# How Learning Data Impacts Privacy

## Scenario

Like virtually all colleges and universities, Emerald University has experienced an explosion in the use of learning-related data over the past decade. In tandem with its efforts to provide the right technologies, Emerald has also worked hard to develop the right portfolio of policies and procedures to guide its use of learning data. In particular, protecting data privacy has been a central focus of that work.

Data privacy issues have become more complicated as data technology has advanced and expanded. Like many of her peers, for example, Emerald CIO Jane Parsons has noted that more of the university's learning data are flowing through tools and services provided by third parties. She knows that means that Emerald data are increasingly traveling through networks and systems outside the campus, including the cloud, and that such data transfers lack the controls that Emerald's internal processes and policies typically provide. She knows, too, that vendors in the private sector are finding new ways to monetize data—a focus that may be at odds with institutional interests in ensuring data privacy.

As at many institutions, different campus stakeholders at Emerald are raising more issues about the security and privacy of the university's learning data. More students are expressing concern about the privacy of their own information—including essays and other intellectual property uploaded to cloud systems—and more are choosing to keep parts of their academic records private. Faculty report that more students are opting out of assignments that use their personal, identifying information. Prompted by concerns about potential liabilities, campus risk managers and even some of Emerald's board members have been asking Parsons about the university's data privacy practices and policies.

Recognizing that this spate of recent developments constitutes a burgeoning change in the learning data landscape at Emerald, Parsons has ordered a broad audit to determine how the university is using third-party vendors for learning data. She plans to use insights gained from the audit to then lead a review of all institutional data policies. One specific goal will be to identify how changes in the learning data landscape might suggest ways in which existing policies need to be amended to help ensure data privacy and whether new privacy policies are needed.

## 1 What is it?

Institutional programs and third-party vendors provide an expanding number of instructional tools and services that generate copious amounts of learning data. These data can be combined and shared to improve learning and increase student success, but these **opportunities complicate the privacy landscape**. Service providers can derive economic value from the data they harvest, creating tension between the pedagogical benefits to students and the business interests of vendors. As higher education increasingly finds uses for instructional tools and services, and as "cloud first" strategies increasingly become the default, colleges and universities should actively consider their governance and policy structures to address the privacy of learning data.

## 2 How does it work?

Privacy is the right of an individual to control his or her data and specify how those data are collected, used, and shared. When learning data were only shared between faculty and students—as in a student essay with instructor feedback—privacy was relatively straightforward. Learning data today can be derived from and shared between numerous systems, services, and tools. Learning data encompass learning content, assessment and evaluation data, human added-value content, and platform tracking content. Some—but not all—of these data are protected by statute, regulation, or policy. Because learning data increasingly reside outside campus networks and systems, internal privacy controls typically provided by institutional processes and policies are sometimes circumvented. Complicating matters, vendors might not adequately protect privacy rights, and many faculty, students, and staff do not take the time to read the fine print of click-through agreements regarding data use. Given the **multiplicity of sources and uses of learning data**, and in light of the evolving interests of service providers, managing the privacy of learning data requires user education and awareness, as well as policies and enforcement aligned with the emerging impacts of learning data on privacy.

# How Learning Data Impacts Privacy

## 3 Who's doing it?

Some colleges and universities have begun the process of shining light on the privacy risks that are introduced by the many new opportunities to share and combine learning data, though higher education overall remains in the early phases of building awareness. Institutions such as the University of California and the University of Michigan are **drafting learning data privacy principles**. The University of California at Berkeley is developing a tool that allows students to specify which analytics services can access the student's learning data and which cannot. The Asilomar Convention, EDUCAUSE, and IMS Global are contributing to conversations in North America and Europe around how higher education institutions and third-party vendors can best collaborate to address emerging concerns about how best to capitalize on the expanding opportunities of learning data while meeting commitments to transparency and privacy.

## 4 Why is it significant?

Learning data are important for measuring learning efficacy and student outcomes, and such data are increasingly required as evidence for accreditors, federally funded grant agencies, state performance-based funding, regulatory agencies, and similar applications. Policies governing the privacy of such data are vital for protecting the interests of students, faculty, staff, and the institution as a whole. But the explosion of learning data and the acceleration in the development of new uses for such data suggest that higher education may need **new and perhaps more nuanced practices and policies concerning learning data usage and privacy**. Specifically, new conversations with vendors may be needed about who controls those data, who has access to which data, and how those data are used (including whether data may be sold). Vendors' interests in monetizing data must be tempered by higher education's interest in upholding principles and practices around privacy.

## 5 What are the downsides?

Trends in learning data raise many privacy questions. What types of data are collected and stored? Who owns the data, and what guidelines are appropriate for whether data may be bought and sold? Who makes those decisions? A highly restrictive privacy policy might deny access to a valuable tool and could be seen as an infringement of academic freedom, limiting the opportunities that the wide sharing of learning data might otherwise offer. Applications designed to improve learning and student success typically work better with more data, but **opting in exposes users to privacy risks, while providing an opt-out choice could result in skewed and incomplete data**. It might prove challenging to develop privacy policies that satisfy the diverse needs and interests of stakeholders including students, faculty, administrators, and vendors. Ethical concerns arise, too, such as bias that can creep into predictive models and whether an institution has an obligation to assist all of the students it identifies as at-risk.

## 6 Where is it going?

Although the question of whether the *expectations* of privacy are changing is a matter of some debate, the *conditions* and *forms* of privacy are certainly in flux. Students, faculty, and staff need ongoing training, and campuses should cultivate an ongoing discussion around the **cultural shifts related to the privacy of learning data**. Better education might be needed to help campus stakeholders assess when the privacy of learning data is or is not protected. As institutions frame the terms of partnerships with vendors, clarifying the dimensions of shared data will be imperative. Students and faculty might be given more opportunities to decide whether to participate in the collection of learning data. Standards for integrating learning data across systems are seeing increased adoption, and vendors are beginning to offer feature sets that include aggregation, export, and analysis of usage data.

## 7 What are the implications for teaching and learning?

Developments in learning applications and services have the potential to significantly improve student learning and success, but many of those opportunities depend on learning data that can be sensitive and possibly protected. Moreover, the value of data persists beyond the scope of a course or any student's academic career. As colleges and universities develop policies to balance learning innovations with user privacy, students and faculty might increasingly be expected to take an active role in deciding what data are collected and how they are used. Higher education might influence the marketplace so that learning tools and services use learning data in ways that align with the interests of learners and institutions. Although the answers are far from clear, institutions should initiate discussions and governance conversations if they haven't already done so—the only real mistake is to do nothing.