveniences, or high costs. Because researchers are often working with an agency directly, the researcher can match security protocols to that agency’s needs. This type of research supports work that benefits the public good as well as furthers academics' research interests. For those researchers who invest the time, these relationships can establish trust that often lasts for many years, which allows for a more nuanced understanding of an agency's data. With some exceptions, the individual setting is often low-cost to set up.

Challenges. The challenge of this arrangement is that that the researcher is in many ways solely responsible for the privacy, security, and uses of the data; yet, that researcher is still within a larger university setting. These relationships can often move quickly, which is a pro, but do not always gain the

[8] A caveat to this is that most policy labs are housed within universities. This home can provide many resources in information technology in some settings or be a hindrance to innovation in others.

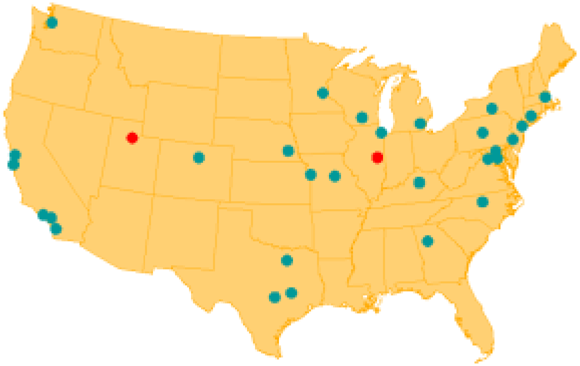
attention of the university's security and compliance offices. As such, within a unit or school, there can be a range of unknown protocols, information technology requirements, and risks. The direct data-sharing relationship is often the least costly to set up but also has many potential hazards.

Setting up these relationships and security systems can be time intensive for the researcher, especially if she is doing it on her own. For those relationships where the government entity does not provide much guidance on the security or privacy requirements, researchers can face a steep learning curve. Once data is received, there can be some limitations, such as not being able to share data with research co-investigators or assistants. Finally, this type of secure setting does not lend itself to any scaling or shared access, as protection and individual relationships are prioritized.

Federal Statistical Research Data Facilities

There are currently 29 Federal Statistical Research Data Facilities (RDCs) across the country that provide secure access to federal agency microdata (see Figure 1 for locations[9]). These facilities are housed within government, nonprofit research, and university settings. For RDC member universities,[10] approved faculty and students can access these data at no cost. For researchers at non-member institutions, there is a high cost.

Figure 1. RDC Locations – September 2019



Source: www.census.gov/fsrdc

Opportunities. The main advantage of using RDCs is that researchers can use individual-level microdata without having to learn the intricacies of data security or find significant funding (for researchers at member institutions). The RCDs manage the access, physical rules, data infrastructure, and output reviews so that universities do not have to do that in-house. This type of controlled environment that includes in-depth background checks, output checks by multiple parties, and locked-down physical locations is ideal for protecting privacy and security. These facilities allow specific data access along with a suite of analytical tools to conduct research within the RDC. Outputs and results are emailed to researchers after a series of checks to confirm safe outputs.

Challenges. Despite this ideal setting for security and outputs, RDCs are limited in many ways. First, they only house limited federal data (see census.gov/fsrdc for available datasets). Secondly, as

[9] Blue dots indicate existing facilities; red dots are locations opening in 2019.
[10] See www.census.gov/about/adrm/fsrdc/about/partners.html for a list of current member institutions.

seen in Figure 1, they are not convenient for a large proportion of researchers across the country, as you must access the facility in-person. While it is not uncommon for some researchers to drive many hours each way to visit the facility, that approach is not tenable for many. Start-up time to gain approval to access the data, as well as wait times for output reviews for disclosure avoidance, can be challenging. As mentioned above, for researchers at non-member institutions, there is also a hefty fee to be able to utilize the facility.

State Longitudinal Data Systems

Most states have at least one state-based integrated data system, although the purposes and scope of each differ significantly (www.ecs.org/state-longitudinal-data-systems). Supported by federal grants, many states have built databases such as State Longitudinal Data Systems focused on early learning through workforce outcomes. These data systems often focus on public reporting, interagency data alignment, and key metrics for public policy. Some also share de-identified microdata with researchers on a project-by-project basis.

Opportunities. While these systems are costly, the costs generally do not fall on researchers to receive the data. The data from the more sophisticated systems, such as Georgia's, is cleaned, organized, linked, and spans many years. For approved researchers, these data and associated codebooks allow for research to be conducted quickly once data are received. Dependent on the governance models, systems may require agency approval for each dataset provided, which ensures each agency understands the project.

Challenges. Because these data systems and governance models are not built purely for research, significant challenges in gaining access to the data may arise. These challenges can be based on wait time for approvals, fit between research questions and state priorities, and limitations on researcher qualifications (such as in-state status or no dissertation research). The systems tend to have precise data security requirements for storing de-identified microdata, but researchers may still be unfamiliar with the terminology or mandated procedures. In that way, there are many similarities to individual data-sharing agreements. Because these systems have already linked data from multiple sources, deidentified data can be used efficiently but will not work for all projects. Receiving data from a statewide integrated data system also places the burden of secure outputs on the researcher, as the data provider may not require review or think about risks of re-identification.

Summary

No matter the setting, the costs in dollars and time to build these systems or learn how to secure them are high. When the government bears the cost for building and maintaining the facility or access, researchers are limited by the datasets provided. When researchers must learn how to develop or maintain secure data facilities, the investment is significant. It is also easy to fall behind ever-advancing technology, which poses risks for security and privacy. Furthermore, no network exists to circulate solutions to the "five safes" systematically. In effect, governments and researchers must invest substantial resources to address these concerns on their own.

Sharing Resources

We believe there is an opportunity for coordinated sharing of best practices and tools among researchers using microdata to improve public policy. Several organizations are moving toward open-source sharing of solutions and technical assistance, but we believe there is still significant room for strategic sharing of practical tools, how-to guides, and checklists. This sharing should invite constructive criticism, redesign, and innovation to improve microdata facilities. It should also discourage high-cost and proprietary solutions. We believe the NASPAA network may be one way to share innovations used by public policy schools.

In addition, some organizations already pushing out innovations, resources, and training to fill this gap include:

- Actionable Intelligence for Social Policy: AISP works with state and local government entities to develop and improve integrated data systems. (www.aisp.upenn.edu)
- Administrative Data Research Facilities Network: ADRFN works to connect researchers, data holders, and intermediaries to improve access to and ethical use of administrative data. (www.adrf.upenn.edu)
- Administrative Data Research Facility and Coleridge Initiative: ADRF is a secure platform for microdata, and the associated Coleridge Initiative builds capacity to conduct research using those data with a project-based approach.
- Policy Labs: Policy labs are collaborations among government agencies and academic researchers to analyze, test, and improve the effectiveness of public policies and programs through long-term, co-learning relationships and rigorous research using microdata. Some provide training for government agencies, and all are committed to open science. (www.arnoldventures.org/work/policy-labs)
- Commission on Evidence-Based Policymaking: This federal initiative was launched to increase the availability and use of data for policymaking while protecting privacy and confidentiality. (cep.gov)
- California's Health and Human Service Agency Data Playbook: This resource provides resources for state government employees to better access and use microdata. (chhsdata.github.io/dataplaybook/resource_library)

Importance of Training

There are high costs for any researcher who is learning how to be a good microdata steward, especially for those building microdata facilities or starting relationships with government entities.[11] We believe public policy schools must also train students how to be competent data stewards. This training should include the principles, best practices, and vocabulary needed to use microdata for good. For scale, it can be offered as a class in a data science curriculum or through hands-on training for schools that have policy labs or other microdata facilities.

Currently, as students are trained in public policy, economics, and other applied policy fields, they come to understand the importance of data and how to analyze it. However, typically, the management of that data is passed over. Post-graduation, this education is borne out of necessity or grit to do applied

[11] In addition to data stewardship training, we encourage advanced analytical training both in traditional methods as well as new data science methods for students and government employees. For a thorough review of the curricula needed and an applied problem-solving approach, see a forthcoming article by Kreuter, Ghani, and Lane entitled, "Change through Data: A Public Extension Program for Government Employees."

policy research, and graduates must confront the barriers of building trust, privacy, security, and governing frameworks.

The benefits of microdata research for evidence-based policy have yet to be fully realized. Microdata and other big data are created at increasingly rapid rates, and computing power no longer stands as a limitation. Now, public policy schools must ensure that students emerge with the analytical skills and data stewardship knowhow to gain access to this data and responsibly use it for good. We suggest that, at a minimum, students graduate with an understanding of:

- Common issues preventing data sharing and best practices about the five safes
- Pros and cons of various data access settings
- Tools and methods used to protect the security and privacy of microdata
- A basic information technology lexicon
- Skills, resources, and staffing models required to support microdata facilities, policy labs, or individual research

This training is crucial for students who will likely serve as public policy researchers in the future. Whether working as a faculty member, consultant, government employee, or policymaker, our students will inevitably need to balance the constraints of data security and access. Public policy schools need to train them so that microdata research is more efficient and safer and can contribute to evidence-based policy.

Call to Action

We call on NASPAA and its members to support evidence-based policymaking by sharing the tools and standards needed for successful microdata access and use.[12]

- NASPAA should create a library of open-source resources and tools for creating and maintaining a secure microdata facility. Schools should actively contribute to this shared resource.
- Public policy schools should create a course that is widely offered on data stewardship.

[12] The Georgia Policy Labs is happy to share its security protocols, data infrastructure design, and training/onboarding checklists with interested parties. Email the authors for more information.

References

Abraham, K. G., Haskins, R., Glied, S., Groves, R. M., Hahn, R., Hoynes, H., Liebman, J., Meyer, B. D., Ohm, P., Potok, N., Rice Mosier, K., Shea, R. J., Sweeney, L., Troske, K. R., & Wallin, K. R. (2017). *The promise of evidence-based policymaking*. Commission on Evidence-Based Policymaking. Retrieved from www.cep.gov/report/cep-final-report.pdf

ADRF Network Working Group Participants. (2018, June). *Data sharing governance and management* (Rep.). Retrieved from repository.upenn.edu/admindata_reports/2

Booker, L., Conaway, C., & Schwartz, N. (2019, May). *Five ways RPPs can fail and how to avoid them: Applying conceptual frameworks to improve RPPs*. Retrieved wtgrantfoundation.org/library/uploads/2019/06/Five-Ways-RPPs-Can-Fail.pdf

Callahan, M. E. (2012, March). *Handbook for safeguarding sensitive personally identifiable information* (United States, Department of Homeland Security, The Privacy Office). Retrieved from www.dhs.gov/sites/default/files/publications/handbookforsafeguardingsensitivePII_march_2012_webversion_0.pdf

Culhane, D., Fantuzzo, J., Hill, M., & Burnett, T. (2018). Maximizing the use of integrated data systems: Understanding the challenges and advancing solutions. *The Annals of the American Academy of Political and Social Science*, *675*(1), 221–239. doi.org/10.1177/0002716217743441

Desai, T., Ritchie, F., & Welpton, R. (2016). *Five safes: Designing data access for research* (Economics Working Paper Series, Working paper No. 1601). Bristol, England: University of the West of England. doi:10.13140/RG.2.1.3661.1604

Foster, I. (2018). Research infrastructure for the safe analysis of sensitive data. *The Annals of the American Academy of Political and Social Science*, *675*(1), 102–120. doi.org/10.1177/0002716217742610

Goerge, R. M. (2018). Barriers to accessing state data and approaches to addressing them. *The Annals of the American Academy of Political and Social Science*, *675*(1), 122–137. doi.org/10.1177/0002716217741257

Hafner, H.-P., Ritchie, F., & Lenz, R. (2015). *User-focused threat identification for anonymised microdata*. *User-focused threat identification for anonymised microdata*. University of the West of England. Retrieved from www2.uwe.ac.uk/faculties/BBS/BUS/Research/Economics Papers 2015/1503.pdf

Lane, J. (2018). Building an infrastructure to support the use of government administrative data for program performance and social science research. *The Annals of the American Academy of Political and Social Science*, *675*(1), 240–252. doi.org/10.1177/0002716217746652

Liebman, J. B. (2018). Using data to more rapidly address difficult U.S. social problems. *The Annals of the American Academy of Political and Social Science*, *675*(1), 166–181. doi.org/10.1177/0002716217745812

Singh, P., & Kaur, K. (2015). Database security using encryption. In *2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE)*. doi:10.1109/ABLAZE.2015.7155019

United Nations Economic Commission for Europe. (2007). *Managing statistical confidentiality & microdata access: Principles and guidelines of good practice* (Rep.). Retrieved from www.unece.org/fileadmin/DAM/stats/publications/Managing.statistical.confidentiality.and.microdata.access.pdf

About the Authors

Maggie Reeves is the founding senior director of the Georgia Policy Labs, where she leads the strategy, state partnerships, and operations of the organization. She also participates on the school's Policy in the Digital World Task Force. She previously served in the Andrew Young School of Policy Studies as the interim director and associate director for the Center for State and Local Finance. Prior to her roles at Georgia State, she was a policy and program analyst and a communications specialist for the Georgia Administrative Office of the Courts. Here, she created an online, unified statewide system to licensing court professionals and managed several court improvement pilot projects. She also acted as the spokesperson and policy analyst for the Georgia Commission on Family Violence—with a range of policy analysis, research, and communications duties. Maggie received her master's degree in public administration from the Andrew Young School of Policy Studies at Georgia State University and her bachelor's degree in women's studies from Emory University. She is a proud alumna of Georgia State's Executive Leadership Academy for Women.

Robert McMillan serves as the Director of Data Integrity. He leads a team of data scientists who are responsible for curating data for research opportunities. He provides technical guidance for the collection and storage of information from our local and state partners. He has over 25 years of experience with enterprise system design, integrations, and operation. Twenty of these years have been at the director level or as a principal consultant focusing on delivery of large complex new builds, legacy migrations and more recently with cloud solutions. He was a senior policy analyst at the Georgia Governor's Office of Planning and Budget. He previously held director-level positions at the State Data and Research Center at Georgia Tech and the Department of Early Care and Learning. As a principal consultant in the private sector, he has led multiple state agencies through modernization efforts as well as having done work with the Department of Defense to coordinate system-of-systems requirements management for U.S. Army training and simulation solutions.

About the Georgia Policy Labs

The Georgia Policy Labs (GPL) is a collaboration between Georgia State University and a variety of government agencies to promote evidence-based policy development and implementation. Housed in the Andrew Young School of Policy Studies, GPL works to create an environment where policymakers have the information and tools available to improve the effectiveness of existing government policies and programs, try out new ideas for addressing pressing issues, and decide what new initiatives are promising enough to scale up. The ultimate goal is to help government entities more effectively use scarce resources and make a positive difference in the lives of children and families. GPL contains three focus areas: The Metro Atlanta Policy Lab for Education (MAPLE) works to improve K-12 educational outcomes in metro Atlanta; the Career and Technical Education Exchange (CTEx) focuses on high-school-based career and technical education in multiple U.S. states; and the Child and Family Policy Lab looks at issues of the whole child and whole family with Georgia's state agencies. In addition to conducting evidence-based policy research, GPL serves as a teaching and learning resource for state officials and policymakers, students, and other constituents. See more at gpl.gsu.edu.