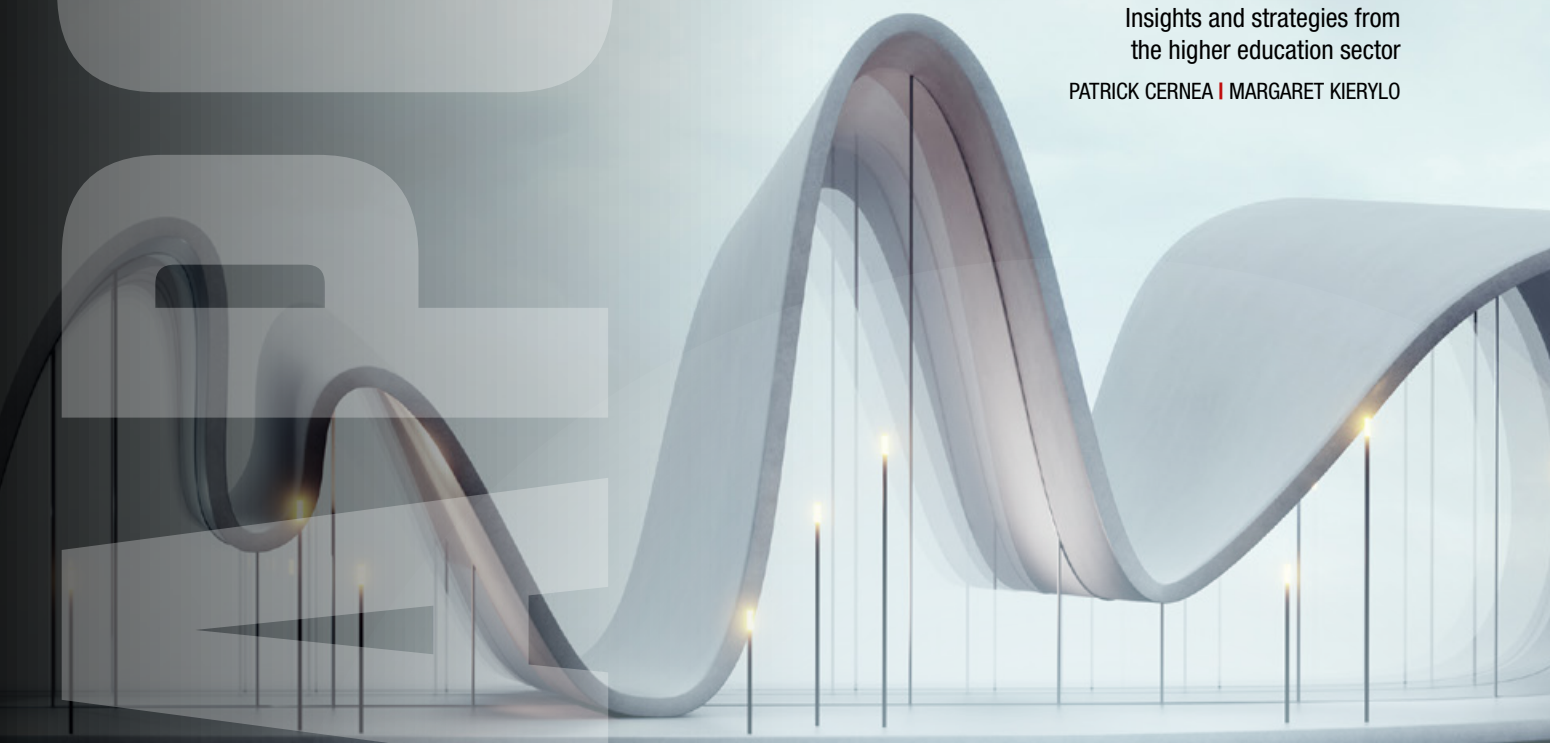
THE CAPCO INSTITUTE
JOURNAL
OF FINANCIAL TRANSFORMATION

GOVERNANCE OF TECHNOLOGY

Implementing data governance: Insights and strategies from the higher education sector

PATRICK CERNEA | MARGARET KIERYLO



BALANCING
INNOVATION & CONTROL

a wipro company

#59 JUNE 2024

THE CAPCO INSTITUTE

JOURNAL OF FINANCIAL TRANSFORMATION

RECIPIENT OF THE APEX AWARD FOR PUBLICATION EXCELLENCE

CONTENTS

GOVERNANCE OF TECHNOLOGY

GOVERNANCE OF SUSTAINABILITY

GOVERNANCE OF CORPORATES

DEAR READER,

In my new role as CEO of Capco, I am very pleased to welcome you to the latest edition of the Capco Journal, titled **Balancing Innovation and Control**.

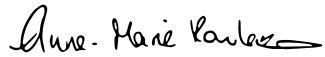
The financial services and energy sectors are poised for another transformative year. At Capco, we recognize that this is a new era where innovation, expertise, adaptability, and speed of execution will be valued as never before.

Success will be determined based on exceptional strategic thinking, and the ability to leverage innovative new technology, including GenAI, while balancing a laser focus on risk and resilience. Leaders across the financial services and energy industries recognize the transformative benefits of strong governance while needing to find the optimal balance between innovation and control.

This edition of the Capco Journal thus examines the critical role of balancing innovation and control in technology, with a particular focus on data, AI, and sustainability, with wider corporate governance considerations. As always, our authors include leading academics, senior financial services executives, and Capco's own subject matter experts.

I hope that you will find the articles in this edition truly thought provoking, and that our contributors' insights prove valuable, as you consider your institution's future approach to managing innovation in a controlled environment.

My thanks and appreciation to our contributors and our readers.



Annie Rowland, **Capco CEO**

IMPLEMENTING DATA GOVERNANCE: INSIGHTS AND STRATEGIES FROM THE HIGHER EDUCATION SECTOR

PATRICK CERNEA | Director, Data Strategy and Governance, York University, Canada
MARGARET KIERYLO | Assistant Vice-President, Institutional Planning and Chief Data Officer, York University, Canada

ABSTRACT

This article explores the critical role of data governance in the context of higher education. The authors highlight the strategic importance of establishing comprehensive data governance frameworks to enable data-informed decision-making and argue that data governance can enhance strategic enrollment management, efficiency and effectiveness, and enable innovation. The authors present a detailed exploration of the strategies for building data governance capabilities within higher education institutions. They outline the process of setting data governance goals, selecting operating models for data governance, considering resourcing models, defining roles and responsibilities, establishing data governance committees, and identifying metrics to assess progress. Practical applications of data governance, including metadata management, data quality management, and ensuring regulatory compliance and ethical use of data, are discussed to illustrate how institutions can enhance their data environment. The authors conclude by exploring future trends and emerging issues in data governance within higher education, pointing to the lag in data governance advancement compared to other sectors and the imperative for post-secondary institutions to adopt robust data governance frameworks to remain competitive and innovative.

1. INTRODUCTION

1.1 The data environment

Organizations – including institutions of higher education – agree that data is critical to management and decision making. Over the past several decades, digitalization and the use of data have expanded into every corner of higher education institutions, including universities and colleges. Departments have become increasingly data-driven, relying on data for their daily operations. As the number of data creators and users has grown, so too has the variety of approaches to collecting, producing, distributing, and analyzing data. A fully data-driven post-secondary institution extends beyond using data for routine tasks, embedding data-derived insights into strategic and operational decision making at all levels, from front-line units to senior leadership.

Despite the widespread acknowledgment of the importance of these developments, many institutions are still at the formative stages of developing and implementing comprehensive data governance frameworks. In Canada, for example, a 2023 survey conducted by the Canadian Association of University Business Officers (CAUBO) revealed that 47% of higher education institutions are just beginning their data governance journeys [Al-Hussein et al. (2023)]. This statistic highlights that nearly half of Canadian post-secondary institutions are beginning to realize the importance of systematically managing their data assets.

Using data well is a core competency for all successful organizations, and all sectors must continually get better at it. In the post-secondary context, beyond enabling reporting, institutions are leveraging data to grow enrollment, enhance efficiency and effectiveness, and drive innovation. This

strategic use of data for growth and efficiency sets the stage for adopting advanced technologies, notably artificial intelligence (AI). AI algorithms and machine learning techniques are now being employed to analyze complex datasets, offering insights that previously required extensive effort. These technologies are enabling post-secondary institutions to predict trends, automate tasks, and personalize student experiences. This evolution in data utilization marks a significant shift in how institutions approach their strategic goals, illustrating the dynamic nature of data as a tool for impact and innovation.

Data is fundamental in how we prepare for the future, share knowledge, and make good decisions. However, transitioning to a fully data-driven model entails significant planning and coordination. For many organizations, the full potential of data and analytics capabilities remains untapped. The key requirement is not simply more data or more assessment, but the establishment of systematic data and analytics practices.

In this context, the formulation of comprehensive data strategies is vital for organizations to harness the power of their information assets effectively [Powers (2020)]. These strategies should not only outline the overarching framework for data utilization but also emphasize the development of key institutional functions: data literacy and data governance. By enhancing data literacy, organizations equip their community with the skills necessary to interpret, analyze, and leverage data effectively. Simultaneously, data governance ensures the integrity, security, and ethical use of data [Eryurek et al. (2021)]. Together, these elements create a strong foundation for a data-driven culture, one that facilitates informed decision making and fosters innovation across all levels of the organization. In the post-secondary context, the aim is to transform data into a strategic asset that enhances teaching and learning, improves student experiences, allows for the delivery of smarter student services, enables effective strategic enrollment management, promotes operational effectiveness, excellence, and efficiency, and drives institutional innovation.

1.2 What is data governance?

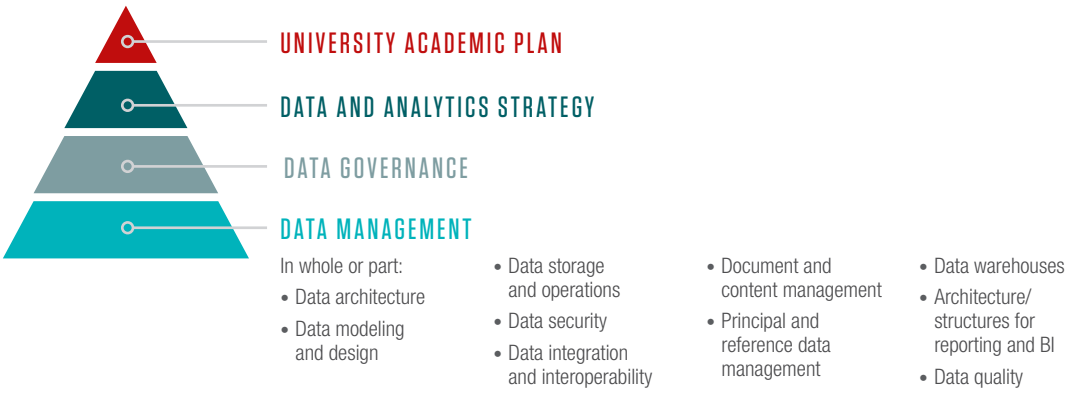
Data governance encompasses the practices, policies, processes, standards, and metrics required to manage data as an asset [DAMA (2017)]. Its purpose is to utilize data effectively and efficiently in support of organizational goals. Data governance has a symbiotic relationship with data management: data governance establishes data policies and procedures, while data management implements those policies and procedures to aid in decision making.

We can illustrate this distinction by using data quality as an example:

- Data governance involves establishing data quality standards (e.g., the acceptable levels of data accuracy, completeness, and consistency), specifying the roles and responsibilities related to data quality (e.g., the responsibilities of “data stewards” and “data custodians”), and creating policies and processes that dictate how data quality should be maintained and enhanced.
- Data management involves the practical implementation of these policies, standards, and procedures. This includes day-to-day operations and technologies used to ensure data meets established quality standards. Activities under data management may include data quality controls (e.g., using software tools to monitor and report on data quality), conducting data quality audits, and undertaking data cleansing and improvement efforts.

An overview of data governance and its relationship to a typical university planning framework is provided in Figure 1.

Figure 1: Data governance in a typical university planning framework



1.3 Key data governance terms

To provide a foundational understanding of data governance, it is important to understand the core terms and concepts that make up this field. This section delves into key data governance terminologies, offering insights into their meanings and implications within an organization. Key data governance terms include:

- **Data:** refers to all quantitative and qualitative information that is collected, stored, managed, analyzed, and utilized across various organizational departments and functions.
- **Data custodian:** generally, the individual who is responsible for the technical management and security of a particular data system or dataset.
- **Data domain:** generally, a specific category or subject area of data within the organization. It is a broad area of data that contains a set of similar or related data elements, such as financial data and human resources data.
- **Data steward:** generally, a senior manager who is responsible for the data in one or more data sub-domains. Data stewards are usually required to be experts in data within their data sub-domain(s).
- **Data sub-domain:** generally, a subset or a specific aspect of a data domain. It is a smaller, more specific area of data (that is part of a larger data domain), such as staff profile data, payroll data, and employee safety data.
- **Data trustee:** generally, a senior executive who is accountable for the data in one or more data domains. Data trustees usually have decision-making authority regarding the authoritative sources of data that are managed and created by the central unit.
- **Metadata:** structured, descriptive information about data elements and data assets that provides context, facilitates understanding, and enables effective management, discovery, and usage. For instance, for a "student ID" data element, the metadata might include a definition and a validation rule.
- **Principal data:** core data that is essential for the organization's operations and decision-making processes. This data is used to identify, describe, and manage the primary aspects of the organization. At higher education institutions, principal data includes information about students, alumni, staff, faculty, academic programs and services, organizational and financial structures, and physical space. Principal data is often synonymous with "master data".
- **Reference data:** the sets of predefined, permissible values or categories that are used within the organization's systems and databases to classify, organize, and ensure the consistency of data. This data provides context and structure to transactional and operational data, enabling accurate data interpretation, reporting, and analysis. At higher education institutions, reference data includes country codes, currency codes, and program classification codes.

1.4 Making the case for data governance

Data governance facilitates better decision making and operational efficiency and effectiveness primarily through improved metadata, data quality, data protection and compliance, and other data management policies and procedures. There are two main reasons to embrace data governance: increasing the value of data and reducing risks associated with poor data management.

1.4.1 INCREASING THE VALUE OF DATA

Today's data environment, characterized by growing volume, complexity, and AI breakthroughs, makes data governance essential for organizations to maintain or gain a competitive advantage. Enhanced visibility over data assets, increased data literacy, standardized data language, and improved data quality all contribute to making data more valuable, thus improving business outcomes. Companies such as Airbnb, GE Aviation, and Uber all leverage data governance to enhance decision making [Atlan (2023)]. In a 2023 survey of the Canadian higher education sector, 41% of universities and colleges highlighted supporting decision making as a key outcome of data governance at their institution [Al-Hussein et al. (2023)].

1.4.2 REDUCING RISKS

The evolving data landscape brings increased regulatory requirements, exemplified by the European Union's General Data Protection Regulation (GDPR), which highlights the legal implications of data mismanagement and risks like security breaches, revenue loss, and reputational damage [E.U. (2016)]. Data governance is crucial for risk management. In the Canadian higher education sector, 18% of universities and colleges highlighted ensuring compliance as a key outcome of data governance at their institution [Al-Hussein et al. (2023)].

2. BUILDING DATA GOVERNANCE CAPABILITIES

This section highlights key considerations in the development of data governance programs. Outlined below are five principles that guided the implementation of data governance at our university.

- **Strategy should drive data governance efforts:** data strategies should inform data governance processes as they provide a comprehensive roadmap to effectively manage and leverage data assets. By aligning data governance processes with an overall data strategy, institutions can ensure that their data governance efforts are focused and relevant to their specific needs and objectives. In the Canadian higher education sector, 21% of universities and colleges highlighted the lack of a business case as a roadblock to data governance adoption [Al-Hussein et al. (2023)].
- **Communication is essential:** since data governance fundamentally involves people, institutions often find that it is crucial to focus on clear and frequent communication. Effective communication strategies are vital to successfully implement data governance, as they ensure that stakeholders are aligned and engaged with institutional data governance initiatives.
- **Data governance efforts should be focused:** institutions should start their data governance journey by focusing on key data domains and initiatives. Data governance is often misunderstood or seen as a high-level strategic initiative, making quick wins essential. Focusing data governance efforts also streamlines resource allocation. In the Canadian higher education sector, 45% of universities and colleges highlighted capacity risk as a roadblock to data governance adoption [Al-Hussein et al. (2023)].
- **Data literacy should be prioritized:** data literacy is crucial for effective data governance. All members of the organization, regardless of their role or unit, should develop a basic understanding of data and its significance. This facilitates better collaboration, informed decision making, and more effective data governance practices. In the Canadian higher education sector, 55% of universities and colleges highlighted a lack of data literacy as a roadblock to data governance adoption [Al-Hussein et al. (2023)].
- **Data governance implementation requires robust change management:** institutions should pay special attention to change management, especially in the early stages of data governance implementation when goals and deliverables are ambiguous. In the Canadian higher education sector, 58% of universities and colleges highlighted a lack of change management as a roadblock to data governance adoption [Al-Hussein et al. (2023)].

2.1 Data governance vision, mission, and goals

Organizations typically begin their data governance journey by crafting a vision and mission for the program, ensuring alignment with their overarching data strategy. Although each organization is unique, Canadian higher education institutions often have vision and mission statements that emphasize leveraging data as a strategic asset, ensuring data quality and security, and fostering data-informed cultures within their institutions.

The data governance program derives its goals from its vision and mission. In the context of Canadian higher education institutions, these goals commonly prioritize enhancing data quality, maintaining metadata, developing data literacy, ensuring compliance, fostering a culture of data sharing, and protecting data assets. These goals help operationalize data governance initiatives and facilitate tracking progress against targets.

2.2 Operating models for data governance

Implementing an operating model that aligns with the organization's unique data environment and business objectives is paramount. Centralized, decentralized, and federated models each have distinct advantages and challenges.

- **Centralized data governance:** typically consolidates data governance authority in one department or unit. Although this model may benefit from uniformity and consistency, it may suffer from slower decision making and a lack of functional subject-matter expertise.
- **Decentralized data governance:** distributes authority across the organization and various departments. This model promotes agility and customization but may lead to inconsistencies around data handling, data definitions, and the implementation of policies and processes.

- **Federated data governance:** blends centralized oversight with decentralized execution. It offers a compromise between uniformity and flexibility. This model works well for most organizations, including higher education institutions.

At our university, we implemented a federated data governance model in which the central data governance team, in consultation with key partners such as data trustees and data stewards, defines policies, standards, and guidelines for data management. The central data governance team is also responsible for managing the institutional metadata repository and developing the university's data literacy program. Day-to-day data management activities are handled at the sub-domain level by data stewards and data custodians, with support from information technology (IT).

2.3 Resourcing models for data governance

Resource allocation varies by organization size and needs. Organizations typically consider in-house teams, outsourcing, or hybrid models based on available resources and expertise. Decisions should be guided by the nature and duration of the data governance work. If an organization prioritizes data quality, does it have the required expertise to undertake that work? If not, can it hire that expertise? The duration of the work also plays a crucial role: short-term projects might be better suited for outsourcing, whereas long-term engagements may benefit from developing and retaining expertise internally. For example, organizations might choose to do the metadata work in-house and outsource for data quality assessments and cleansing.

2.4 Roles and responsibilities in data governance

With the resourcing model in place, organizations must consider additional roles and responsibilities related to data governance. Stakeholder maps and engagement plans allow for the identification and involvement of key groups in data governance activities.

As data governance is often perceived as yielding less tangible results, organizations typically begin their journey by securing an executive sponsor to advocate for the data governance program. Another crucial role is the chief data officer (CDO), who is responsible for overseeing the data governance strategy, ensuring data quality, and driving the cultural change toward data-driven decision making across the organization.

Organizations subsequently assign accountability for enterprise data by segmenting it into data domains and sub-domains. Typically, in the higher education sector, data domains represent broad categories like student, financial, and human resources data, while sub-domains are more specific areas within these domains.

Generally, as noted in Section 1.3 above, there are three roles found in higher education institutions:

- **Data trustees:** usually senior executives who oversee one or more data domains.
- **Data stewards:** often senior managers who manage one or more data sub-domains and are considered experts in their areas.
- **Data custodians:** typically IT professionals who handle the technical aspects of data sub-domains, systems, or both.

Fundamentally, data governance hinges on people and cross-functional collaboration for success and it permeates the entire organization.

2.5 Data governance committees

Data governance committees are essential, as they facilitate decision making and policy development and help advance operational work. The specific composition of data governance roles may require creating dual committees alongside various working groups:

- A highly strategic data trusteeship committee or data governance council that sets the direction for the data governance program and ensures alignment with the institutional data strategy.
- A more operational data stewardship committee that plans and undertakes project work.
- Working groups that reflect data sub-domains, involve data custodians, or both.

Whatever structure is established, committees should have clear mandates and goals, involve all relevant stakeholders, and ensure consistency through regular meetings.

2.6 Metrics: Measuring the effectiveness of data governance

Data governance is often underprioritized as an institutional function because of two key issues: misalignment with strategic and operational objectives and a lack of tracking data governance initiatives. Identifying key performance indicators for each data governance goal is crucial for monitoring progress and assessing the impact on business operations.

Metrics used to track the progress of data governance goals might include the number of data stewards identified, the number of data definitions approved, and measures of engagement in data governance workshops.

Metrics that measure the impact on operations are harder to quantify, but they are ultimately critical to showcase value. They can include efficiency gains due to data governance, decreased penalties from avoiding regulatory non-compliance, or time saved due to improved data quality [for examples of other metrics, see Pansara (2023)].

Data governance measures should be actionable and top-of-mind for those accountable for data domains and sub-domains, and progress should be shared with institutional stakeholders.

3. THREE PRACTICAL APPLICATIONS

With foundational data governance elements established, organizations can start to undertake data governance work to enhance their data environment. Three practical applications include metadata management, data quality management, and regulatory compliance and ethical use of data. It is important to note that implementing these initiatives requires collaboration across various teams. Specifically, improving data quality requires significant collaboration between the data governance and data management teams and business units.

To prioritize data governance projects and sub-domains of focus, organizations typically conduct a business impact analysis to understand the value and sensitivity of their data assets while identifying high-risk areas. An instrumental part of this process is the development of a prioritization matrix that incorporates criteria such as strategic and operational alignment, business value, revenue potential, risk mitigation, and resource availability.

Higher education institutions typically start this prioritization process by identifying the data most critical to key institutional initiatives, such as strategic enrolment management. In the Canadian higher education sector, 68% of universities and colleges highlighted performance data collection and analysis, and 63% highlighted enrollment management as key use cases of data governance at their respective institutions [Al-Hussein et al. (2023)]. Often, this involves prioritizing student sub-domains such as “student profile”, “recruitment and admissions”, and “student advising”, as well as principal and reference sub-domains such as “principal academic programs and services” and “reference geographic locations”. The work undertaken in those sub-domains often begins with a focus on managing metadata, ensuring data quality, promoting data literacy, and developing relevant data policies and procedures.

3.1 Metadata management

As noted in Section 1.3, “metadata” is information that describes and provides context for other data. In essence, it describes the various aspects of data, like its content, format, source, and context. Organizations gain visibility and understanding of their data once they inventory it, define it, and track its lineage. Metadata management, therefore, involves ensuring that data is managed with the same rigor as any other valuable asset. A 2023 survey of post-secondary institutions found that only 25% of respondents reported that their institutions have clear and comprehensive data definitions that adequately cover the nuances of data [Al-Hussein et al. (2023)].[1]

Why should organizations undertake metadata management? Consider the following scenarios:

- Senior management does not fully grasp the intricacies of the data in an institutional performance report. The solution might be to define data elements in an institutional metadata repository.
- An organization is implementing a new customer relationship management (CRM) software and notices unnecessary duplication of data assets. The first step might be to catalog data assets, their content, formats, and sources.
- An organization wants to improve the quality of its financial data but does not know where to start. The first step might be to inventory data elements, enabling them to document data quality rules and conduct a data quality assessment.

[1] Results include “strongly agree” and “somewhat agree” out of a 5-point Likert scale ranging from “strongly disagree” to “strongly agree”.

Given that metadata management is crucial for the proper governance of an organization's data assets, applying a set of best practices will facilitate its successful implementation. The following are recommendations developed through operationalizing metadata management at our university:

- **Establish clear policies and standards:** a metadata guidelines document, for instance, should define the types of metadata accepted, prescribe standards for writing definitions, and set standards for cataloging data assets.
- **Involve stakeholders across departments and units:** metadata management should involve data trustees, data stewards, data custodians, and subject matter experts. In creating institutional data definitions, it is essential to include relevant experts. For instance, when standardizing reference geographic data, institutions should involve those who manage and use this data to ensure a consistent organizational standard.
- **Provide training opportunities and conduct awareness campaigns:** educate staff on the importance of metadata management and its effective uses. Institutions could decide to provide every data trustee, data steward, data custodian, and subject matter expert with an onboarding session on the institutional metadata repository.
- **Use metadata management tools:** specialized tools and software such as Informatica Enterprise Data Catalog, Collibra, and Data Cookbook are specifically designed to help create, store, and retrieve metadata effectively. Tools should be chosen based on the size and complexity of the institution and its budget.
- **Enforce data security and privacy:** institutions should ensure their metadata management practices comply with security and privacy policies. Additionally, they should use their metadata guidelines document and onboarding sessions to communicate these standards. For instance, data definitions should not include sensitive information.

We began our metadata management journey by defining terms required for strategic enrollment management and our digital transformation program. This meant primarily involving teams from the "student" and "principal academic programs and services" domains and ensuring integration of data definitions with relevant dashboards and reports, such as the university's enrollment management dashboard.

3.2 Data quality management

High-quality data leads to better decisions, facilitates strategic planning, and reduces the time employees spend on ad-hoc assessments, data manipulation, and cleansing. Implementing transparent processes for managing data quality can also significantly enhance trust in organizational data. A recent survey of Canadian post-secondary institutions found that only 66% of respondents believe their institutions' data is trustworthy [Al-Hussein et al. (2023)].

The key to improving data quality sustainably is to establish an institutional data quality management program. This approach allows organizations to focus their scope, align data quality improvements to business outcomes, and streamline resource allocation. In essence, this enables organizations to:

- Enhance decision making and drive strategic impact by improving the value and usability of their data.
- Increase efficiency and productivity by streamlining data processes and minimizing delays caused by data inaccuracies.
- Improve customer and stakeholder satisfaction through reliable data-driven services and interactions.
- Reduce risks associated with poor data quality, such as compliance issues and reputational damage.
- Enable a modern ecosystem of integrated information platforms and applications, which requires high-quality data to function properly.

Data quality improvements should have a specific and focused scope. Since data quality efforts can become costly and time-consuming, organizations should initially aim for quick wins and impact on critical business areas. Typically, this is done by surveying data producers and users to identify the most significant data quality issues. Improvements should then be prioritized based on business value. Higher education institutions usually begin their data quality journey by addressing student data. According to the Data Management Body of Knowledge, a systematic approach to data quality includes:

- Defining high-quality data.
- Defining a data quality strategy.
- Identifying critical data and business rules.
- Performing an initial data quality assessment.
- Identifying and prioritizing potential improvements.

[2] Ibid.

- Defining goals for data quality improvement.
- Developing and deploying data quality operations including:
  - managing data quality rules
  - measuring and monitoring data quality
  - developing operational procedures for managing data issues
  - establishing data quality service level agreements
  - developing data quality reporting [DAMA (2017)].

Building on this systematic approach, enhancing an organization's data quality typically involves implementing a set of best practices tailored to its specific needs and goals. To ensure the effectiveness of their data quality management program, organizations should:

- Prioritize data quality work based on sub-domains and systems critical to the business.
- Leverage use cases and data quality process maps as catalysts for data quality.
- Ensure adequate resources are in place in IT to enable data management functions.
- Implement measures to keep low-quality data out of the organization's data ecosystem. This often involves establishing data entry controls and defining data quality rules.
- Leverage tools for data quality profiling (e.g., IBM InfoSphere Information Analyzer, Collibra), modeling and "extract, transform, and load" (ETL) processes (e.g., Informatica, AWS Glue), metadata management (e.g., Informatica Enterprise Data Catalog, Data Cookbook), incident tracking (e.g., Jira, Zendesk), and data quality reporting (e.g., Power BI, Tableau).
- Empower stakeholders responsible for data quality (e.g., data stewards) to:
  - decide if their data is sufficiently complete and accurate to support business process needs
  - set up targets for specific attributes
  - set up thresholds for the level of quality acceptable
  - establish measures and metrics to track improvements [HealthIT].

Our university began enhancing data quality in the student information system by prioritizing the data assets and elements critical to strategic enrollment management. This effort primarily involved data stewards and data custodians from the "student profile" data sub-domain. Initial tasks included documenting existing data quality processes, capturing data quality rules in the institutional metadata repository, and performing an initial data quality assessment.

3.3 Regulatory compliance and ethical use of data

In the context of data management, regulatory compliance and ethical considerations in data usage are of paramount importance, especially with the rise of technologies like AI and cloud computing, which have increased the possibilities for innovative (and unethical) data use. Adhering to legal standards and ethical guidelines is essential not only to avoid legal repercussions but also to maintain public trust and safeguard the rights and privacy of individuals. For readers of this article, various legislation and regulations could apply, including:

- the General Data Protection Regulation (GDPR) in the European Union
- the California Consumer Privacy Act (CCPA) in the United States
- international standards like ISO/IEC 27001 for information security management
- sector-specific regulations, such as those in finance and education (e.g., Ontario's Ministry of Training, Colleges and Universities Act).

Ethical use of data goes beyond legal compliance; it encompasses respect for confidentiality, consent, fairness, and transparency in data handling. A commitment to ethical data usage is crucial in building a responsible and sustainable data culture within organizations.

3.3.1 ETHICAL PRINCIPLES FOR THE USE OF DATA

At our university, this work was undertaken by the Principles for the Ethical Use of Student Data Working Group, which had representation from across the university. The group was tasked with developing guiding principles to ensure ethical use of student data, aligning with the university's commitment to decolonization, equity, diversity, and inclusion (DEDI), and compliance with institutional values and policies.

Between Fall 2022 and Spring 2023, the working group focused on developing high-level principles for future projects and activities involving student data. They balanced various values, such as student privacy and the duty to act, and recognized the ongoing debate around ethical principles. The group emphasized the need for a continuous, fact-informed discussion.

The scope of the principles covers all data related to current and prospective students, and alumni. It includes a wide range of activities from advising to the use of AI and learning analytics.

Key principles include:

- **Consent:** this principle refers to the explicit, informed, and meaningful agreement given by an individual for their personal data to be collected, processed, or analyzed for a specific purpose.
- **Transparency:** this principle refers to the obligation to be open and honest about organizational data practices. The principle of transparency facilitates the ability of students and alumni to provide free and informed consent.
- **Duty of care:** this principle refers to the obligation to take steps to ensure that data is collected, processed, and used in a way that does not cause harm.
- **Obligation to act:** this principle is about ensuring that personal student data is collected, processed, and used in a way that is aligned to the best interests of students.
- **Data minimization:** this principle refers to collecting, processing, and using the minimum amount of personal data necessary to achieve a specific purpose.
- **Stewardship of data:** this principle refers to the responsible management of data.

These principles reflect emerging practices for ensuring compliance and ethical data usage. To ensure continued alignment with best practices, the working group's recommendations include developing resources for faculty and staff to build literacy in interpreting student data, conducting an annual review of the principles, and implementing a communication plan. The operationalization of the principles is currently underway. For instance, when making data requests, the requestor is required to review and align their request with the "Principles for the ethical use of student data". This step ensures compliance with the established ethical guidelines set forth for student data usage.

3.3.2 METHODS FOR PRIORITIZING COMPLIANCE AND ETHICAL USE OF DATA IN BUSINESS OPERATIONS

As AI continues to permeate various sectors including finance, transportation, healthcare, and higher education, it will become increasingly important to balance these technologies against potential ethical risks [Kaushikkumar (2024)]. To effectively prioritize ethical data usage within an organization's operations, organizations need to adopt a strategic approach. Formulating comprehensive policies, guidelines, or frameworks is essential; these should meet legal standards and embody ethical principles that align with the organization's strategic directions while balancing benefits and risks.

Furthermore, it is essential to initiate regular training programs for employees, focusing on data ethics and legal compliance. These training sessions are crucial in cultivating a culture of awareness and responsible data usage.

Another key method is conducting thorough data audits. These audits play a vital role in verifying adherence to both regulatory requirements and ethical principles, thereby reinforcing overall compliance.

Finally, the integration of ethical considerations into decision-making processes, especially for projects involving personal data, is critical. This practice ensures that decisions are made with an ethical lens, not just a legal one, thus embedding a sense of trust and integrity in business operations.

Together, these methods form a strong framework for ensuring that compliance and ethical use of data are central to an organization's operations [Braunack-Mayer et al. (2020)]. The integration of ethical considerations into data usage is not just a legal necessity; it is a cornerstone of building trust and integrity in business operations. By prioritizing these aspects, organizations not only protect themselves from legal risks but also establish themselves as responsible and trustworthy entities in the eyes of their customers and the public.

4. CONCLUSION AND FUTURE TRENDS

Data governance represents a fundamental shift in how organizations value and manage one of their most critical assets: data. This article has provided insights into data governance within the context of higher education. As we have demonstrated, data governance is a critical function as institutions prepare for the future of learning; insights derived from well-governed data can lead to transformative outcomes for students, faculty, and the broader community. Without data governance, an institution's ability to leverage data as a strategic asset is limited. As we have explored, the data governance journey is multifaceted and dependent on collaboration and coordination, and involves comprehensive metadata management, a rigorous pursuit of data quality, and adherence to ethical standards.

Through the establishment of comprehensive data strategies, higher education institutions can harness the full potential of their data, enhancing decision making, operational efficiency

and effectiveness, and innovation. The journey involves not only the technical aspects of data management but also a cultural shift towards data literacy and a shared responsibility for data governance across all levels of the institution.

As we look to the future, several trends and emerging issues are evident:

- Compared to sectors such as health and finance, post-secondary institutions have lagged in advancing data governance as a core competency. Universities and colleges will seek to unlock the full potential of their data assets through the implementation of data governance frameworks [CAUBO (2023)].
- The pace of technological innovation may outstrip the ability for regulatory frameworks to adapt. AI and machine learning rely heavily on large datasets. Data governance ensures that the data used to train these systems is not only high quality and relevant but also ethically used. This underscores the importance of data governance as organizations navigate the complexities of an AI-driven future [Kaushikkumar (2024)]. As regulations evolve to catch up with technology, organizations with robust data governance frameworks will be better positioned to remain agile and competitive in their respective fields.
- Given identified research gaps in data governance, future research trends will explore data privacy and the evolving landscape of data governance, with emphasis on the interplay between AI and data stewardship practices. Other topics of interest include organizational challenges related to governance implementation, the impact of AI on data governance, and cross-border regulatory compliance [Pansara (2023)].
- Concerns about bias in "AI algorithmic decision support" will continue. Data governance can assist by ensuring that datasets are diverse and representative. By monitoring and managing the composition of datasets (metadata), data governance can help anticipate and prevent biases that arise from underrepresented groups or skewed data samples, leading to more equitable AI outcomes [Davidson (2023)]. Ethical considerations will become as important as legal compliance.
- Institutions committed to decolonization, equity, diversity, and inclusion (DEDI) and Indigenization face the challenge of navigating an even more complex data governance landscape. Understanding Indigenous ways of knowing and Indigenous data systems will be crucial in supporting Indigenous data sovereignty [Animikii, 2022].

As we have demonstrated, the role of data governance in higher education will only continue to grow in importance. Implementation requires a commitment to continuous improvement, collaboration, and alignment with institutional priorities. Moreover, data governance efforts should be focused, targeting areas of greatest impact and importance. Prioritizing data literacy is essential, as it empowers individuals across the institution to effectively interpret data and apply insights to enable informed decision making. Finally, change management is required to navigate the complexities of data governance implementation successfully. By prioritizing data governance and committing to effective data management, post-secondary institutions can ensure they are well-positioned to meet the challenges of the 21st century, driving innovation and excellence in higher education.

REFERENCES

Al-Hussein, R., P. Cernea, Z. Chan, M. Kierylo, and M. Morgado. 2023, "Data governance capabilities," Webinar, Canadian Association of University Business Officers, October 12, http://tinyurl.com/2y2dkrf3

Animikii, 2022, "#DataBack: asserting and supporting indigenous data sovereignty," http://tinyurl.com/yfnu7h7r

Atlan, 2023, "5 data governance examples: case studies, takeaways and more." May 25, http://tinyurl.com/2v5uv6xn

Braunack-Mayer, A. J., J. M. Street, R. Tooher, X. Feng, and K. Scharling-Gamba, 2020, "Student and staff perspectives on the use of big data in the tertiary education sector: a scoping review and reflection on the ethical issues," Review of Educational Research 90:6, http://tinyurl.com/y6b7feps

CAUBO, 2023, "New resource: data governance framework for Canadian universities," Canadian Association of University Business Officers, June 27, http://tinyurl.com/38anamj3

DAMA, 2017, DAMA-DMBOK: Data management body of knowledge, 2nd edition, Technics Publications

Davidson, E., L. Wessel, J. Sunrise Winter, and S. Winter, 2023, "Future directions for scholarship on data governance, digital innovation, and grand challenges," Information and Organization 33:1, http://tinyurl.com/67f4dbae

Eryurek, E., U. Gilad, V. Lakshmanan, A. Kibunguchy-Grant, and J. Ashdown, 2021, Data governance: the definitive guide: people, processes, and tools to operationalize data trustworthiness, O'Reilly Media

E.U. 2016, "General Data Protection Regulation," Regulation (EU) 2016/679 of the European Parliament and of the Council, April 27, http://tinyurl.com/32yzry5j

HealthIT, "Data quality assessment," Office of the National Coordinator for Health Information Technology Health IT Playbook, http://tinyurl.com/997pb2kh

Kaushikkumar, P. 2024, "Ethical reflections on data-centric AI: balancing benefits and risks," International Journal of Artificial Intelligence Research and Development (IJAIRD) 2:1, http://tinyurl.com/27b2ss83

Pansara, R. R., 2023, "Unraveling the complexities of data governance with strategies, challenges, and future directions," Transactions on Latest Trends in IoT 6:6, http://tinyurl.com/m7sdny4s

Powers, K., 2020, Data strategy in colleges and universities, Routledge

ABOUT CAPCO

Capco, a Wipro company, is a global technology and management consultancy focused in the financial services industry. Capco operates at the intersection of business and technology by combining innovative thinking with unrivalled industry knowledge to fast-track digital initiatives for banking and payments, capital markets, wealth and asset management, insurance, and the energy sector. Capco's cutting-edge ingenuity is brought to life through its award-winning Be Yourself At Work culture and diverse talent.

To learn more, visit www.capco.com or follow us on Facebook, YouTube, LinkedIn and Instagram.

WORLDWIDE OFFICES

APAC
Bengaluru – Electronic City
Bengaluru – Sarjapur Road
Bangkok
Chennai
Gurugram
Hong Kong
Hyderabad
Kuala Lumpur
Mumbai
Pune
Singapore

MIDDLE EAST
Dubai

EUROPE
Berlin
Bratislava
Brussels
Dusseldorf
Edinburgh
Frankfurt
Geneva
Glasgow
London
Milan
Paris
Vienna
Warsaw
Zurich

NORTH AMERICA
Charlotte
Chicago
Dallas
Hartford
Houston
New York
Orlando
Toronto

SOUTH AMERICA
São Paulo

WWW.CAPCO.COM

CAPCO
a wipro company