Searching for a Smoking Gun, Chasing a Silver Bullet

Data Breaches in Higher Education

ECAR Research Bulletin | February 21, 2017

D. Christopher Brooks, EDUCAUSE
Joanna Lyn Grama, EDUCAUSE

We have to play by the rules, and the hackers don't! It's a lot like the war on terrorism. We have to be right all the time. The hackers only have to get lucky once. We always seem to have one hand tied.

—Lynn Carroll, *Entertaining War: Let the Games Begin*

Overview

EDUCAUSE members frequently seek information about data breaches in higher education: Who is experiencing data breaches? How do they happen? Do institutions successfully remediate following a first data breach (so that no more breaches happen)? Are there any factors that make higher education data breaches more or less likely? The EDUCAUSE Center for Analysis and Research (ECAR) published its first look at data breaches in higher education in 2014.[1]

The 2014 research showed that, compared to other industries, higher education had experienced a larger number of reported breaches from 2005 to 2014 but that fewer records were exposed per breach. It also found that while roughly 7% of all U.S. institutions had had at least one data breach, doctoral institutions were responsible for the majority of data breaches during that time period. Hacking and unintentional data disclosures were the most common types of breaches during this period.[2]

Our current research looks at whether any factors increase or decrease the likelihood of a higher education data breach. Is there a smoking gun, something found in every higher education data breach? And, conversely, is there a silver bullet—a control or controls that higher education institutions can employ to prevent data breaches?

Highlights

The short answer to both questions is no. No smoking gun is common to the breaches we studied, and we are unable to identify any silver bullet that can be employed to prevent future institutional data breaches.

Our current research reviews higher education data breaches reported in the Privacy Rights Clearinghouse (PRC) Chronology of Data Breaches and the EDUCAUSE Core Data Service (CDS) benchmarking information for the institutions experiencing a breach. Comparing data from these two sources allows us to see if any information security factors or controls are consistently present (or missing) from institutions that have experienced a breach. This approach is complicated by the possibility that not all higher education data breaches are reported in the PRC and the fact that not all higher education institutions contribute data to CDS.[3] The current research is further complicated because the number of data breaches in higher education seems to be falling, making valid analysis more difficult (see figure 1).[4]

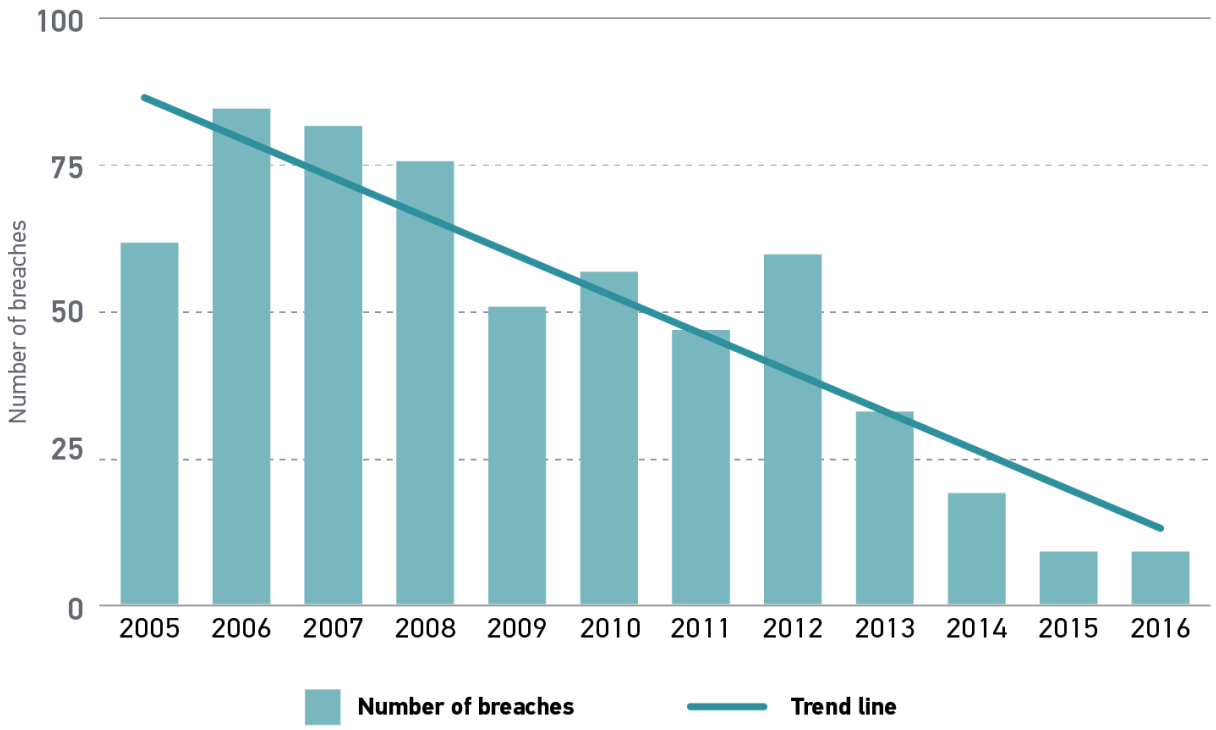

Figure 1. Number of data breaches in higher education, by year

Source: Privacy Rights Clearinghouse Chronology of Data Breaches

Our research for this report is confined to data breaches reported during the years 2014–16 because CDS first asked institutions to share data related to information security maturity and technology deployment in 2014. We chose to use these particular CDS factors in our analysis because we thought they might yield helpful information about an institution's capability to deliver information security that is related to incidents of data breach.

From a pool of 35 institutions that had experienced a data breach during this time (a total of 37 breaches), we were left with a research data set of 20 institutions (22 breaches). These 20 institutions experienced at least one breach from 2014 to 2016 *and* had completed CDS either in the year of the breach or in the year immediately preceding the breach.[5]

Because the number of cases in the data set was too small to conduct a proper statistical analysis, we resorted to a more qualitative comparative method that employs Boolean algebraic techniques.[6] First, we explored CDS to identify variables that could inform our understanding of breaches. For this project, we identified the following nine items:

1. Does the highest-ranking person with primary responsibility for information security hold the title of chief information security officer (CISO)? (CISO variable)[7]
2. Does the designated CISO devote 60% or more of his or her time to information security activities? (TIME variable)[8]

3. Does the institution have mandatory information security training for faculty and staff? (TRAINING variable)
4. Has the institution deployed data loss prevention technologies? (PREVENTION variable)
5. Has the institution deployed network intrusion detection technologies? (DETECTION variable)
6. Does the institution have an incident response process and designated staff roles? (ROLES variable)
7. Are incident response staff aware of legal or compliance requirements surrounding evidence collection? (LEGAL variable)
8. Does the institution have an information security policy? (POLICY variable)
9. Does the institution have an information security risk management process? (RISK variable)

We chose these variables because they offered the most interesting mix of policy, technology, and best-practices data relevant to a data breach and institutional information security maturity. We specifically did not include variables such as information security spending or information security staffing numbers because these data points were similar across all institutions.[9]

Second, we recoded each of the maturity or deployment-level variables[10] to coarsely reflect the simple presence or absence of the variable in question (see table 1 for maturity levels and table 2 for deployment levels; the coding schemes for the CISO and TIME variables are detailed in endnotes 7 and 8).

Table 1. Comparison of maturity-level variable to research variable recoding for TRAINING, ROLES, LEGAL, POLICY, and RISK variables

Maturity Level	Maturity Definition	Research Variable Recoding
Absent	Capability components are largely not achieved. Little to no planning is under way.	Not present
Initial	Capability components exist either latently or slightly. Early planning and discussions may be under way.	Not present
Developing	High-priority capability components may be largely or fully achieved, while other components are still maturing. Active planning and strategic attention are under way.	Present
Established	Capability components have been developed but may not yet be incorporated into institutional culture and practices. Efforts to improve sustainability or scalability are under way.	Present
Optimized	Capability components have been developed with an eye toward sustainability, adaptability, and scalability. Components are fully integrated into institutional practices and culture (and may be influencing both).	Present

Table 2. Comparison of deployment-level variable to research variable recoding for PREVENTION and DETECTION variables

Deployment Level	Deployment Definition	Research Variable Recoding
None	None of this technology or service is in place, and no work is under way or resources committed for this technology or service.	Not present
Tracking	Staff are assigned but restricted to monitoring and understanding this technology or service (much more than just reading articles).	Not present
Initial deployment	This technology or service is not yet available to users; however, meaningful planning for deployment is under way. A plan for deployment is either in development or in place.	Not present
Partial institutional deployment	Full, production-quality technical capability or service is in place, including ongoing maintenance, funding, etc., with potential access by selected users but not institution-wide.	Present
Full institutional deployment	Full, production-quality technical capability or service is in place, including ongoing maintenance, funding, etc., with deployment supporting potential access institution-wide.	Present

Third, we analyzed those recoded variables to glean both the number and type of patterns manifest in the 22 cases of data breaches. As part of this process, we also looked to see if any of the variables was present or absent for all of the cases. However, from the 22 cases we analyzed, we found 14 different patterns and only 4 patterns represented by more than one case (table 3). The sheer number of patterns precluded the application of further comparative techniques (e.g., comparing cases of breach to cases of nonbreach). More importantly, these patterns were so diverse that no single variable could be identified whose presence or absence was a necessary condition for the breach to occur. That is, we failed to find a smoking gun common to all breaches in our data set. This does not mean, however, that one does not exist; it just means that among the limited number of cases for which we have data, there appears to be no common indicator.

Table 3. Variable patterns present across data breach cases analyzed

Number of Institutions Showing Pattern	CISO	TIME	TRAINING	PREVENTION	DETECTION	ROLES	LEGAL	POLICY	RISK
5	X	X	X		X	X	X	X	X
3	X	X			X	X	X	X	X
2		X		X	X	X	X	X	X
2		X	X		X	X	X	X	
1		X			X		X	X	
1		X				X	X	X	X
1			X	X	X	X	X	X	
1			X					X	X
1				X	X		X	X	
1								X	
1		X	X	X	X	X	X	X	X
1	X	X						X	
1	X	X			X		X		
1	X	X	X		X			X	X

Note: Cells marked with an "X" indicate the presence of the variable; blank cells indicate the absence of the variable.

Few Patterns Emerge

Although our data and analysis do not reveal the presence of any smoking gun that explains why or how these breaches were possible, we were able to observe some loose patterns in the data that may prove helpful for institutions to consider.

Do Data Loss Prevention Technologies Matter?

In the most common pattern (n = 5), all of the variables were present except data loss prevention systems or technologies (PREVENTION variable). The next most common pattern (n = 3) also showed the absence of data loss prevention systems and technologies, as well as the absence of developing, established, or optimized information security awareness training (TRAINING variable). Moreover, although the presence of data loss prevention systems did not stop breaches from occurring, in 17 of the 22 cases of security breaches the meaningful deployment of such systems was absent. The absence of such systems could be of interest to higher education institutions, given that 19 of the 22 cases reviewed for this research involved electronic entry by an outside party or data loss via malware.[11]

Implementation of data loss prevention technologies was not common among the institutions represented in this research, and it is not common among higher education institutions in general. The 2016 EDUCAUSE Strategic Technologies research showed that only 4% of responding institutions had content-aware data loss prevention technologies in place institution-wide (see figure 2).[12] Over 70% of responding institutions were unfamiliar with the technology, had no deployment of the technology, or were merely tracking the technology for possible implementation. That report predicted that by 2021, implementation of this technology will still only be emergent, meaning that it will be deployed institution-wide in just 21–40% of institutions.[13] With this information for context, it is not surprising that most institutions in the current research did not have data loss prevention technologies in place.

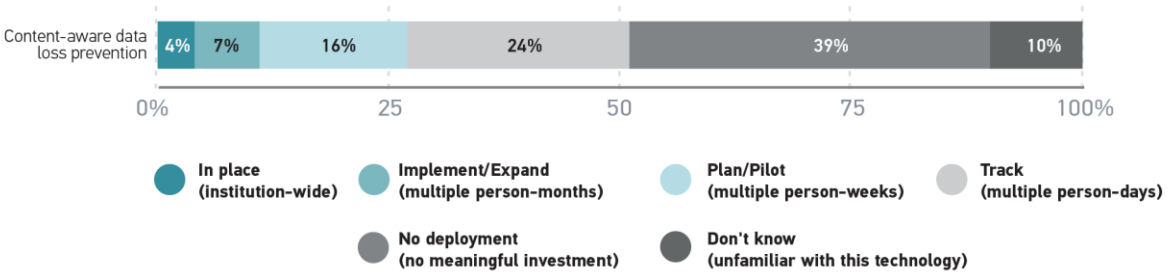

Figure 2. 2016, content-aware data loss prevention deployment

Are CISOs Critical?

Previous EDUCAUSE research has indicated that institutions with a chief information security officer or other full-time staff member devoted to information security are more likely to have implemented information security controls.[14] Of the 22 cases analyzed for this research, half occurred at institutions that had no CISO designated in the year that the breach occurred (CISO variable). Of the 11 institutions without a CISO, 7 of them had a person who spent 60% or more of his or her time dedicated to information security activities (TIME variable).

The use of the CISO title in higher education is still somewhat uncommon. In 2014, just 34% of U.S. institutions had a person whose time was 100% devoted to IT security, and only 32% of those individuals held the CISO title.[15] While information security leadership *titles* come in many forms, forthcoming ECAR research shows that the *functions* of CISOs are relatively consistent—each of six main tasks (one of which is incident management) is the responsibility of 90% or more of CISOs.[16] Coupled with our breach analysis, these data suggest the importance both of having a CISO in place and of having a significant amount of the CISO's time directly dedicated to information security activities.

It is interesting to note the interplay between the CISO and PREVENTION variables. In the 11 breaches where a CISO was present, data loss prevention systems were absent. But, 5 of the 11 institutions that reported *not* having a CISO *did* have data loss prevention systems deployed in a targeted or institution-wide manner. Our data do not suggest that data loss prevention systems could replace a CISO, but it is an interesting pattern to note for comment by the information security community. Is it possible that the implementation of certain information security technologies alleviates the need for a CISO or full-time information security leader? Or do institutions without CISOs view data loss prevention systems as a legitimate substitute, or *in loco CISO*?[17]

Does Information Security Awareness Training Really Work?

In half of the breaches we analyzed there was a reported dearth of information security awareness training for staff, students, and third parties who were interacting with institutional systems. That is, in 11 instances of a data breach, institutions either had no information security awareness training in place or engaged in training in only an informal manner.[18]

This variable result is a bit surprising, especially because in 2015, almost 75% of U.S. institutions required information security training for faculty and staff.[19] While the sample size in this research is too small to draw a strong conclusion about the relationship between a data breach and information security awareness training, such training is widely viewed as an essential part of any institutional information security strategy.

What It Means to Higher Education

For the second year in a row, information security is the top issue on the EDUCAUSE Top 10 IT Issues list. It garners the top spot on the list because IT leaders recognize that institutional data—be they research data, student data, health center data, or business data—are a valuable commodity, not just to the institution but also to its partners and the world at large. Protecting the institution and the IT systems and resources that use, process, transmit, and store institutional data from information security threats is a top priority. Data breaches are not the only information security threat to an institution, but they often receive the most attention because data breach stories can be sensational and the tangible and intangible consequences of a data breach can be wide-ranging.

Our current research shows that data breaches and the circumstances contributing to those breaches are unique. Even though this research points to no single security control or technology that can be implemented to prevent a higher education data breach, there are basic steps that an institution can take to improve its information security posture:

- *Designate an individual responsible for information security.* An effective leader who can communicate information security issues across the institution is essential for information security program success.[20]
- *Implement an institutional information security policy based on recognized best practices*. Information security is an institutional issue and must be addressed from an institutional perspective, not from a silo. An institutional policy based on recognized best practices sets the foundation for improving the institution's information security posture.
- *Use risk management techniques to identify and respond to information security risks*. There is not enough time or money to eliminate every information security threat. Instead, employ risk management practices to identify and respond to the institution's most pressing information security risks.
- *Implement information security technologies to identify, track, and protect sensitive IT resources and data*. Technology changes quickly, and it is often difficult for higher education institutions to keep up in times of great technological change. However, a careful review of the institution's IT infrastructure to identify and track IT resources containing sensitive data is a must. That review, coupled with a risk assessment, can indicate the best areas in which to introduce information security technologies to effectively enhance the institution's information security posture.

- *Implement an incident response policy*. "If you fail to plan, you are planning to fail."[21] Every hour counts in recognizing, responding to, mitigating, and resolving an information security incident such as a data breach. Create an incident response process with identified roles for information security staff and make sure the plan is regularly tested, reviewed, and updated. Make sure that the plan also ensures that staff tasked with executing the plan are familiar with evidence-collection laws and requirements to ensure adequate law enforcement follow-up (if indicated) after a breach. Test the plan frequently with both IT and non-IT staff to make sure that the process works as expected and that key staff understand their responsibilities during an incident.
- *Train faculty and staff on information security policies and practices*. Ensuring basic levels of end-user information security awareness is critical, especially for those faculty and staff members who handle sensitive institutional data on a daily basis. In order to properly protect institutional data, they need effective information security awareness education to understand the institution's information security policies and know how to recognize potentially dangerous situations that could lead to a data breach.

So how does an institution prevent a data breach? There is no easy or clear answer. Indeed, our limited exploration of this phenomenon suggests that no single measure of prevention is enough by itself to prevent a breach. Furthermore, even if everything that could be done is done, it will not be enough to guarantee absolute security. IT leaders recognize this as well. As the 2017 IT Issues article reported, "Information security is not binary: there is no state of complete security."[22] The 2016 IT Issues article noted that "information security [is] now acknowledged as a field in which 'perfection isn't nearly good enough.'"[23] In both years, higher education IT leaders acknowledged that information security capabilities must be agile and constantly evolve to reduce institutional risk.

Where to Learn More

- EDUCAUSE Cybersecurity Initiative
- EDUCAUSE Core Data Service
- Privacy Rights Clearinghouse Chronology of Data Breaches

Acknowledgments

The authors wish to thank Cathy Hubbs, Chief Information Security Officer, and David Swartz, Vice President and Chief Information Officer, at American University for their encouragement and support of this research.

About the Authors

D. Christopher Brooks is Interim Director of Research for ECAR. Joanna Lyn Grama is Director of Cybersecurity and IT GRC programs for EDUCAUSE.

Citation for This Work

Brooks, D. Christopher, and Joanna Lyn Grama. *Searching for a Smoking Gun, Chasing a Silver Bullet: Data Breaches in Higher Education*. Research bulletin. Louisville, CO: ECAR, February 21, 2017.

Notes

1. Joanna Lyn Grama, *Just in Time Research: Data Breaches in Higher Education* (Louisville, CO: ECAR, May 2014).
2. Ibid.
3. This research is made possible through institutional participation in the EDUCAUSE Core Data Service. Colleges and universities use CDS to inform their IT strategic planning and management activities. Visit the Core Data Service to learn more about the service and participate in data collection.
4. Note that the Privacy Rights Clearinghouse Chronology of Data Breaches may not be exhaustive. It is limited to breaches reported in the United States. For more information on the Chronology of Data Breaches resource, visit the FAQ page.
5. There was an exception to this in our data set. One institution had no data for either the year preceding or the year of the breach but did complete CDS the following year.
6. For more details on the comparative method used here, see Charles C. Ragin, *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*, revised edition (Oakland, CA: University of California Press, 2014).
7. The question asked, "In the prior fiscal year, what was the title of the highest-ranking person with primary responsibility for information security across your institution?" If the response was CISO, then we coded the variable as 1; if the response was other than CISO, we coded the variable 0.
8. The question asked, "What percentage of full time did this person devote to information security?" If the response was 60–100%, then we coded the variable 1; if the response was less than 60%, we coded the variable 0.
9. Joanna Lyn Grama and Leah Lang, *CDS Spotlight: Information Security*, research bulletin (Louisville, CO: ECAR, August 2016). We also explored a measure of IT centralization, but the data for these cases were not conducive to dichotomization.
10. Maturity factors contribute toward a mature capability in institutional information security programs and include organization, policy, and procedural considerations. Deployment factors measure the state of institutional deployment of information security technologies. For more information on EDUCAUSE maturity and deployment indices, visit Higher Education IT Assessment and Benchmarking Projects.
11. According to the classification schema used by the Privacy Rights Clearinghouse.
12. Susan Grajek, *Higher Education's Top 10 Strategic Technologies for 2016*, research report (Louisville, CO: ECAR, January 2016).
13. Ibid.
14. EDUCAUSE, *Foundations of Information Security: Institutional Implications for Safeguarding Data* (Louisville, CO: EDUCAUSE, September 2014).
15. Joanna Lyn Grama and Leah Lang, *CDS Spotlight: Information Security*, research bulletin (Louisville, CO: ECAR, July 3, 2015).
16. Jeffrey Pomerantz, The Higher Education IT Workforce, 2016: The Scope of the CISO Role, forthcoming 2017.
17. *In loco parentis* is a term used to reference "in place of a parent." We took some liberties to create a new term to mean "in place of a CISO."
18. Unlike data loss prevention systems, the lack of training does not appear to be systematically related to the presence of a CISO.
19. EDUCAUSE Core Data Service Almanac, All U.S. Institutions, 2015 data, February 2016.
20. Cathy Bates et al., *Technology in Higher Education: Information Security Leadership*, HEISC/ECAR working group paper (Louisville, CO: ECAR, March 2016).
21. This quote is often attributed to Benjamin Franklin.
22. Susan Grajek and the 2016–2017 EDUCAUSE IT Issues Panel, "Top 10 IT Issues, 2017: Foundations for Student Success," *EDUCAUSE Review* 52, no. 1 (January/February 2017).
23. Susan Grajek and the 2015–2016 EDUCAUSE IT Issues Panel, "Top 10 IT Issues, 2016: Divest, Reinvest, and Differentiate," *EDUCAUSE Review* 51, no. 1 (January/February 2016).