FIVE SAFES

Protecting Individual Privacy When Sharing Postsecondary Student Data

Data sharing across government agencies can allow consumers, policymakers, practitioners, and researchers to answer pressing questions about how postsecondary students access and succeed in institutions and programs, how they fare after college, and whether outcomes differ between student groups. To be effective, a streamlined and secure data infrastructure must address legal, privacy, technical, and perception challenges.

The "Five Safes" framework seeks to address the challenges of data sharing by providing a simple, practical, and valuable approach to managing data access and use.[1]

SAFE PROJECTS Safe data-sharing projects include governance protocols to control project requests, review, and approval processes, and may require institutional board or ethics board review and approval.

SAFE PEOPLE Data users should be screened and trained to become "safe people." For example, obtaining access to a data system may require background checks, fingerprinting, and proof of research competence, as well as training to understand and follow data governance protocols.

SAFE SETTINGS Ensuring data sharing is done in safe settings requires regulation of data inputs, computation, and outputs. The most important control factors involve the data user's interface and the equipment that is handling the data.

SAFE DATA Hand-in-hand with "safe settings," data users[2] must create "safe data." Practices impose restrictions on what data an analyst can access, what an analyst can do with the data, and the analyst's computing environment, including physical location.

SAFE OUTPUTS Ensuring individuals are not re-identified through the results in the data outputs requires statistical disclosure limitation methods such as rounding, aggregating, and suppressing results to obscure unique observations in tables, figures, or maps.

THIS SUMMARY DRAWS ON FINDINGS FROM "POSTSECONDARY DATA INFRASTRUCTURE: WHAT IS POSSIBLE TODAY" BY AMY O'HARA.

IHEP

The Institute for Higher Education Policy (IHEP) is a nonpartisan, nonprofit organization committed to promoting access to and success in higher education for all students. Based in Washington, D.C., IHEP develops innovative policy- and practice-oriented research to guide policymakers and education leaders, who develop high-impact policies that will address our nation's most pressing education challenges.

Data Sharing Takeaways for Postsecondary Education

Practices and tools exist today to manage postsecondary data access, analysis, and analytic results securely and responsibly. Agencies that act as intermediaries, such as the National Center for Education Statistics (NCES), are necessary to establish and implement a more robust postsecondary data infrastructure and to share data across a number of stakeholders—including students, institutions, states, and government agencies. When building cross-agency linkages, the "Five Safes" framework and insights from other sectors can enable NCES to ensure secure, responsible data access and use, on top of current practices.

The "Five Safes" framework also has implications for postsecondary data legislation. For example, under the College Transparency Act, the Commissioner of NCES would coordinate data sharing agreements with other federal agencies to securely match student information across systems to answer critical questions about student success.

Agencies Should Consider These Recommendations:

SUPPORT SAFETY AND SECURITY THROUGH DATA INTERMEDIARIES

Postsecondary institutions and the federal government should facilitate use of data in secure ways through the use of data intermediaries, like NCES.

Across industries, intermediaries support research access through data standardization, linkages, and secure data hosting. University-based and non-profit research and data intermediaries process large volumes of data, including confidential student-level data, and can address complex governance and security issues. Such examples prove that data access challenges are surmountable and can be addressed in higher education. The Kilts Center at the University of Chicago hosts research on marketing data including Nielsen consumer panel and scanner data. Only subscribers at qualifying institutions can download data, subject to data security provisions, and the data may only be used for research.

SECURELY LINK AND ANALYZE DATA

Government agencies should implement linkages and protocols to make postsecondary education data more available for productive analyses.

Government agencies already use a variety of structures to securely link and analyze data in order to ensure access to highly sensitive and highly curated data in safe ways. Their protocols permit linkages to auxiliary data and access to personally identifiable information, while protecting privacy. For example, the Centers for Medicare and Medicaid Services (CMS) curates and supplies extracts of administrative data based on an analyst's needs, including personal identifiers when necessary, through a virtual research data center. CMS can link files for analyses, and researchers can use their own laptop to log into the CMS safe setting, a secure environment, from which no data leaves.

IMPLEMENT SAFE DATA ACCESS MODELS

Agency leads, policymakers, and data architects should build safe projects by outlining clear and thorough governance protocols and agreements.

Data system architects can look to existing government agencies, such as the Census Bureau's Federal Statistical Research Data Center (FSRDC), to determine how to implement safe data access models, including issues like processing bottlenecks for data hosting, analyst credentialing, shifting disclosure review requirements, and providing remote access for analysts living far from current labs. FSRDC acts as a data intermediary, harmonizing and linking data from many sources for agency staff and the FSRDC labs.

1 O'Hara, A. (2019). Postsecondary data infrastructure: What is possible today. Retrieved from www.ihep.org/sites/default/files/uploads/docs/pubs/ihep_privacy_brief_data_sharing_v2.pdf.

2 The terms "analyst" and "data user" are used interchangeably.