:: IT Management

GAME CHANGER

RETHINKING HOW TECHNOLOGY IS USED IN EDUCATION

DATA SECURITY

High Stakes: Research Requires Smart(er) Secure Data Protection

Managing sensitive research data takes savvy security, storage, and recovery practices. The right mix of tools and business processes can help.

RESEARCH IS BIG BUSINESS. Fiscal 2021 numbers released in December 2022 showed U.S. universities reported the largest growth in federally funded R&D expenditures since fiscal 2011, to the tune of $89.0 billion, up 4% from fiscal 2020, **according to the Higher Education Research and Development Survey sponsored by the National Science Foundation's National Center for Science and Engineering Statistics.**

That growth goes hand-in-hand with policy and technology developments that add yet another layer of complexity. Government increasingly calls upon colleges and universities to follow stringent rules and protocols defining the protection and maintenance of data collected through the course of government-funded research, for instance, and cloud computing's rapid expansion has ushered in significant changes to the ways research teams store and manage data and access.

PRODUCED BY: CAMPUS TECHNOLOGY

SPONSORED BY: COMMVAULT®

Data now stands as the primary product of research and the lifeblood of the researcher's work. And in this era of Big Data, with massive amounts of information that must be rapidly accessed and stored on prem or in the cloud, even maintaining a handle on where it's all hosted within a Research-1 or Research-2 institutional setting can prove a daunting task. Without support from central IT, researchers find themselves working more as full-time data managers or auditors rather than focusing on the crux of their research.

Support from groups like EDUCAUSE and resources like the **Higher Education Community Vendor Assessment Toolkit (HECVAT)** go a long way toward standardizing approaches and sharing best practices for managing research data. And a growing class of tools that leverage AI and machine learning to help IT and researchers readily catalogue, classify, tag, and ultimately protect data rather than simply back it up or recover it provide new and improved solutions for ensuring compliance with rules governing access, privacy, and security.

Model Success

The University of Michigan provides a robust example of how guidelines around research data can be established and communicated. The institution spent $1.64 billion on research in fiscal 2020, according to the NSF survey, second only to Johns Hopkins University, which spent $3.11 billion in that same time. Michigan's **Data Security Guidelines** from Research Ethics and Compliance offer detailed guidance to principal research investigators and their study teams for maintaining human subject data securely, while also ensuring anonymity, confidentiality and de-identification. Even in instances where personally identifiable data is not collected or used, research teams are advised to follow core data security controls, codified in the IT department's Data Management and Security Protocol. The university has also established rules for **Third Party Vendor Security & Compliance**, which guide teams in vendor selection wherever non-university products or services will be used to store, process or transmit university data.

This type of approach ensures that research teams can take the lead on reviewing the mix of tools and solutions — on prem, cloud-based, or hybrid — that make the most sense for their work, while also consulting and informing central IT, and ensuring compliance with university policies and risk management in addition to individual grant requirements.

Keep It Simple

Research grants stipulate rules that govern data protection and maintenance, but the solutions and tools research teams and funding recipients rely on to meet those requirements vary from institution to

A growing class of tools that leverage AI and machine learning to help IT and researchers readily catalogue, classify, tag, and ultimately protect data rather than simply back it up or recover it provide new and improved solutions for ensuring compliance with rules governing access, privacy, and security.

institution, and even within distinct research groups within an institution. What should be near the top of their list of requirements for data protection? Simplicity, security, and autonomy and cost typically top their lists.

When the IT team at Illinois State University in Normal, IL, took the opportunity to simplify its approach to data protection as it shifted to a hybrid cloud environment, a key benefit was found in protection tools and documentation that readily met grant funding requirements, according to Tim Walsh, a sysadmin at the university: "I've had people come to me to ask what our backup strategy is because they have a grant that has a requirement. Now I can just say, 'It's all taken care of. As long as it's here, as long as you're putting it in your normal spaces, it's taken care of.' I think that's been huge for a lot of our researchers, who can now think more about like smashing stars and vortex beams. They don't think about where their data goes and how they have to handle it in case of an emergency. They just have to think about their research."

Refine the Mix

While no one solution will work for every research team across the institution, generally the mix includes:

- Data discovery and classification software;
- Backup and recovery;
- Security information and event management (SIEM);
- Data loss prevention (DLP);
- User and entity behavior analytics (UEBA); and
- Network security.

Maintaining ease of use when deploying any data protection solution is a top line goal for most teams but given the level of protection that may be required for different data sets, rapid retrieval isn't always possible. It's critical to artfully mix the ingredients above to achieve an effective and efficient security of the research data.

From a risk management perspective, not all data requires the same level of protection. Data discovery and classification tools have shifted to the forefront of the data protection mix for that very reason, ensuring that all types of data are scanned and accounted for, sorted, labeled and ranked or prioritized based on compliance and protection needs or standards. These tools tag data with unique signatures that enable security to be customized around each unique set, while also providing analytics and insight into use and movement. AI can monitor data in the organization's network or research sandbox and protect it using specific guidance or rules frameworks, as granular as required.

DATA COMPLIANCE AND GOVERANCE

Unraveling Risk and Compliance

Here's where to start for guidance on data management, classification, governance, and more.

UNDERSTANDING ALL THE RESEARCH DATA AN INSTITUTION has and where it is stored, isn't as simple a task as it perhaps once was, given the proliferation of cloud-based platforms, rogue or shadow IT, and the just plain massive amounts of data generated not only through the course of research but also administration (to name but one source), in addition to various storage types.

When teams treat or manage data as a strategic asset rather than simply a security risk or liability, they start to approach data management as a means of achieving mission-critical priorities.

Asset and data management must include means of discovery, ownership, value, acceptable use, protection, and disposal of information-related assets, **according to guidance from EDUCAUSE**. Those assets may be hardware (tangible) or software and data (intangible). Regardless of an institution's size or research funding status, EDUCAUSE's "4 Knows" offer a strong start for any asset and data management initiative:

- **Know what** you have.
- **Know where** it is.
- **Know who** owns it and who maintains it.
- **Know how** important it is to the institution.

EDUCAUSE has also developed a **Data Classification Toolkit**, which serves as a helpful primer for getting started with addressing the following questions:

- **Need:** Why is it necessary or mandatory to classify data?
- **Roles:** Who should classify what data?
- **Methods:** How should data be classified?
- Are there any **best (or common) practices** available?
- **Impact:** What processes are dependent or impacted by data classification?

Data governance is a critical component of the classification exercise and must be detailed within any data cataloguing activity, ensuring research teams understand and can communicate ownership, the specific terms that apply to it, and any other limitations.

Where previously a grant funding audit might have looked at where funds were allocated and tools that were purchased, audits now require robust security and data documentation, detailing who has access, and more. Empowering researchers with the ability to respond to those queries quickly and easily, without requiring additional support from IT or HPC teams, falls squarely within the capabilities of the latest intelligent, cloud-based data management and protection tools.

Assessing Solutions

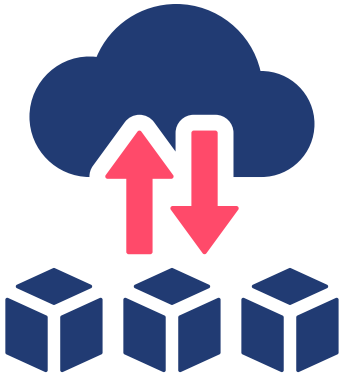
The **Higher Education Community Vendor Assessment Toolkit (HECVAT)** offers a definitive framework for U.S. higher education to measure vendor solution risk, particularly within the areas of data protection and cybersecurity (and originally focused on cloud computing vendors). Created by the Higher Education Information Security Council and Shared Assessments Working Group, in collaboration with Internet2 and REN-ISAC, the tool "attempts to generalize higher education information security and data protection questions and issues regarding cloud services for consistency and ease of use."

HECVAT helps institutions ensure they appropriately assess cloud services for security and privacy needs, including many capabilities or functions unique to research and higher education. The framework also provides a consistent and easy-to-adopt methodology for reducing costs without increasing risks and reduces cloud providers' burden to respond to requests for security assessments from higher ed customers, according to REN-ISAC.

"The HECVAT is an example of how increasing collaboration across higher education institutions and organizations can facilitate advances in security risk management and streamline procurement processes," said Nick Lewis, program manager, Security and Identity, Internet2.

EDUCAUSE advises institutions to request any third-party solution providers complete a HECVAT questionnaire to confirm alignment with required policies governing sensitive information and PII. More than 150 colleges and universities now rely on HECVAT to reduce risk and improve efficient data and cloud management, and more than 50 solution vendors have already completed the assessment and shared it within the Cloud Broker Index.

When teams treat or manage data as a strategic asset rather than simply a security risk or liability, they start to approach data management as a means of achieving mission-critical priorities.



GAME CHANGER

RETHINKING HOW TECHNOLOGY IS USED IN EDUCATION

DATA PROTECTION

Ensure More Agile Protection

Simple is best when it comes to securing research data, boosting collaboration, and meeting grant requirements.

MANY CHANGES ARE AFOOT AT RESEARCH INSTITUTIONS today, not least of which is their approach to data management and storage, brought about by demands for more agile computing and storage as well as more stringent data protection requirements and privacy regulations. As these complex, sometimes competing requirements put pressure on institutions to modernize infrastructure and rely more heavily on the cloud, new ways of protecting and managing data expand IT's ability to more deftly meet researchers' needs.

And while the requisite infrastructure rapidly evolves, in many ways researchers' needs haven't changed: "Simplicity in use, security of their data with dynamic collaboration, and autonomy to conduct their research while meeting the terms of the grant that funds their work," said David DeVries, SLED CTO and senior director of strategic initiatives for SLED at Commvault, a provider of intelligent data management tools and services. (DeVries is also the former director of the State of Michigan's Department of Technology, Management and Budget, and the former CIO for the U.S. Office of Personnel Management, and former deputy CIO, Department of Defense.)

"From the perspective of a university, often central IT provides storage and compute services to the various departments and groups," DeVries noted. "This is becoming even more centralized to realize the efficiencies of providing that from a central, professional group, with security in-depth to protect their data and systems. Out of necessity, there's governance — meeting the security controls, meeting the retention and compliance requirements, and efficient use of resources. From the researchers' perspective, ease of use is very important, along with autonomy of tagging their data, understanding and controlling who has access to the data. At the end of the day, they are responsible for ensuring compliance with the rules laid down by the issuing grant."

Design With Growth in Mind

Craig Jackson, director of Infrastructure, Operations and Networking at Illinois State University, said in a recent **case study** that his team selected a Commvault solution for data backup and management when the institution shifted to a hybrid cloud environment. Last year, the university announced

it would expand, and plans to open a new college of engineering in 2025.

"Our HPC is really cranking up, and that's a campus service for our faculty, to use for their research without charging them for that. We've got a ton of AR/VR coming with new esports coming online as well. So there's a whole lot coming to Illinois State, and we have the solution that can protect the data for all of those things. The infrastructure has been built with growth in mind."

Commvault HyperScale X™ allows IT to simplify all of the university's otherwise complex backup infrastructure, Jackson said.

In the ISU **case study**, Tim Walsh, a system administrator at ISU, agreed: "Commvault allows us to live in this weird hybrid environment, mostly because Commvault lives in both environments. If I need something dealt with in AWS, it's already in AWS and I have the data there, I can spin it up from AWS .... It's the seamless meshing between on prem and the cloud, in a single click."

"The goal is to bring simplicity to those who must host, secure, maintain, and protect the data, and at the same time give the capability to the researchers so that they go and find something," Commvault's DeVries said. "Maybe they can find it in the lake, retrieve it and use it for their research, and keep on going."

More Complete Protection

Commvault works with higher education, government, business, and other industries around the world to deploy the appropriate, partner-agnostic data protection solutions and tools, ensuring efficient, secure, and agile data protection, no matter where it's hosted.

Commvault's portfolio of Intelligent Data Services includes:

- **Data Protection** — to rapidly recover data cost-effectively and at scale, on-prem or in the cloud
- **Data Security** — to detect, protect and recover data from ransomware attacks and other data breaches
- **Data Compliance and Governance** — Manage data access to drive regulatory compliance and mitigate data privacy risks
- **Data Transformation** — to seamlessly move data across environments for app modernization and flexible data use
- **Data Insights** — applies machine learning and artificial intelligence to optimize and automate IT processes

Learn more at **commvault.com**.

The goal is to bring simplicity to those who must host, secure, maintain, and protect the data, and at the same time give the capability to the researchers so that they go and find something.