University of Toronto

Towards an Institutional Data Governance Program

Report of the Data Governance Committee February 2020

UNIVERSITY OF
TORONTO

Table of Contents

Contents

Background and Context

The increasing availability of rich and plentiful data, together with powerful tools for collecting and processing data, is transforming the global community. The University of Toronto is no exception. Its mission increasingly relies upon the adept and efficient use of institutional data – everything from enrolment numbers and information about financial aid, to recruitment details, learning analytics, research funding statistics, and bibliometrics. Such data are abundant and growing exponentially in the modern university context, and the tools for analyzing them are steadily becoming more sophisticated and accessible.

These are positive developments. Institutional data, properly understood and studied, can help guide decisions, policies, and strategies. Such data can produce unexpected insights or ideas, revealing opportunities for the university to improve its operations and better serve its broad community – students, faculty, staff, alumni – and society at large.

However, if institutional data are to be useful, they must be well-managed, accurate, understandable, and accessible. Institutional data that are ambiguous, erroneous, inconsistent, or unavailable are problematic. There is an inherent tension between the *value* created by the appropriate and widespread use of institutional data and the *risks* – of data corruption or misinterpretation, for example – occasioned by greater openness and access. Managing this tension, safeguarding and enriching the value of institutional data, is at the heart of what has come to be called *data governance*.

Moreover, behind most institutional data are people. Individuals or groups are affected, directly or indirectly, by the way data are used, the inferences they yield, and the decisions to which they lead. It is not surprising, therefore, that an evolving regulatory landscape is setting requirements and limitations for how institutions and individuals interact with data. For example, the *Freedom of Information and Protection of Privacy Act* (FIPPA) and the *Simpler, Faster, Better Services Act* (SFBSA) in Ontario – or the *General Data Protection Regulation* (GDPR) in the European Union – influence the way data are handled across every sector. Facilitating regulatory compliance, along with promoting the responsible use of institutional data more generally, is an essential part of data governance. Good data governance fosters the trust of all those who interact with institutional data.

For these reasons, the University of Toronto is developing a formal *Institutional Data Governance Program*. Until now, where it has been practiced at all, institutional data governance at the University has been largely informal and piecemeal. Its success has been uneven. In place of the current approach, the University will implement a collaborative and coordinated program. The program will rest upon a set of guiding principles, definitions, and goals. And it will build upon the multitude of knowledgeable partners and deep expertise already existing across the institution.

Data Governance: Definitions and Goals

What is data governance?

Data governance, in the University of Toronto context, is the process that both establishes and provides consistent and clear guidance around creating, accessing, using, and defining the University's institutional data, while managing the associated risks, responsibilities, and legal obligations.

The University of Toronto Institutional Data Governance Program will expressly respect – and, indeed, rely upon – the decentralized nature of the University and divisions' local authority and expertise.[1] At the same time, it represents a collective and strategic, institution-wide approach to managing the University's institutional data.

What are institutional data?

Institutional data comprise all of the data that are held by the University for the purpose of supporting its administrative operation, broadly understood.[2]

For the purposes of the Institutional Data Governance Program, research data are data that are held at the University for conducting scholarly research, and are outside the scope of the Institutional Data Governance Program.

What is the Program's goal?

The goal of the Data Governance Program at the University of Toronto is to promote and support the responsible use of high-quality institutional data, to facilitate informed and insightful use of these data, and to increase their value to the University community and beyond.

What are the Program's desired outcomes?

The primary, long-term objective of the Institutional Data Governance Program is to create a culture in which everyone who collects, manages, or uses institutional data follows good data governance principles and practices. Adopting these principles and practices will improve data quality and promote responsible

---

[1] The University of Toronto uses the word division to include each of its faculties on the St. George campus, the University of Toronto at Scarborough, the University of Toronto at Mississauga and each of the shared service portfolios (Vice Presidents' portfolios, the Library, the offices of the President, Government Relations, and Governing Council).

[2] Examples include: enrolment and financial aid data; performance indicators; research funding statistics; course listings and class sizes; advancement data; data about the University's physical plant, occupancy, and capital projects; human resources data; and the University's financial statements and records.

data use throughout the University community. Ultimately, the Program should help the University increase inter- and intra-divisional collaboration, and lead to better evidence-informed decision-making.

High Quality Institutional Data

The Institutional Data Governance Program will improve the quality of the University's institutional data by systematically addressing accuracy, interpretation, completeness, and reliability. For example, inconsistency or ambiguity in data definitions can make analysis unreliable and impedes collaboration. A good Institutional Data Governance Program will include attention to metadata[3], promoting broad agreement and clarity on how key variables in institutional datasets are defined and used.

Responsible Institutional Data Use

Good data governance also promotes responsible data use. Institutional data should be managed securely throughout their lifecycle (collection, storage, access, transfer, retention, and destruction) and in compliance with relevant laws. Access to institutional data must be governed by appropriate procedures, with clear guidance as to the conditions for accessing data and the nature of their use. Access also needs to be timely. Responsible data use includes endorsing recognized standards for data literacy, data analytics and reproducibility.

The Institutional Data Governance Program's *Guiding Principles*

Data governance at the University of Toronto should be guided by the following overarching principle:

> *Institutional data are a valuable university resource over which the university community has a duty to exercise good stewardship – that is, the careful and responsible management and use of the data entrusted to its care.*

This overarching principle is expanded by the following specific principles:

Inclusivity

The development, implementation, and administration of institutional data governance guidelines and procedures should involve input from the University community, including end-users and data contributors.[4]

---

[3] Metadata are data that provide context and additional information about data.

[4] A data contributor is a person or organization/unit that submits data for inclusion in a collection of data.

Accountability and responsibility

The collection, use, and maintenance of institutional data should have clear lines of responsibility and accountability.

Transparency

Guidelines and procedures regarding institutional data governance should be freely accessible to members of the University and to the public. In addition, notification about collection, use, availability, and disclosure are important features of transparency.

Integrity

The University's data governance guidelines and processes should advance uses of institutional data that demonstrate operational, analytical, and ethical integrity, free of actual or perceived conflicts of interest.

Clarity and consistency in access

Institutional data – whether held centrally or in the divisions – should be subject to clear and transparent access guidelines that are consistent across all divisions.

Efficiency

Institutional data governance guidelines and procedures should support an environment wherein institutional data are not unnecessarily duplicated. Institutional data should be managed in an efficient fashion, taking into account human, digital, and physical resources.

Partnership

Data governance is a collective responsibility. Everyone who manages, uses, or provides data should be a partner in its governance.

Partners in Institutional Data Governance

The Institutional Data Governance Program should be integrated into the University's administrative infrastructure and draw on existing structures and processes, to the extent possible. As with all new programs, it will take time to reach maturity; however, once established, the Program's guidelines and practices will become intrinsic to the way in which the University community engages with institutional data.

As the principles of data governance emphasize, data governance should be a collective responsibility, shared in varying degrees, by everyone who manages, uses, or provides data. This responsibility will manifest itself in different ways, of course, depending on an individual's role and expertise.

Knowledgeable subject matter experts, for example, can help keep the data they work with accurate and current; data analysts can assist one another in communities of practice; data contributors can help inform decisions about appropriate or reasonable uses of institutional data; transparency and responsibility in managing institutional data will foster public trust; government oversight and smart regulation can promote inter-jurisdictional collaboration and improve outcomes not just locally, but across the sector. Thus, the University of Toronto community – broadly understood to include everyone who interacts with the University's institutional data – is a partner, to varying degrees and in various respects, in data governance.

The success of such a Program will depend upon the University's model of delegated authority and accountability in which authority and accountability typically reside close to the relevant operational unit, from departments to divisions to campuses to the institution as a whole. A network of partners should support this infrastructure.

An Institutional Research and Data Governance (IRDG) Team

At the centre of the network should be a small Institutional Research & Data Governance team. The IRDG team should focus on collaborating with, and supporting, the divisions in the development and operation of the Program. The team should promote institutional consistency, sharing and assessing best practices to ensure that the University's institutional data assets are managed so as to enhance their value, in accordance with the University's data governance principles.

A broad network of partners

An effective Institutional Data Governance Program will depend upon a broad and diverse network of units and offices. Data governance requires collaboration and coordination among a host of cross-divisional partners with distinct but intersecting mandates:

- *Information Technology* colleagues in the University's ITS portfolio and within divisions provide and support digital infrastructure and the tools and guidelines for data security, integration, analysis, and transmission. They also host, maintain, and share many sources of data;

- *Human Resources* colleagues in the University's Human Resources & Equity portfolio and within divisions are critical to building the necessary human capacity at the heart of a successful Institutional Data Governance Program. They work with divisions to hire staff with essential skills and to offer ongoing training and opportunities for professional development;

- *Library* colleagues, not least in University of Toronto Archives and Records Management Service (UTARMS), are valuable resources, advising on the creation, maintenance, storage, access, and preservation of data;

- *Vice President Research & Innovation*, the *Vice President & Provost*, the *Vice President Human Resources & Equity* and the *Research Ethics Board* engage as required where institutional data or surveys of faculty, staff, students or alumni are requested for research purposes;

- *Communications* professionals across the University assist with change management and can help communicate – and illustrate – some of the many significant benefits good data governance will realize;

- *Legal*, *Internal Audit, FIPPO, Information Security Council* colleagues, and similar resources work closely on regulatory compliance and advise on the creation, interpretation, and implementation of policies and practices;

- *Equity, diversity, inclusion, indigenous* colleagues provide resources and guidance on specific issues as they arise.

Cross-divisional partners are vital contributors to a successful Institutional Data Governance Program. The Institutional Research & Data Governance team must work alongside these partners in developing and implementing the University's approach to data governance.

Conclusion

Data governance at the University of Toronto is still in its infancy. Insightful and creative leaders at every level and across every division are already leading the way and motivating changes in the University's culture. These changes will strengthen and improve how the University community manages and interacts with institutional data, as new or emerging norms around data stewardship, access, transparency, and collective accountability gain greater acceptance. A formal Institutional Data Governance Program to help organize and support the community is overdue. Building such a Program will amplify the value of institutional data to the University community, harness those data in the service of better decisions, and advance the University's mission.

Appendix A: Committee Members

Data Governance Committee

Caroline Abrahams   Director, Policy, Planning & Systems Management, Faculty of Medicine

Sally Garner (co-chair) Senior Strategist, Operation Initiatives, University Operations & Provost

Tony Gray   Director of Strategic Research, Office of the President

Helen Lasthiotakis   Executive Director, Strategic Partnerships/Office, VP Research & Innovation

Aimy Lieu   Director, Academic Operations, OISE

Neil Neebar   Associate Registrar A&S, Designate for Deborah Robinson

Philip Oreopoulos   Professor, Economics, UTM

Deborah Robinson   Faculty Registrar & Director UG Academic, Faculty of Arts & Science

Jamie Stafford (co-chair)   Professor, Statistical Sciences, Faculty of Arts & Science

Sean Suleman   Executive Director, CAHRS & Workplace Investigation, Human Resources & Equity

Nathan Taback   Teaching Stream Professor, Statistical Sciences, Faculty of Arts & Science

Jeff Waldman   Manager Institutional Data Governance, Planning & Budget

Geoff Wichert   Divisional Reporting & Information Analyst, Faculty of Applied Science &
Engineering

Don Willison (advisor) Associate Professor, IHPME, Dalla Lana School of Public Health

Institutional Data Hub Executive Steering Committee

Vivek Goel   Vice-President Research & Innovation

Scott Mabury   Vice-President Operations and Real Estate

Kelly Hanna-Moffat   Vice-President Human Resources & Equity

Cheryl Regehr   Vice-President & Provost

Bo Wandschneider   CIO

Business Intelligence Committee Co-Chairs

Trevor Rodgers   Assistant Vice President, Planning & Budget

Towards a Data Governance Program at the University of Toronto – February 2020

Susan McCahan   Vice-Provost Academic Programs & Innovations in Undergraduate
Education