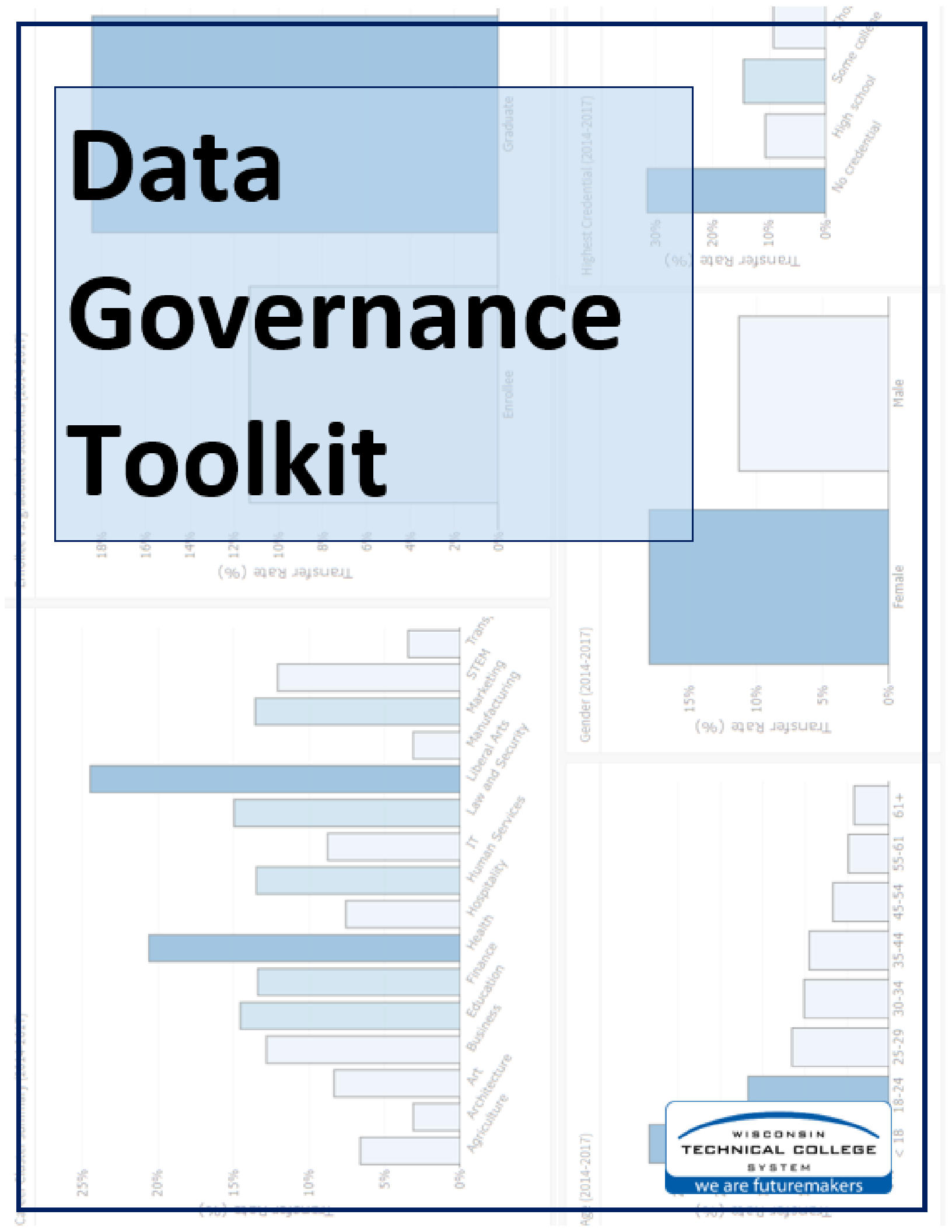
Data Governance Toolkit



WISCONSIN TECHNICAL COLLEGE SYSTEM

we are futuremakers

Data Governance Toolkit

Published April 2019

Wisconsin Technical College System
4622 University Avenue
Madison, WI 53705

Table of Contents

If you have questions regarding the information in this toolkit, contact hilary.barker@wtcsystem.edu

[1] The Case for Data Governance

Data governance is an organization-wide process for effectively managing data and information. This management includes data storage, integration, documentation, quality, security, use, access, ethics, and destruction/retention. Data governance can be an institution's greatest asset or greatest liability.

MORE RESOURCES

Read **What's Your Data Strategy** by DalleMule and Davenport (2017) for more information on data defense and offense.

Read **Hard Costs of a Data Breach: Five moves to make immediately following a data breach and what they'll cost your college** by Negrea (2015).

Assess your college's data breach risks using **IBM's assessment tool**.

The most powerful and effective data governance systems strike a balance between the college's data **defense** (security, protections) and **offense** (analysis, research; DalleMule & Davenport 2017). Without adequate data defense, sensitive records (e.g., social security numbers of students and staff) can be stolen and leaked costing the college millions of dollars. Without effective data offense, a college fails to harness data for continuous improvement, which impedes student success (retention and graduation). This result can then jeopardize the viability of the college with declining student enrollments.

Data Defense:

From 2005 to 2018, there have been 861 reported data breaches in the education sector within the United States (Privacy Rights Clearinghouse, Fig. 1). In total, 65.9 million education records have been breached, of which 176,000 records were breached in Wisconsin education institutions. The average cost of a data breach within the United States education sector was $8.1 million in 2018 (IBM 2018).

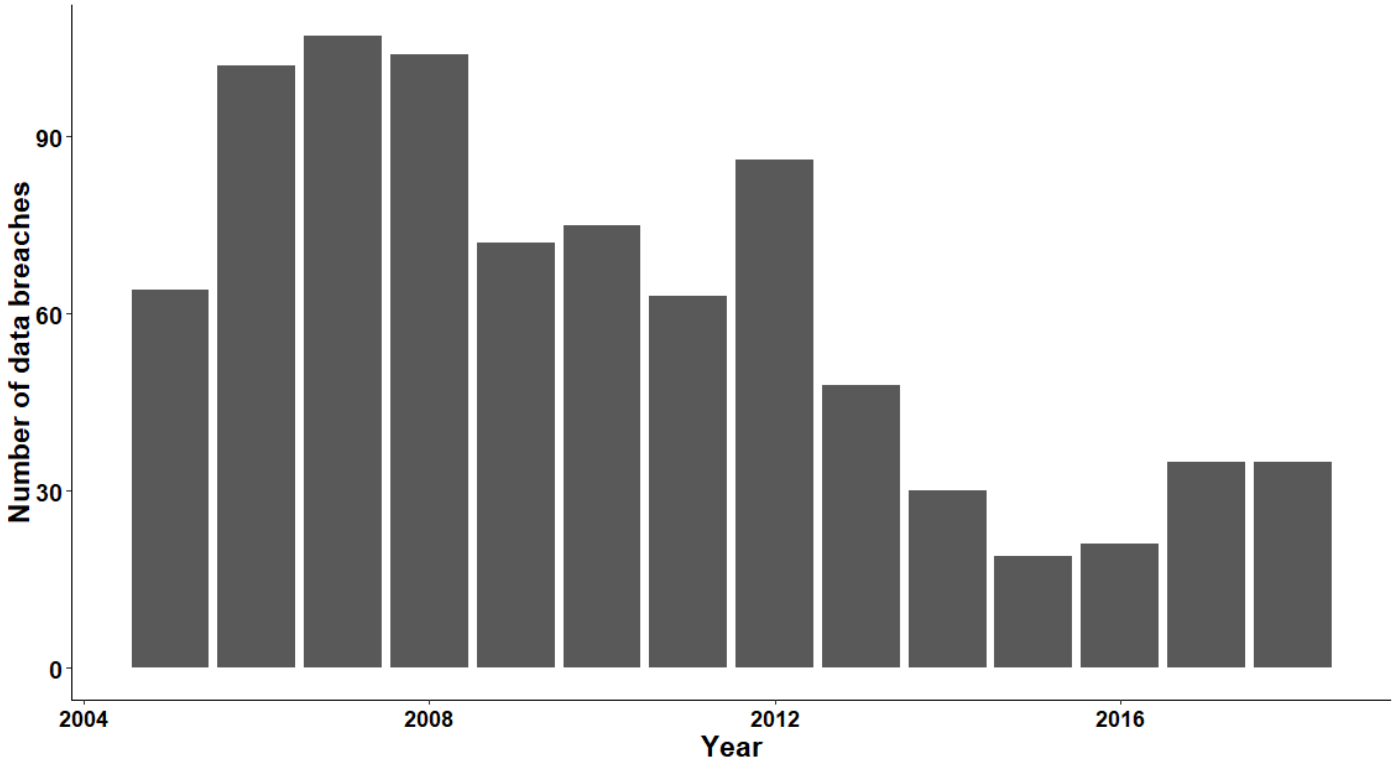


Figure 1. Number of data breaches in the United States education sector from 2005 to 2018. Data derived from Privacy Rights Clearinghouse.

Not only does a data breach cost money, a breach can damage public trust within an institution and the college's image. For instance, previous data breach incidents have resulted in national headlines, lawsuits from students, grievance filings from faculty unions, and multimillion dollar costs (e.g., CBC News 2019, Smith 2019, Sundqvist 2018, AZ Central 2014).

Data Offense:

Data-informed continuous improvement can completely transform a college. In 2008, Odessa College (Texas) had the lowest graduation rate in the nation according to **IPEDS**, and was at risk of being defunded by their state (Phillips & Wood 2018). To turn things around, they looked at course dropout data to better understand what factors helped keep a student enrolled in class and what factors led to students dropping out. With instructor interviews they identified four key practices that led to student success. Effective instructors (1) interacted with their students by name, (2) monitored their students' progress, (3) conducted one-on-one meetings with their students, and (4) allowed flexibility when needed (e.g., extending homework deadlines for students who had a family emergency).

Given these results, Odessa College administrators provided professional development to accelerate the use of these best teaching practices across the college. These changes immediately increased course completion success and in 2014, Odessa College had the highest 3-year graduation rate in Texas. In 2017, Odessa College was awarded the **Aspen Prize** for community college excellence and the college has one of the best graduation rates in the country. This dramatic increase in student performance was due not only to results from research on dropout rates, but also from other data-informed practices (e.g., retention and semester length).

Data governance holds incredible opportunity and incredible risk, depending on implementation and integration across a college. The following chapters in this toolkit can be used as a guide to help set up or improve an existing data governance program. This toolkit is not comprehensive but provides a succinct summary of important topics with links and resources for more information.

[2] Planning for Change

Improving an existing data governance program or creating a new one can require substantial change in policies, procedures, norms, workflows, and even the culture of your college.

Organizing Change:

Research suggests that 70% of organizational change initiatives fail (Ewenstein et al. 2015). Thus, care must be taken in how your team rolls out your college's data governance program so that it will most likely succeed. Here are a few best practices for leading change and resources to learn more:

- Gather engagement at all levels of your college to build a shared vision.
    - Provide options that allow stakeholders to choose their level of engagement (Kislik 2018).
    - To engage and inspire others, start with why (the vision), rather than what (e.g., data security) or how (e.g., data policies; **Simon Sinek's TED talk on 'How great leaders inspire action'**)
- Consider how your college's various stakeholders will respond and feel about this change. What are their concerns? Lead with empathy (Sanchez 2018).
- Small nudges can bring about large change (Tams 2018).
- For a listing of data-informed best practices in organizational change, see 'How the implementation of organizational change is evolving' by Blake Lindsay, Eugene Smit, and Nick Waugh (2018).

MORE RESOURCES

Read **Focusing on Organizational Change**, an open textbook by William Judge and published by Saylor Foundation (2012) for more information on building the capacity for change across your college.

Read about ideas for incorporating technology to help improve the success of your change projects in **Changing change management** by Boris Ewenstein, Wesley Smith, and Ashvin Sologar (2015).

One Step at a Time:

To ensure sustainability, break down the change process into manageable parts. Data Governance 1.0 does not have to be perfect; instead it builds the foundation for Data Governance 2.0 and beyond. This pace and iterative approach will also allow for more engagement across the college (i.e., feedback and input on each step in the Data Governance development) without overwhelming stakeholders with too much change at once. In addition, an iterative approach to data governance mirrors the progression of analytics maturity (Fig. 2). Each college will be at a different stage in analytics and data governance maturity, and that is okay.

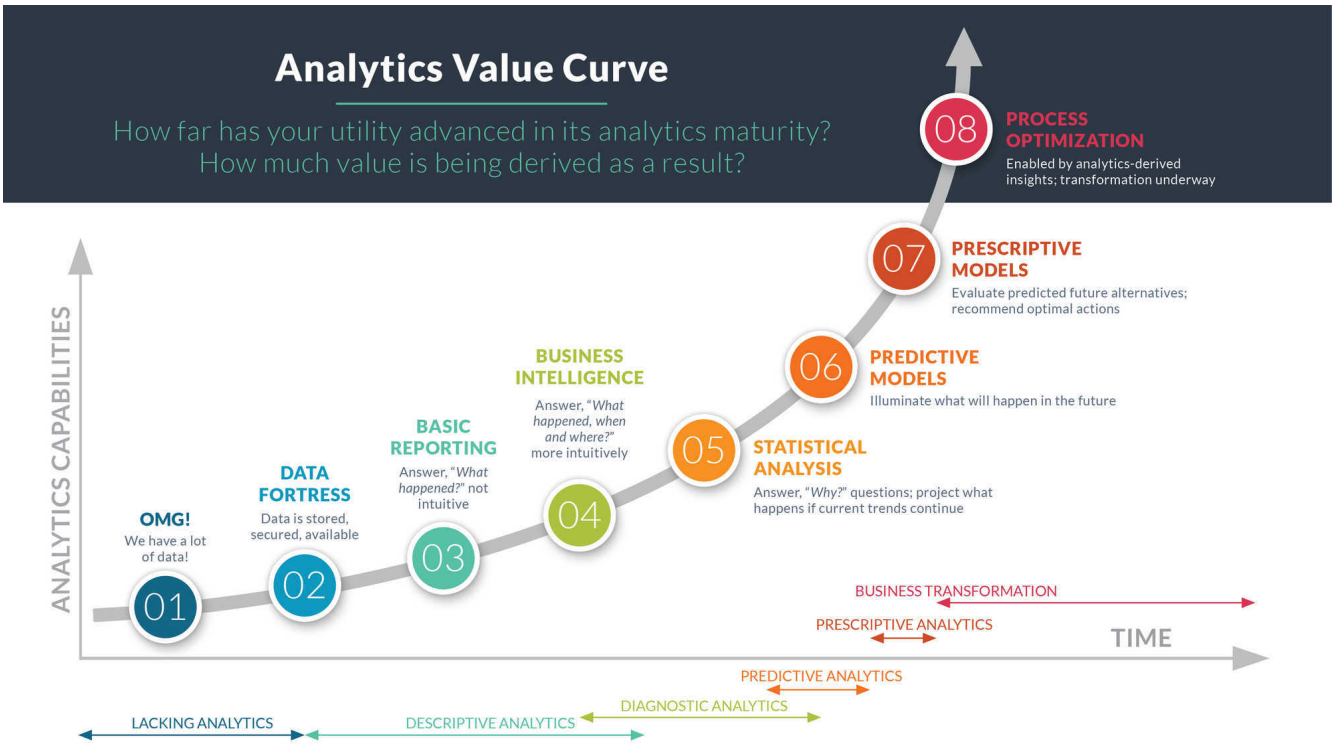


Figure 2. Model for analytics maturity/capabilities. Figure copied from Energy Central blog post.

Build Your Network:

Remember that your college is not unique in needing to start or improve a data governance program. Make sure to leverage your colleagues and resources at other colleges that can help you in this work. Create a community of practice to share ideas and experiences in developing data governance at an institution.

[3] Building a Data Governance Framework

Below are recommendations and guiding questions for you to consider in establishing a data governance framework for your college. Note that this process and resulting framework will need to be tailored to best fit your college.

- ✓ Establish the data governance team (steering committee, data stewards, etc.).
- ✓ Conduct a **data inventory** (elements, technology, documentation, policies, etc.) and needs assessment.
- ✓ Steering Committee sets the data governance goals and vision (yet provide opportunities for input from stakeholders across all levels of your college – see [2] Planning for Change).
- ✓ Create a data governance **charter** (i.e., project management tool).
    - Includes objectives and constraints of the project, stakeholders, defines the scope of the project, timeline, documents potential risks, benefits, and costs
- ✓ Meet with stakeholders.
- ✓ Refine and execute action plan to meet the data governance goals. This may include:
    - Develop/refine the college's data model/map (schematic of data flow)
    - Develop/refine the college's data dictionary
    - Develop/refine data policies and procedures
        - Data inventory cycle
        - Data content management
        - Data quality
        - Data access
        - Data security
        - Data sharing agreements
    - Upgrade/develop data warehouse
- ✓ Meet with stakeholders.
- ✓ Conduct professional development to train college staff in relevant data governance policies and procedures (e.g., data security, data quality, effective use of data).
- ✓ Review data governance goals and assess how your current data governance framework can be improved.
- ✓ Iterate and improve your data governance program with regular touchpoints for reviewing data policies and procedures and meeting with stakeholders.

Guiding Questions:

- Who are your most important stakeholders in data governance across your district?
- How will you measure success within your data governance program? How will this be readily tracked?
- What is working well within your college's current data systems and processes? What needs exist within this system?
- What sources of data exist across your district? How are these data integrated and used?
- How are records created, checked, managed, secured, retained, and destroyed across your college?
- Why are these records collected? Is this justification transparent and communicated to stakeholders?
- Who has access to what data and why?
- How will your college balance data defense and offense within your data governance program?
- How will you communicate and provide transparency about your data governance program across your college?
- How will your data governance team break down your governance goals into iterative and achievable milestones?
- What resources and connections can you leverage to accelerate your data governance program?

MORE RESOURCES

Project management resources to help keep your work focused and effective:

- Collaborative/team platforms (*e.g.,* Microsoft Teams, Slack, Asana, etc.)
- **Project Management** open textbook by Adrienne Watt (2014)

[4] Building a Data Governance Team

A data governance team includes a steering committee, data owners, data stewards, data custodians, data quality analysts, and data architects (Table 1). Successful data governance teams are cross-functional, include college leadership, and build collaboration between Institutional Research and Information Technology.

Steering Committee:

This committee should be comprised of key data stakeholders, including college leadership. The committee will help set the goals and direction for the college's data governance framework, which the data stewards will follow.

Data Owners:

Each data owner is responsible for one data domain (e.g., Client data, curriculum, financial, facilities), and they report to the steering committee.

Data stewards:

Data stewards are the most involved in designing and implementing the data governance framework. In this way, their role is similar to a project manager. Part of their day-to-day job responsibilities should allow for working on the college's data governance. Data stewards are accountable to the data owners and steering committee.

Data Custodians:

A data custodian manages the systems, technology, and infrastructure for data warehousing, storage, and security. They are in a decision-making position (e.g., Director of IT), interact with data stewards, and report to the steering committee.

Data Quality Analyst:

A data quality analyst is tasked with assessing and tracking data quality. They discover data errors and inconsistencies and conduct root-cause analysis to uncover problems with the data management lifecycle.

MORE RESOURCES

See an example **Data Governance Council Roles and Responsibilities** guidelines from the California State University Channel Islands.

See Educause's article on **Speaking the same language: Building a data governance program for institutional impact** (Chapple, 2013) for helpful information, including the use of a RACI (Responsible, Accountable, Consulted, Informed) matrix to organize stakeholders in data governance at a college.

Priority	Steering Committee	Data Owners	Data Stewards	Data Custodians	Data Quality Analysts	Data Architects	Stakeholders
Data Governance Vision and Goals	Accountable	Informed	Informed	Informed	Informed	Informed	Informed
	Responsible	Consulted	Consulted	Consulted			
Data Governance Framework & Policies	Informed	Accountable	Accountable	Consulted	Consulted	Consulted	Informed
	Consulted	Responsible	Responsible				Consulted
Data Quality	Informed	Accountable	Accountable		Accountable	Consulted	Informed
	Consulted	Responsible	Responsible		Responsible		Consulted
Data Ethics	Informed	Accountable	Accountable	Accountable	Accountable	Accountable	Accountable
	Consulted	Consulted	Responsible	Consulted	Consulted	Consulted	Consulted
Data Security	Informed	Accountable	Accountable	Accountable		Consulted	Informed
	Consulted	Responsible	Responsible	Responsible		Responsible	Consulted
Data Warehousing, Storage & Integration	Informed	Accountable	Accountable	Accountable		Consulted	Informed
	Consulted	Responsible	Responsible	Responsible		Responsible	Consulted
Communication & Training	Informed	Consulted	Accountable	Accountable	Consulted	Consulted	Informed
	Consulted		Responsible	Responsible			Consulted

Table 1. Data governance team roles and responsibilities.

Data Architects:

A data architect is responsible for building, maintaining, and upgrading the systems, technology, and infrastructure for data warehousing, storage, and security. They report to the data custodians and interact with the data stewards and quality analysts.

[5] Data Warehousing

Data warehousing can help integrate and connect data across a college, thereby eliminating data silos which can be vulnerable to security threats and hinder data analysis for continuous improvement. In addition, data warehousing can help make your college less dependent upon various software applications (e.g., learning management systems, human resources information systems).

A data warehouse can store information from various parts of your college (marketing, academics, finance, etc.) in one repository for data access and analysis (Fig. 3). This system can break down data silos across your district and make data more readily available for data offense and maintained and stored in a way that supports data defense. Several kinds of data warehouse platforms are available. Thus, your data governance team will need to identify which platform is the best fit for your college.

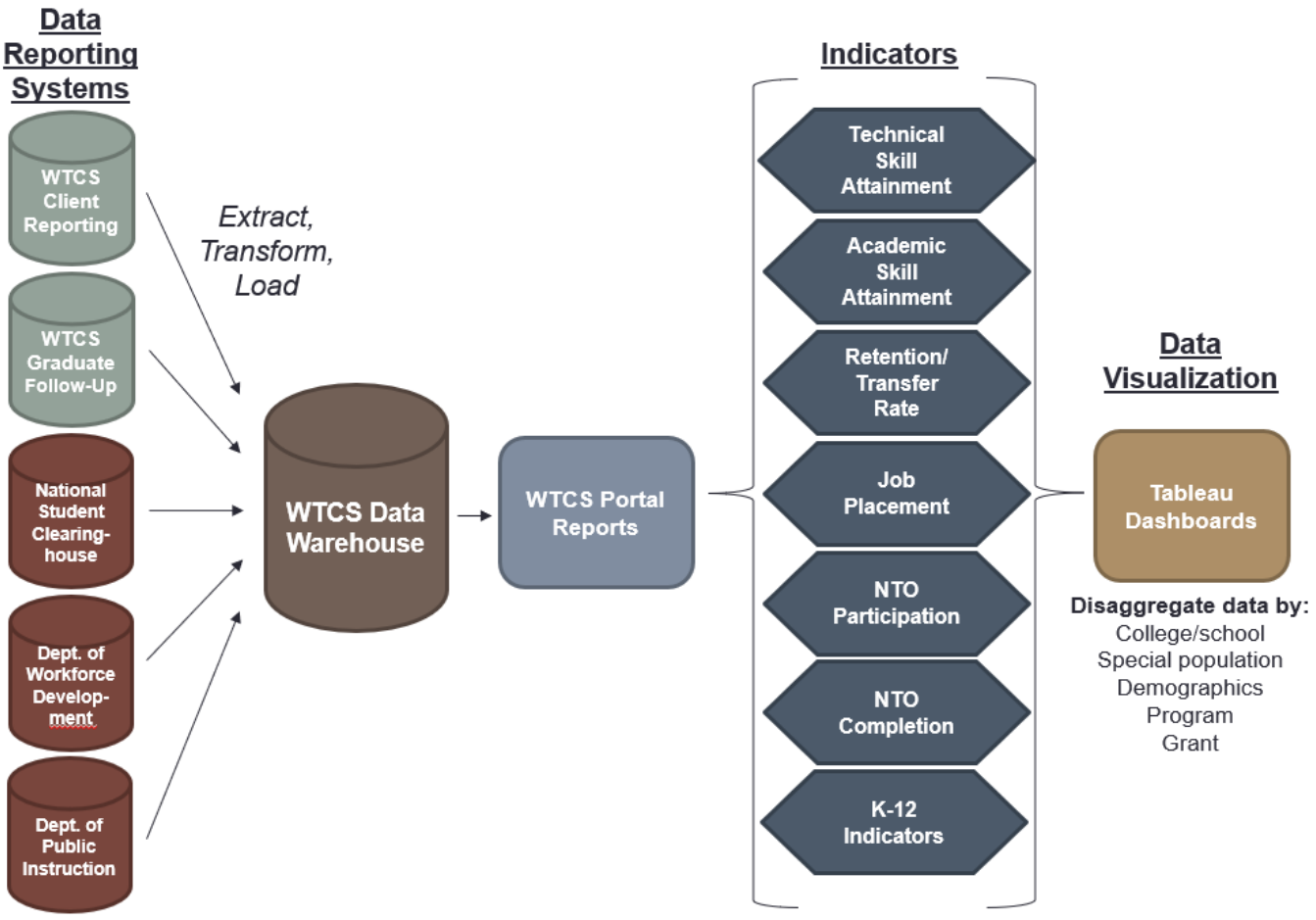


Figure 3. Example WTCS data workflow for Perkins IV, which highlights the data sources, staging (extract, transform, load), the data warehouse, and resulting reporting and analysis for continuous improvement.

Selecting a Data Warehouse Platform:

If your college is starting from scratch, first set a data warehouse goal that best fits your college's needs. For instance, do you want one enterprise data warehouse for all college data or would a 'best of breed' model (selecting the best business solution for each department/division's data needs) work? Once you have a clear vision of your college's data warehouse goals, then build your implementation plan. For this plan, consider starting with small wins (i.e., integrating a few databases together to create a departmental data mart), which then build up to a more comprehensive data warehouse platform (**Mullins 2016**). For information on selection criteria, evaluating your data warehouse needs, and a summary of top data warehouse vendors, visit:

- TechTarget's **buyer's guide** to selecting the best data warehouse product (2016)
- Xplenty's blog post on 'What to consider when selecting a data warehouse for your business' (2017)

MORE RESOURCES

Read **Breaking Down Data Silos** by Edd Wilder-James (2016) for information on the impact of data silos and strategies for data integration.

Read **Supporting Analytics through Data Integration and Governance** by Eckles, Gill and Riley (2017) for perspectives from college IT leaders on their data integration strategies and experiences.

Consider becoming a member of the **Higher Education Data Warehousing Forum** for information on data warehousing and governance and a chance to connect with and learn from colleagues at other institutions. Membership is free.

If you are considering a cloud-based data warehouse as a service (DWaaS) vendor, see the Department of Education's **FAQ document about cloud computing**. This document addresses FERPA issues and recommended best practices for cloud computing. Also, learn more about the costs of cloud computing in higher education to help discern whether this option will fit your college's needs (**McKenzie 2018**).

Data Storage and Integration:

While data warehouse platforms can provide data integration (e.g., extract, transform, and load) in addition to data storage, other data integration tools are available and may be more effective for your college's needs. For perspectives from college IT leaders on their experiences and strategies for data integration, read Integrating Data and Systems to Support Next-Generation Enterprise IT (Berman et al. 2017) and Addressing New Challenges of Data Ingestion (Flerlage 2018). Also, visit Data Integration Info for more information and guidance for choosing a data integration vendor.

[6] Data Quality

To ensure data quality and accuracy, your college should implement a continuous improvement cycle (Fig. 4). Your college's Data Quality Analyst(s) should be at the forefront of this work.

Data Dictionary:

Data quality starts with clear and transparent data definitions (also called business rules). A data dictionary clearly lays out and organizes these definitions to minimize confusion and inconsistencies. When creating or modifying a data dictionary, Data Stewards should meet with stakeholders across the college to collect data terms and identify and remove inconsistencies. Once a drafted data dictionary is available, key stakeholders should be consulted with and sign off on the finished product. This dictionary should then be published and readily available for work across your college. Within the data quality cycle (Fig. 4), the data dictionary should be updated and maintained as needed.

For more information on data dictionaries, explore:

- Carl Anderson's **Data Dictionary: a how to and best practices** (2015)
- **Common Education Data Standards** from the Department of Education

Also, check out **WTCS's Continuous Improvement Data Library**.

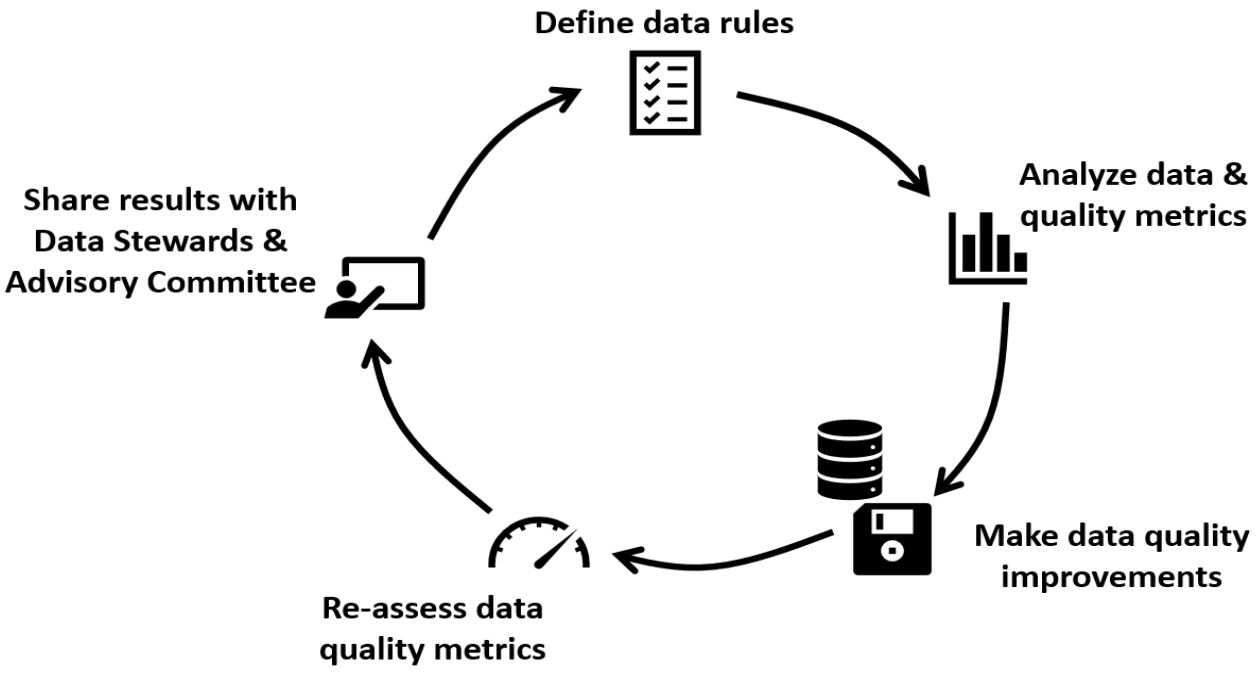


Figure 4. Continuous improvement cycle for data quality (adapted from **Templar 2017**).

Data Quality Best Practices:

Monitor data quality metrics (listed below) in an interactive dashboard to identify data elements that need improvement (Fig. 2). Information for data quality metrics can be collected from regular data audits. To make this information most useful, make sure to audit data at various stages within the data management lifecycle (e.g., data entry/creation, processing, storage, joining, transformation, analysis, destruction/removal).

Data Quality Metrics from **Templar 2017**:

- *Conformed*: does the record follow data standards/structures (e.g., MM/DD/YYYY for a date)
- *Valid*: does the information make sense (e.g., is the year within an expected range, 2018 vs 2108)
- *Complete*: there are no missing values within a record
- *Accurate*: the data have been checked (manually or against an approved source) and are correct
- *Consistent*: is the information the same across systems, datasets?
- *Unique*: if expected, is the information unique and unduplicated within a dataset
- *Available*: is the information accessible to those who need it (according to data access criteria – who has access to what information and when?)
- *Timely*: were the data entered and stored in a reasonable timeframe so that they can be used in decision-making and analysis?
- *Current*: does the information reflect what is happening now at your college?

MORE RESOURCES

See the National Cooperative Education Statistics (NCES) **Forum Guide to Education Indicators** (2005) for data definitions and best practices. Also check out their **Forum Guide to Metadata: The Meaning Behind Education Data** for information on metadata and how these records can be used to improve data quality.

For criteria for selecting a data lineage tool, see Loshin's report, **How Data Lineage Tools Boost Data Governance Policies** (2018).

Data lineage tools can be helpful in tracking your data and its metadata and any potential issues (e.g., security breaches, errors within the data) throughout the data lifecycle (Loshin 2018).

Within data processing, transformation, and analysis steps, follow these best practices:

- *Code review*: implement checks within data extraction/processing/analysis steps by having at least one other programmer/analyst review the code, provide feedback, and testing
- *User testing*: once code has been reviewed, include another testing step. For instance, in a data extraction, have a person who is knowledgeable within the area (e.g., general education) test and check the information to ensure that the data are reasonable, and no red flags are apparent
- For all steps of the data management lifecycle, provide clear, consistent, and detailed documentation (e.g., comments within code) that describe each step
- To maintain consistency, adapt and follow *code style guides* (e.g., Google style guides) and *statistical standards* (NCES Statistical Standards) across your college

[7] Data Ethics

To ensure that your use of data is ethical, adapt a code of data ethics across your college (see below) that enhances transparency, protects sensitive and identifiable personal information, and strives to use data for continuous improvement.

Transparency:

Communicate with students and college stakeholders about what information is collected, what this information is used for and why, and to whom this information is presented. For information and best practices for education data transparency, read the Department of Education's Privacy and Technical Assistance Center's (PTAC) **report on transparency**.

Code of Data Ethics (NCES 2010)

"Integrity

- Demonstrate honesty, integrity, and professionalism at all times.
- Appreciate that, while data may represent attributes of real people, *they do not describe the whole person*.
- Be aware of applicable statutes, regulations, practices, and ethical standards governing data collection and reporting.
- Report information accurately and without bias.
- Be accountable, and hold others accountable, for ethical use of data.

Data Quality

- Promote data quality by adhering to best practices and operating standards.
- Provide all relevant data, definitions, and documentation to promote comprehensive understanding and accurate analysis when releasing information.

Security

- Treat data systems as valuable organizational assets.
- Safeguard sensitive data to guarantee privacy and confidentiality."

Check out the National Cooperative Education Statistics (NCES) Forum Guide to Data Ethics (2010) for more information and recommended practices to help implement this code across your college.

Identifiable Information:

When sharing or presenting student data, always make sure that the information is de-identified (i.e., cannot be linked to an individual). Information that can be used to identify a student includes student name, telephone number, student ID, Client ID, email address, birth date, etc. In addition, when displaying disaggregated student data, be careful to suppress information for small subgroups (i.e., fewer than 10 students) since this information could still be used to identify a student. To better understand what information is both directly and indirectly identifiable, check out UW-Madison's guide to '**identifiability**'. Also, read PTAC's brief on **Statistical Methods for Reporting Personally Identifiable Information in Aggregate Reporting**.

MORE RESOURCES

Read **Setting the Table: Responsible Use of Student Data in Higher Education** by Kurzweil and Stevens (2018) for recommendations.

Read Virginia Eubank's book on **Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor** (2018) for information on the use of predictive analytics and how this use can lead to negative unintended consequences.

For more information on predictive analytics in higher education as well as recommendations and best practices, read New America's guides:

- **The Promise and Peril of Predictive Analytics in Higher Education** by Ekowo and Palmer (2016)
- **Predictive Analytics in Higher Education** by Ekowo and Palmer (2016)

Predictive Analytics:

Predictive tools in education can be incredibly useful for personalizing education, informing continuous improvement strategies, and targeting student support services so that services reach students who need support when they need it. Yet, predictive tools can also create serious problems (e.g., 'retention plan' scandal Jaschik 2016). Also, remember that each college's analytics maturity will be at different stages (see Fig. 2 on pg. 4); you are not expected to engage in predictive analytics.

When developing a predictive analytics approach, make sure to consider:

- The data source – Is a vendor using data from other colleges and/or outdated data to create a prediction? Are the data accurate and of high quality? If using a vendor with your college's data, who own's the data (answer – this should be your college)? Can the vendor keep your college's data after the contract/services expire (answer – no)?
- Transparency – Is the prediction algorithm transparent? What variables are feeding into the model?
- The prediction outcome – What is the algorithm truly predicting? Is this the outcome that your college is interested in or a proxy? Is this transparent? Is the proxy effective?
- Predictor variables – What variables are used to predict the outcome? Are any of these variables demographic information (e.g., gender, race/ethnicity, socioeconomics)? If so, could the model reinforce inequalities across student groups and discriminate against students?
- Limitations – What are the limitations of the prediction model? Are these limitations clearly stated and communicated to those who use it?
- Training and use – How is the model being used within your college? How are college staff trained to use the prediction results?

[8] Securing Data

Data defense protects sensitive and personally identifiable information from security threats and vulnerabilities (Fig. 5).

Security Measures

Security Threats:

In 2018, the most common security threats and attacks in the education sector included phishing scams, attacks through web-applications, and user errors (e.g., sending information to the wrong person, databases that are not secure; **Verizon 2018**). These breaches mostly (72%) targeted personal data for financial gain (e.g., selling on black markets, applying for loans and credit cards, blackmail, filing tax returns; **Verizon 2018**). Cybercriminals can hack into and compromise a company's data systems in a matter of minutes, yet it typically takes several months for the company to discover these security breaches (**Verizon 2018**).

Multi-layered Security Systems:

To best protect your college's information from these threats, you will need to deploy a multi-layered security system with comprehensive and regular training for college students and staff. This multilayered security system should include:

- Physical security (e.g., locked server rooms)
- Spam blocker to catch phishing scams
- Anti-virus and anti-malware

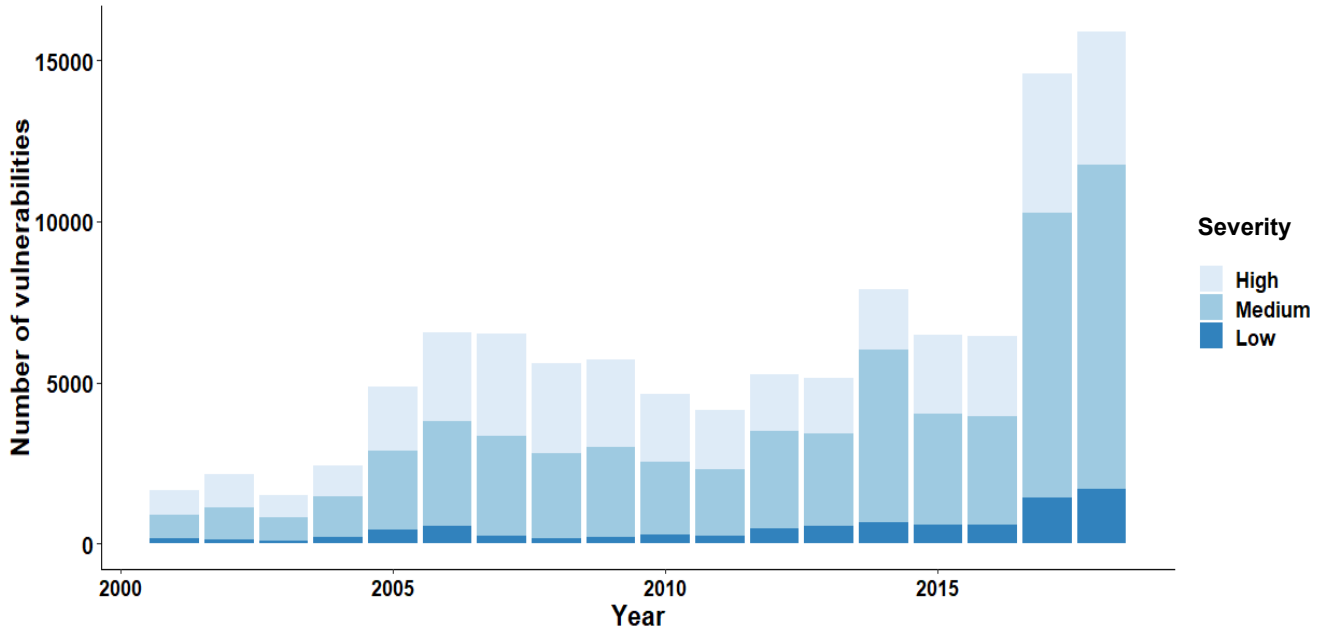


Figure 5. Number of software and hardware security vulnerabilities from 2000 to 2018. Data are derived from NIST.

MORE RESOURCES

The Department of Education's Privacy Technical Assistance Center (PTAC) has a **Data Security Checklist** along with other helpful **resources**. PTAC can also provide **technical assistance and guidance and visit** your college to audit your security systems. For resources and information on protecting ***online*** student data, visit PTAC's website on **Education Technology** (cloud computing FAQs, model terms of service, etc.).

Visit the National Institute of Standards and Technology's (NIST) **Computer Security Resource Center** for information on security and privacy, technology, applications, laws and regulations, and more. Also check out their **Vulnerability Database** for up to date information on security risks (*e.g.,* buffer errors, information leaks).

Check out the **National Cybersecurity and Communications Integration Center** (NCCIC) website for information and publications on computer and internet security.

- Firewall protection
- Vulnerability scanners
- Intrusion detection systems
- Two-factor authentication with strong passwords
- Encryption
- Regular data backups
- Patch management

(**Department of Homeland Security NCCIC 2016**; **Department of Education PTAC 2015**)

Security Best Practices:

- Provide clear and transparent **policies for data users** that outlines prohibited and allowed uses of data with relevant information on federal (*e.g.,* FERPA), state, and local student data laws.
- Administer regular security training for students and staff that is contextualized and relevant to the individual's role, habits, and needs. For recommendations, see PTAC's **Data Security and Management Training: Best Practice Considerations.**
- Deploy regular **security audits** and data inventories to ensure that all of your college's data are protected.
- Conduct regular drills to practice responding to various data breach scenarios (**PTAC's Data Breach Response Checklist**)

Responding to a Data Breach

PTAC provides information and training for responding to a data breach (see their **checklist** and **training kit** for more information). Effectively responding to a data breach includes (**PTAC**):

- Developing a comprehensive data breach plan that designates staff responsibilities, documents the complete data breach process (from reporting the breach to reviewing the data breach response for continuous improvement), and describes specific tasks within the process with step by step instructions
- Broadly communicating, sharing, and testing/practicing the data breach plan
- Assembling a response team, which may include staff from information technology, institutional research, legal counsel, marketing/communications, finance, and human resources
- Notifying staff (e.g., data owners), agencies (e.g., Family Policy Compliance Office and PTAC), and law enforcement as needed

- Collecting data breach evidence in a manner that is well documented, secure, and could be used in a court of law
- Clearly communicating information to individuals whose data were compromised and providing assistance as needed (e.g. credit monitoring)
- Reviewing the data breach response process to identify areas for improvement and minimize future risks

Also, see information on **Wisconsin's Data Breach Notification Law** (e.g., what information is covered, 45 day timeframe to give notice of a breach).

Data Retention and Destruction

Data Retention Policies:

Your college's data policies should outline data retention rules, which describe how long different types of data should be retained for and who is responsible for tracking data retention and destruction. For instance, the WTCS state agency policies for Client data is to retain this information for seven years before destroying it.

Data Destruction Best Practices:

"Data destruction is the process of removing information in a way that renders it unreadable (for paper records) or irretrievable (for digital records)" (**PTAC Best Practices for Data Destruction**). Data destruction best practices from the National Institute of Standards and Technology's (NIST) include:

- Overwriting data with non-sensitive information
- Degaussing (demagnetizing) to erase information on magnetic media (e.g., hard drives, zip drives)
- Physical destruction via shredding, grinding, incineration, or melting

For more information, see:

- **NIST's Guidelines for Media Sanitation**
- **PTAC's Best Practices for Data Destruction** and **Cornell's Disk and File Erasure** resources and recommendations

Data Access

Your college's data policies should clearly document data access (Who has access to which datasets and why?) along with clear guidelines for data users (What can the data be used for? What uses are prohibited?). The data governance team should review and update data access policies and guidelines regularly.

Data-Sharing Agreements:

Data-sharing agreements are formal contracts with outside organizations and/or researchers for sharing information for intended uses. A data-sharing agreement includes (**HRIS Creating a Data-Sharing Agreement**, **PTAC**):

- *Data description* – What information will be shared and why?
- *Intended use of data* – What is the permitted use(s) of the data? What is the purpose of the data sharing agreement? How will any personally identifiable information be used?

- *Data constraints* – Can the receiver of the data share the information, and/or resulting reports and/or research findings? Who owns the resulting reports/research findings?
- *Data confidentiality* – If personally identifiable or sensitive information is included, how will this be secured and protected? How will this information be reported? (e.g., are minimum sub-group sizes stipulated?)
- *Period of agreement* – How long can the data be retained? What will happen once the period is over? Will the data be returned and/or destroyed?
- *Data security* – How will the data be secured? How many copies of the data will exist? Who will access the data? How will the data be destroyed and who is responsible for this? (See **PTAC's Best Practices for Data Destruction** for recommendations; e.g., data destruction certification form). How will your college audit/monitor the data sharing agreement? If the agreement is not fully met, how will this be managed? What accountability measures (e.g., penalties, right to sue for data breaches) can be put in place? Is the recipient (and their employees) adequately trained in FERPA and data security? What will happen in the case of a data breach?
- *Methods of data sharing* – How will the data be transferred? Will the data be encrypted for this transfer? Is the connection secure?
- *Financial costs* – What are the expenses of this data sharing agreement (transfer, security, etc.)? Who will cover these costs?

MORE RESOURCES

For information on the use of student financial aid data for program evaluation and research, see **PTAC's guidance report**.

Visit **FERPA SHERPA: The Education Privacy Resources Center** for information on state and federal laws that influence student privacy in higher education. Also, watch **It's Not Just FERPA: Privacy and Security Issues in Higher Education** (Baker Donelson, 2015).

For more recommendations, see **PTAC's Guidance for Reasonable Methods and Written Agreements**.

Learning Management and Analytics Software:

When entering into a contract with a learning management system (LMS), analytics software, or education technology company make sure to clearly define data use and control. Outline who owns the data (this should be your college), how the data will be protected, used and destroyed (similar to a data sharing agreement). For more information, see **PTAC's report on Protecting Student Privacy While Using Online Educational Services: Requirements and Best Practices**.

Legislation

Family Educational Rights and Privacy Act (1974):
FERPA protects the privacy of student records and allows students and parents (for students who are under 18 years old) the right to request and review their (or their child's) records and request that the school/college correct these records if they are incorrect. Under FERPA, schools/colleges must have written consent from the student or parent (for students who are under 18 years old) for disclosing the student's records, yet exceptions apply. These exceptions include records that are disclosed to:

- "School officials with legitimate educational interest;
- Other schools to which a student is transferring;
- Specified officials for audit or evaluation purposes;

- Appropriate parties in connection with financial aid to a student;
- Accrediting organizations;
- To comply with a judicial order or lawfully issued subpoena;
- Appropriate officials in cases of health and safety emergencies;
- State and local authorities, within a juvenile justice system, pursuant to specific State law" (from **U.S. Department of Education**)

Schools/colleges may also disclose 'directory' information about a student (e.g., honor roll, address, name, date of attendance) as long as the school has notified students and parents about what constitutes directory information (e.g., letter, student handbook, etc.) and the student (or parent if the student is under 18 years old) has not opted out of directory information disclosures.

For more information, visit **PTAC's FERPA Frequently Asked Questions** webpage.

Health Insurance Portability and Accountability Act (1996):

If your college provides health care to students, then these records are likely subject to FERPA as either an 'education record' or 'treatment record.' If your college's health clinic also serves non-students, then these records will be subject to HIPAA's guidelines for security and privacy. For more information, read the Joint Guidance on the Application of FERPA and HIPAA to Student Health Records (U.S. Departments of Education and Health and Human Services).

For more information on legislation or standards that regulate college data, see Table 2.

Data types	Department/Office(s)	Legislation/Standards
Educational records (e.g., grades, application, Client data)	College-wide	FERPA
Student treatment records	Health clinic	FERPA
Non-student treatment records	Health clinic	HIPAA
Employee health insurance information	Human resources	HIPAA
Credit card transactions	Bookstore, registrar's office, dining halls, etc.	Payment Card Industry Data Security Standards (PCI-DSS)
Employee records	Human resources	Family and Medical Leave Act; Americans with Disabilities Act
Student and/or employee mental health records	Mental health/substance abuse clinic	42 C.F.R. Part 2
Personal records/data (students and employees)	College-wide	Privacy Act; Wisconsin data security, disposal and breach laws (**134.97, 134.98, 134.99, 134.74**)
Information about employees and potential employees	Human resources	Fair Credit Reporting Act & Fair and Accurate Credit Transactions Act
Emails to prospective students	Marketing, registrar's office	CAN-SPAM
College websites	Marketing	**State online privacy policies** (e.g., CA Bus. & Prof. Code 22575-22578, DE Code. Title 6 205C)

Table 2. Relevant legislation or standards that regulate college data. This information is derived from a presentation from **Baker Donelson (2015**).

[9] References

Anderson, C. 2018. Data dictionary: A how to and best practices. Medium. Accessed online: https://medium.com/@leapingllamas/data-dictionary-a-how-to-and-best-practices-a09a685dcd61

Berman, M., P. Bedi, K. Willey, M. Mathews, & K. Reid-Martinez. 2017. Integrating data and systems to support next-generational enterprise IT. EDUCAUSE Review. Accessed online: https://er.educause.edu/articles/2017/6/integrating-data-and-systems-to-support-next-generation-enterprise-it

Bureau of Consumer Protection. Wisconsin's data breach notification law. Wisconsin Department of Agriculture, Trade and Consumer Protection. Madison, WI: Bureau of Consumer Protection. Accessed online: https://datcp.wi.gov/Documents/IDTheftDataBreach607.pdf

California Department of Technology. 2017. The Agile project charter. Accessed online: https://projectresources.cdt.ca.gov/agile/the-agile-project-charter/

CBC News. 2019. Algonquin College faculty union files grievance over data breach. Accessed online: https://www.cbc.ca/news/canada/ottawa/algonquin-college-faculty-phishing-grievance-1.5046865

Chapple, M. 2013. Speaking the same language: Building a data governance program for institutional impact. EDUCAUSE Review. Accessed online: https://er.educause.edu/articles/2013/12/speaking-the-same-language-building-a-data-governance-program-for-institutional-impact

Chestler, A. & E. Setterlund. 2015. It's not just FERPA. Baker Donelson. Accessed online: https://youtu.be/JrsClBPwHF0

Common Education Data Standards. Department of Education. Accessed online: https://ceds.ed.gov/

DalleMule, L. & T.H. Davenport. 2017. What's your data strategy? Harvard Business Review. Accessed online: https://hbr.org/2017/05/whats-your-data-strategy

Data breaches. Privacy Rights Clearinghouse. Accessed online: https://www.privacyrights.org/data-breaches

Data governance council roles and responsibilities. California State University Channel Islands. Accessed online: https://www.csuci.edu/ir/dgc-documents/governing-documents/dgc-roles-and-responsibilities-7-8-16.pdf

Eckles, J., C. Gill, & M. Riley. 2017. Supporting analytics through data integration and governance. EDUCAUSE Review. Accessed online: https://er.educause.edu/articles/2017/12/supporting-analytics-through-partnerships-and-governance

Ekowo, M. & I. Palmer. 2016. Predictive analytics in higher education. New America. Accessed online: https://www.newamerica.org/education-policy/reports/predictive-analytics-in-higher-education/

Ekowo, M. & I. Palmer. 2016. The promise and peril of predictive analytics in higher education. New America. Accessed online: https://na-production.s3.amazonaws.com/documents/Promise-and-Peril_4.pdf

Eubanks, V. 2018. Automating Inequality: How high-tech tools profile, police, and punish the poor. New York City, NY: St. Martin's Press. Print.

Examining the 2018 Cost of a Data Breach. IBM Security. Accessed online: https://databreachcalculator.mybluemix.net/thankyou/explore/

Ewenstein, B., W. Smith, & A. Sologar. 2015. Changing change management. McKinsey & Company. Accessed online: https://www.mckinsey.com/featured-insights/leadership/changing-change-management

Faller, M.B. 2014. Maricopa County colleges computer hack cost tops $26M. AZCentral. Accessed online: https://www.azcentral.com/story/news/local/phoenix/2014/12/17/costs-repair-massive-mcccd-computer-hack-top-million/20539491/

Family Policy Compliance Office. 2018. Family Educational Rights and Privacy Act (FERPA). U.S. Department of Education. Accessed online: https://www2.ed.gov/policy/gen/guid/fpco/ferpa/index.html

Family Policy Compliance Office. 2008. Joint guidance on the application of the Family Educational Rights and Privacy Act (FERPA) and the Health Insurance Portability and Accountability Act of 1996 (HIPAA) to student health records. U.S. Department of Education and U.S. Department of Health and Human Services. Accessed online: https://www2.ed.gov/policy/gen/guid/fpco/doc/ferpa-hipaa-guidance.pdf

FERPA SHERPA. 2019. The education privacy resource center for higher ed. Accessed online: https://ferpasherpa.org/higher-ed/

Flerlage, K. 2018. Addressing new challenges of data ingestion. EDUCAUSE Review. Accessed online: https://er.educause.edu/blogs/2018/12/addressing-new-challenges-of-data-ingestion

HRIS Strengthening Implementation Toolkit. 2011. Creating a data-sharing agreement. Accessed online: https://www.ihris.org/toolkit/tools/data-sharing.html

IT @ Cornell. 2017. Disk and file erasure. Cornell University. Accessed online: https://it.cornell.edu/security-essentials-it-professionals/disk-and-file-erasure

Jaschik, S. 2016. The questions developed to cull students. Inside Higher Ed. Accessed online: https://www.insidehighered.com/news/2016/02/12/questions-raised-about-survey-mount-st-marys-gave-freshmen-identify-possible-risk

Judge, W. 2012. Focusing on Organizational Change. Saylor Academy. Accessed online: https://open.umn.edu/opentextbooks/textbooks/focusing-on-organizational-change

Kissel, R., M. Scholl, S. Skolochenko, X. Li. 2006. Guidelines for media sanitation. National Institute of Standards and Technology Special Publication 800-88. Accessed online: https://ws680.nist.gov/publication/get_pdf.cfm?pub_id=50819

Kurzweil, M. & M. Stevens. 2018. Setting the table: Responsible use of student data in higher education. EDUCAUSE Review. Accessed online: https://er.educause.edu/articles/2018/5/setting-the-table-responsible-use-of-student-data-in-higher-education

Lindsay, B, E. Smit & N. Waugh. 2018. How the implementation of organizational change is evolving. McKinsey & Company. Accessed online: https://www.mckinsey.com/business-functions/mckinsey-implementation/our-insights/how-the-implementation-of-organizational-change-is-evolving

Loshin, D. 2018. How data lineage tools boost data governance policies. TechTarget. Accessed online: https://searchdatamanagement.techtarget.com/tip/How-data-lineage-tools-boost-data-governance-policies

McKenzie, L. 2018. Reaching for the cloud. Inside Higher Ed. Accessed online: https://www.insidehighered.com/news/2018/07/05/cost-concerns-keep-cloud-services-out-reach-many-small-colleges

Mullins, C. 2016. Evaluating your need for a data warehouse platform. TechTarget. Accessed online: https://searchdatamanagement.techtarget.com/feature/Evaluating-your-need-for-a-data-warehouse-platform

National Cybersecurity and Communications Integration Center. 2016. Recommended practice: Improving industrial control system cybersecurity with defense-in-depth strategies. U.S. Department of Homeland Security. Washington, DC: National Cybersecurity and Communications Integration Center. Accessed online: https://ics-cert.us-cert.gov/sites/default/files/recommended_practices/NCCIC_ICS-CERT_Defense_in_Depth_2016_S508C.pdf

National Forum on Education Statistics. 2005. Forum guide to Education Indicators (NFES 2005-802). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Accessed online: https://nces.ed.gov/pubs2005/2005802.pdf

National Forum on Education Statistics. 2009. Forum guide to metadata: The meaning behind education data (NFES 2009-805). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Accessed online: https://nces.ed.gov/pubs2009/2009805.pdf

National Forum on Education Statistics. 2010. Forum guide to data ethics (NFES 2010-801). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Accessed online: https://nces.ed.gov/pubs2010/2010801.pdf

National Institute of Standards and Technology. 2018. CVSS severity distribution over time. Accessed online: https://nvd.nist.gov/general/visualizations/vulnerability-visualizations/cvss-severity-distribution-over-time

Negrea, S. 2015. Hard costs of a data breach. University Business. Accessed online: https://www.universitybusiness.com/article/hard-costs-data-breach

P., C. 2017. IRB guidance: Identifiability. University of Wisconsin – Madison. Accessed online: https://kb.wisc.edu/sbsedirbs/page.php?id=76643

Phillips, B.C. & D. Wood. 2018. It's better to lead than lag: How leading indicators can drive institutional change. Achieving the Dream. 2018 Data and Analytics Summit. Accessed online: https://www.palmbeachstate.edu/ire/StrategicPlanning/achieving-the-dream/documents/leading_and_lagging_for_AtD_Data_Summit.pdf

Privacy Technical Assistance Center. 2012. Data breach response checklist. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online:

https://studentprivacy.ed.gov/sites/default/files/resource_document/file/checklist_data_breach_response_092012_0.pdf

Privacy Technical Assistance Center. 2014. Best practices for data destruction. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://studentprivacy.ed.gov/sites/default/files/resource_document/file/Best%20Practices%20for%20Data%20Destruction%20%282014-05-06%29%20%5BFinal%5D_0.pdf

Privacy Technical Assistance Center. 2014. Protecting student privacy while using online educational services: Requirements and best practices. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://tech.ed.gov/wp-content/uploads/2014/09/Student-Privacy-and-Online-Educational-Services-February-2014.pdf

Privacy Technical Assistance Center. 2014. Transparency best practices for schools and districts. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://studentprivacy.ed.gov/sites/default/files/resource_document/file/LEA%20Transparency%20Best%20Practices%20final.pdf

Privacy Technical Assistance Center. 2015. Data security and management training: Best practice considerations. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://studentprivacy.ed.gov/sites/default/files/resource_document/file/Data%20Security%20and%20Management%20Training_1.pdf

Privacy Technical Assistance Center. 2015. Data security checklist. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://studentprivacy.ed.gov/sites/default/files/resource_document/file/Data%20Security%20Checklist_0.pdf

Privacy Technical Assistance Center. 2015. Frequently asked questions: Cloud computing. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://studentprivacy.ed.gov/sites/default/files/resource_document/file/FAQ_Cloud_Computing_0.pdf

Privacy Technical Assistance Center. 2015. Policies for users of student data: A checklist. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://studentprivacy.ed.gov/sites/default/files/resource_document/file/Policies%20for%20Users%20of%20Student%20Data%20Checklist.pdf

Privacy Technical Assistance Center. 2015. The Family Educational Rights and Privacy Act: Guidance for reasonable methods and written agreements. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://studentprivacy.ed.gov/sites/default/files/resource_document/file/Guidance_for_Reasonable_Methods%20final_0.pdf

Privacy Technical Assistance Center. 2017. Data breach response training kit. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://studentprivacy.ed.gov/resources/data-breach-response-training-kit

Privacy Technical Assistance Center. 2017. Guidance on the use of financial aid information for program evaluation and research. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://studentprivacy.ed.gov/sites/default/files/resource_document/file/FSA_final_0.pdf

Privacy Technical Assistance Center. Frequently asked questions. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://studentprivacy.ed.gov/frequently-asked-questions

Privacy Technical Assistance Center. Privacy and education technology. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://studentprivacy.ed.gov/Apps

Privacy Technical Assistance Center. Responding to IT security audits: Improving data security practices. U.S. Department of Education. Washington, DC: Privacy Technical Assistance Center. Accessed online: https://nces.ed.gov/programs/ptac/pdf/issue-brief-responding-to-security-audits.pdf

Sanchez, P. 2018. The secret to leading organizational change is empathy. Harvard Business Review. Accessed online: https://hbr.org/2018/12/the-secret-to-leading-organizational-change-is-empathy

Seastrom, M. 2002. NCES Statistical Standards (NCES 2003-601). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Accessed online: https://nces.ed.gov/statprog/2002/stdtoc.asp

Seastrom, M. 2010. SLDS technical brief (NCES 2011-603). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Accessed online: https://studentprivacy.ed.gov/sites/default/files/resource_document/file/2011603.pdf

Sinek, S. 2005. How great leaders inspire action. TEDsPugent Sound. Accessed online: https://www.ted.com/talks/simon_sinek_how_great_leaders_inspire_action?language=en

Smith, M. 2019. Hackers breach admissions files at three private colleges. The Washington Post. Accessed online: https://www.washingtonpost.com/education/2019/03/08/hackers-breach-admissions-files-three-private-colleges/?noredirect=on&utm_term=.8c074f258852

Sundqvist, V. 2018. Yale University sued over 2008 data breach. New haven Register. Accessed online: https://www.nhregister.com/news/article/Yale-University-sued-over-2008-data-breach-13315315.php

Styleguide. Google Style Guides. Accessed online: http://google.github.io/styleguide/

Tams, C. 2018. Small is beautiful: Using gentle nudges to change organizations. Forbes. Accessed online: https://www.forbes.com/sites/carstentams/2018/02/22/small-is-beautiful-using-gentle-nudges-to-change-organizations/#325c6a185a8d

TechTarget. 2016. A buyer's guide to selecting the best data warehouse product. Accessed online: https://searchdatamanagement.techtarget.com/buyersguide/A-buyers-guide-to-selecting-the-best-data-warehouse-product

Templar, M. 2017. Get Governed. Rescue, California: Ivory Lady Publishing. Print.

United States Computer Emergency Readiness Team. Publications. U.S. Department of Homeland Security. Washington, DC: United States Computer Emergency Readiness Team. Accessed online: https://www.us-cert.gov/security-publications

Verizon. 2018. Data breach investigations report. Accessed online: https://www.verizonenterprise.com/resources/reports/rp_DBIR_2018_Report_execsummary_en_xg.pdf

Watt, A. 2014. Project Management. Open Textbook Library. Accessed online: https://open.umn.edu/opentextbooks/textbooks/project-management

Wilder-James, E. 2016. Breaking down data silos. Harvard Business Review. Accessed online: https://hbr.org/2016/12/breaking-down-data-silos

WTCS continuous improvement data library. 2018. Wisconsin Technical College System. Accessed online: https://mywtcs.wtcsystem.edu/data-systems-grp/qrp/data-description

Xplenty. 2017. What to consider when selecting a data warehouse for your business. Medium. Accessed online: https://medium.com/xplenty-blog/what-to-consider-when-selecting-a-data-warehouse-for-your-business-65755d6fcdea

Zetoony, D.A. 2016. How to conduct a data inventory. Bryan Cave Leighton Paisner LLP. Accessed online: https://www.bclplaw.com/en-US/thought-leadership/how-to-conduct-a-data-inventory-data-privacy-and-security-101.html