disc
Data Integration Support Center
WestEd

2023
Issue 1

Adding a Privacy Focus to Security Reviews for Integrated Data Systems

INTEGRATED DATA SYSTEMS connect data over time and across sectors to provide data insights that support leaders in answering policy questions, directing resources, and better supporting individuals.

Ensuring the security of data systems is critical. What adult in this country has not at one time or another been notified that their personal information has been stolen or otherwise compromised? And who has not read or heard news reports of cyberattacks aimed at crippling major data infrastructure systems, sometimes for political reasons, sometimes for ransom? A mid-2022 *Wired* article (Newman) provides examples of successful cyber assaults from just the first 6 months of that year, including a ransomware attack on Costa Rica's Ministry of Finance that shut down the country's import/export businesses, leading the president to declare a national emergency.

Also reported in the article are cases in which large swaths of personal information were accessed with the intent of monetizing the data by selling or otherwise using them to commit fraud. In just one of several known data thefts from health care providers in early 2022, one monthlong breach of the data system of a large Massachusetts health care provider resulted in unauthorized access to data for some 2 million patients. Those data included names, Social Security numbers, and birth dates as well as addresses, billing information, and medical diagnoses.

Building and maintaining a data system that can withstand evolving and increasingly sophisticated cyberattacks is a challenge for any organization whose operations rely heavily on a data infrastructure. But organizations that collect personal information about the people they serve have the added responsibility and challenge of protecting the privacy of those individuals represented by the data. When such organizations merge their data into a large integrated data system (IDS), the security and privacy issues become significantly more complex.

Addressing the Privacy and Security Needs of Integrated Data Systems

Most commonly found in the public sector, an IDS is a system in which data from various sources are brought together as an integrated whole. By linking data across various public agencies, states can create a more holistic picture of outcomes for people in education, health care, the workforce, and beyond. This clearer picture enables a better understanding of the complex needs of individuals and communities, which, in turn, can inform the design of new strategies and interventions to address those needs and the evaluation of how well current programs and policies are achieving desired outcomes.

Data in these systems are generally collected and stored at the individual and identifiable level. The potential to link data to the individuals they represent makes IDSs subject to an array of laws and regulations intended to protect individuals' privacy and the security of the systems. Part of the complexity of ensuring data privacy within an IDS is that these laws and regulations vary by sector and among the agencies contributing or using the data. For example, an education agency and a health agency or provider submitting data to a statewide IDS must each comply with a different federal privacy-related law—the Family Educational Rights and Privacy Act (FERPA) focuses on student education records, while the Health Insurance Portability and Accountability Act of 1996 (HIPAA) focuses on patient health information. Each law has its own rules and regulations that must be considered in designing an IDS that includes both sectors.

Recognizing the need to protect the security of the individual data, the privacy of the humans represented in these data, and the operational integrity of large public-sector data systems, many states mandate an annual security review (or other protective measures, like an audit) for IDSs and other large publicly maintained data systems (e.g., for an individual state agency that keeps personal data). Mandated security reviews are typically aligned with national security frameworks such as those developed by the National Institute of Standards and Technology or the International Organization for Standardization. The intent of applying such frameworks is to identify any security- and privacy-related problems that system designers might have missed, or that the hosting agency might have introduced, so that identified risks may be mitigated. Mandated reviews look at the technological controls and structures in place for how a data system is developed and executed, based on the types of data the system stores, transfers, uses, and reports.

But do mandated reviews, as traditionally conceived and conducted, provide a full enough picture of an IDS and its unique security and privacy requirements?

This paper suggests they do not. Essential as they are for understanding many critical aspects of a data system, mandated reviews tend to operate with a particular—and narrow—definition of security that limits their ability to give a full picture of how well an IDS is able to protect the privacy of those whose personal information it contains. This paper explains those limits and recommends that a mandated security review be augmented by an enhanced system review whose more encompassing focus on privacy can help improve the security-and-privacy posture of a data system.

In writing about the limits of mandated security reviews and the need to do more to protect privacy, the authors have drawn on their more than 30 combined years of experience working in or with data integration efforts in the fields of defense, criminal justice, and education in the for-profit, not-for-profit, and governmental spaces. In addition, the authors have conducted over 300 onsite security-related engagements across every state and territory in the United States. Collectively, they have spent the last 15 years leading national efforts in data integration privacy and security as well as consulting on data-system governance. Finally, two of the three authors have successfully led multiagency data integration efforts and currently serve as lead facilitators of technical privacy and security for large federal and state data integration efforts. In early 2022, to generate the most up-to-date information about the data integration space, the authors and colleagues conducted research through observations, surveys, and interviews with data integration leadership in the public sector.

The Importance and Limits of Mandated Security Reviews

When considering how to safeguard the data for which they are responsible, many data system designers and reviewers see their role primarily, if not exclusively, as making sure the data system can repel those intent on breaching it for malicious purposes, such as stealing personal data with the intent to profit. Many conflate security and privacy, assuming that if unauthorized users cannot gain access to the data, the privacy of those data is ensured.

It is, in fact, critical to fend off would-be intruders. So, much like homeowners or renters who worry about break-ins, those responsible for securing a data system work to harden it, in this case with the technological version of locks and alarms. The intent is to outsmart anyone trying to overcome a system through *brute force*, "a method of accessing an obstructed device [or system] by attempting multiple combinations of numeric/alphanumeric passwords" (National Institute of Standards and Technology, n.d.). Guided by concern about whether a data system has adequate "perimeter protections" in place, a mandated system review is likely to map the system's architecture and security profile and analyze the system's vulnerability through system scans and penetration tests.

In addition to checking that all necessary technology-based protections are in place, a mandated review may consider the potential for human error that can undermine the ability of even the most solid technological solutions to repel unauthorized users but may overlook misuse by authorized users. A 2020 joint study by Stanford University Professor Jeff Hancock and researchers from the data security firm Tessian found that 88 percent of data breaches are caused by human error and that "user error is among the fastest-growing causes of breaches." Additional reporting (Shah, 2020) found that information technology leaders count implementing security-awareness training and "following company policies and procedures" (p. 1) as the most effective ways to prevent data loss. With human error in mind, a mandated review generally looks at what system operators do (e.g., require annual security training) to help ensure that authorized users—those with legitimate access to the system's data—do not inadvertently open the system's door to outsiders. Authorized users might do so by falling prey to social engineering efforts (i.e., various forms of phishing) or by otherwise failing to think about security when they access the data system. For example, after signing into the system, they may walk away to get coffee, leaving their computer unattended, or, when signing in with their password, they may make no attempt to ensure that no one is looking over their shoulder.

This focus on hardening data systems and keeping users from unintentionally undermining a system's built-in protections is important. But an exclusive focus on ensuring that unauthorized users cannot get their hands on protected data leaves system operators without the full range of information needed to effectively carry out their privacy-related responsibilities. Mandated reviews yield a myopic view of the system because they tend to overlook the ways in which both security and privacy can be compromised by how authorized users themselves handle the system's data.

The Case for an Enhanced System Review

An enhanced system review examines those aspects of an IDS in which security and privacy intersect. More specifically, it reviews the system with an eye toward any potential for privacy violations, not just by outside actors (directly and, also, indirectly, as when authorized users fall prey to social engineering efforts) but by unprompted human error in how authorized users manage, report, share, or otherwise use the data. Enhanced system review includes an alignment of privacy requirements associated with data types and ensures appropriate controls are in place to minimize risk in the design of the entire IDS. An enhanced system review is intended to augment rather than replace a mandated security review. Whereas mandated reviews assess the degree to which a system aligns with a national security framework and consider the effectiveness of the system's technology-based security elements, enhanced system reviews assess an IDS's adherence to its legal framework and agreements (e.g., data-use agreements) among its participating agencies and focus on the degree to which the system is set up to ensure appropriate use of the data by authorized users.

Security – ensuring that a data system and the specific data within it are protected from unauthorized access

Privacy – ensuring that the identities of individuals represented by the information in a data system are protected

By following up on a mandated security review, an enhanced system review can help a maturing IDS address both security and privacy risks as it moves forward. An enhanced system review looks primarily, though not necessarily exclusively, at the following aspects of an IDS, principally in the interest of ensuring privacy:

- IDS alignment with its legal framework and other structuring documents (e.g., policies, procedures, interagency agreements)
- training to support privacy and security
- use of technology to prevent or mitigate human error

An IDS's Alignment With Its Structuring Documents

An IDS's legal framework delineates privacy and security requirements mandated by federal, state, and local laws, regulations, and policies—mandates that, as noted above, can vary by each sector that contributes data to the system. The legal framework often (but not always) establishes an IDS's overarching purpose, scope, and use. It also establishes any restrictions related to privacy, security, and/or compliance, such as ownership of source data; hosting requirements; use limitations; indemnification; breach notification requirements; and alignment with federal, state, and local regulatory requirements. (See *Using a Legal Framework Approach for Integrated Data Systems*.)

The legal framework usually includes information on what data may be included in the system and their permitted or prohibited uses. All other structuring documents, such as policies, procedures, and formal agreements (e.g., data-sharing agreements, memoranda of understanding, vendor agreements), must adhere to the legal framework. Mandated reviews look at alignment between the IDS and a national security framework, but because they do not account for an IDS's legal framework, they could overlook potential within the system for data misuse.

As noted, government requirements for data sharing vary by sector. The more sectors participating in an IDS, the more complicated the legal framework (and other supporting documents) will be because of that variance. Below are just two examples of the complex federal regulations to which an IDS may need to adhere depending on what sectors are participating:

- The federal **Workforce Innovation and Opportunity Act** (WIOA) (Pub. L. 113–128) requires that states use education information and quarterly wage records to measure the state's progress in meeting performance reporting and evaluation requirements. An enhanced system review for an IDS that includes this information will ensure that its legal framework, other structuring documents, and system policy and procedures adhere to the complex disclosure requirements for personal information from education records, vocational rehabilitation agencies, and state unemployment compensation agencies.
- The **Higher Education Act** (HEA) restricts the sharing of Financial Student Aid data without a student's written consent, and, in some cases, sharing is prohibited even with the student's permission. An enhanced system review will analyze the IDS's use of Financial Student Aid data to ensure compliance with the legal framework and identify if additional structuring documents (policies, procedures, or written agreements) are necessary to ensure appropriate use of the data within the IDS.

Tightly tied to an IDS's legal framework are its governance system and organization, which enable the IDS to operate. An enhanced system review considers these aspects of an IDS, asking such big picture questions as the following:

Carefully conceived and refined IDS policies, procedures, and related architecture mean little if authorized users do not fully understand how the system works or the variety of ways in which they, as individuals, are responsible for ensuring both privacy and security.

Are policies and procedures in place to support change in leadership, in statutes and regulations, and in IDS membership?

Change is an inevitable part of any IDS. Structuring documents in conjunction with the legal framework ensures continuity when there are significant changes. These can include changes in leadership within agencies or at the legislative or gubernatorial levels, changes when laws or regulations change, and changes when a new sector joins the IDS. An enhanced system review can identify gaps in governance that may inhibit the IDS's ability to adapt easily and quickly to change.

Does the IDS structure ensure that all participating agencies receive the benefits that prompted their participation in the IDS?

An enhanced system review may examine the enacting legislation or the other reasons that prompted participation of member agencies in the IDS to evaluate if each member agency benefits from participation in the IDS. The review may also identify things that might inhibit member agencies' full participation or benefit.

Does the IDS have adequate policies and procedures in place for responding to requests under its state's open records laws without inadvertently violating the privacy of those whose data might be part of the requested information?

Many states have open records laws allowing access to public records, including information that may be housed in an IDS. Because IDS data include individuals' personal information, an IDS must have policies and procedures, aligned with its legal framework, that define what information housed in the IDS may be made available to the public and in what form. An enhanced system review will evaluate the IDS's policies and procedures for open records against the legal framework and verify whether the IDS's responses to such requests sufficiently protect individuals' personal information.

Training in Support of IDS Privacy and Security

Carefully conceived and refined IDS policies, procedures, and related architecture mean little if authorized users do not fully understand how the system works or the variety of ways in which they, as individuals, are responsible for ensuring both privacy and security. High-quality training is essential for staff at the IDS host agency and at agencies participating in and contributing data to the IDS as well as for any individuals who are authorized to use the data. Training programs must be relevant, effective, and monitored.

An enhanced system review checks whether all users (internal and external) are included in the training program and whether the training is aligned with the legal framework and structuring documents. This review provides insight into any gaps in training and suggests and prioritizes steps to improve training. An enhanced system review may also identify important training components that are typically overlooked, such as training on access controls, appropriate disclosure-avoidance procedures, and sector-specific regulations and protections.

Leveraging Privacy-Enhancing Technology to Prevent or Mitigate Human Error

Regardless of policies, procedures, training, or legal frameworks, the potential for human error by authorized users persists. However, privacy-enhancing technologies are increasingly being developed to protect the personal information of the individuals in a data system of any kind. For example, the Massive Data Institute's *Privacy Preserving Technologies In Education* report summarizes the privacy-preserving technologies that have been discussed, tested, implemented—and abandoned—to date in the education field.

While the use of privacy-enhancing technologies assists an IDS in mitigating human error, such solutions have advantages and disadvantages, so decisions about whether to use a privacy-enhancing technology must be based on the particulars of the IDS. An enhanced system review can assist an IDS in identifying the most appropriate solutions unique to its own context for mitigating human error. Additionally, an enhanced security review can validate the use of a previously implemented privacy-enhancing technology to ensure compliance with the IDS legal framework, policies, and procedures.

Resources

The following are some useful resources to consult as organizations consider an enhanced security review:

- **Privacy Impact Assessments:**
  - The Privacy Impact Assessment, developed by the Department of Homeland Security, is a decision tool used to identify and mitigate risks when developing or implementing technologies or systems that handle or collect personal information.
  - A blog post from the Future of Privacy Forum documents how privacy impact assessment policies help cities use and share data responsibly with their communities and includes a Model Privacy Impact Assessment Policy.
  - The Data Protection Impact Assessment is required under the European Union's General Data Protection Regulation law under certain conditions. A Data Protection Impact Assessment is conducted before and during the planning stages of a project in which the processing of data is likely to result in a high risk to the rights and freedoms of people.
- *Guide for Community Training on Data and Technology*: This guide from the National Neighborhood Indicators Partnership describes the process for developing a comprehensive training program that fosters a culture of privacy and reinforces appropriate use of data.
- *Integrated Data Systems and Student Privacy*: This guidance document supports the development of a legal framework by explaining how FERPA must be considered in an IDS.

How the Data Integration Support Center Can Help

The Data Integration Support Center (DISC) at WestEd offers enhanced system reviews that augment mandated security reviews by focusing particularly—though not exclusively—on alignment with the IDS's legal framework and the potential for human error that can compromise IDS security and data privacy. Starting in fall 2023, DISC will offer annual enhanced system reviews to a limited number of IDSs free of cost. The reviews are designed to yield an improvement roadmap for the IDS. The DISC is also available to assist the IDS in implementing the recommendations from that roadmap.

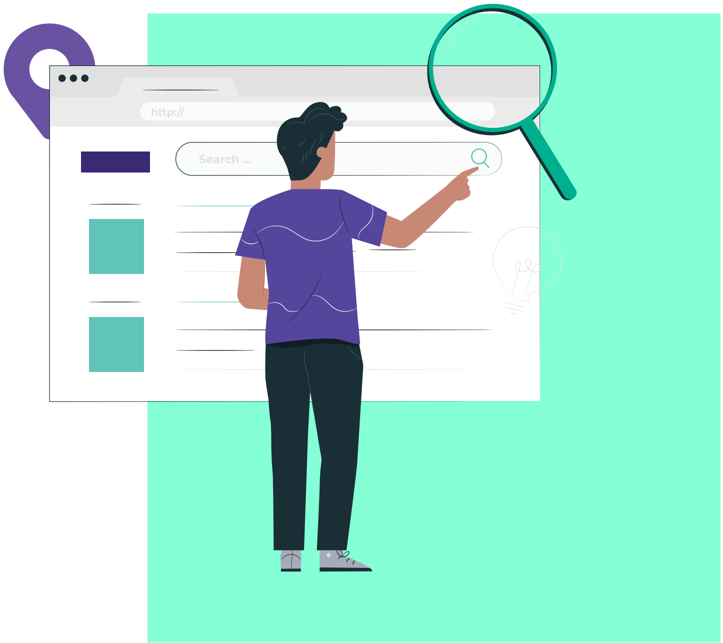
References

Hancock, J. (2020). *Psychology of human error.* Tessian. https://www.tessian.com/research/the-psychology-of-human-error/

National Institute of Standards and Technology. (n.d.). *Computer Security Resource Center: Glossary*. https://csrc.nist.gov/glossary/term/brute_force_password_attack

Newman, L. H. (2022, July 4). The worst hacks and breaches of 2022 so far. *Wired*. https://www.wired.com/story/worst-hacks-breaches-2022/

Shah, S. (2020, May 31). More cyber training does not mean fewer data breaches. *Forbes*. https://www.forbes.com/sites/soorajshah/2020/05/31/more-cyber-training-does-not-mean-fewer-data-breaches/?sh=4359370a640c



A project of WestEd
WestEd.org

Suggested citation: Rodriguez, B., Cottrell, S., & Tiderman, L. (2023). *Adding a privacy focus to security reviews for integrated data systems*. WestEd.

WestEd, a nonpartisan research, development, and service agency, works with education and other communities to promote excellence, achieve equity, and improve learning for children, youth, and adults. WestEd has more than a dozen offices nationwide. More information about WestEd is available at WestEd.org.