UNIVERSITY OF TORONTO

Data Classification Standard

Overview

Data classification is a fundamental aspect of data management and cyber security, helping organizations protect their sensitive information and ensure compliance with regulatory requirements.

The Information Security Council developed the University of Toronto's Data Classification Standard. The standard groups U of T data into four levels based on its importance, sensitivity and potential for misuse. The guidance is endorsed by the University of Toronto Data Governance Council (now represented online through the Institutional Data Strategy) under the authority of the Information Security Council.

Level	Definition	Explanation and examples

Level 4	Level 4 data is non-public information designated by the university that requires substantially greater protection measures than level 3 data.	Level 4 data is highly sensitive, and its disclosure poses substantially greater risk of harm to the university than level 3 data. It should not normally be stored on general-purpose systems or handled as ordinary office paperwork.
		
		Examples (not exhaustive):
		• Government-issued ID:
		  o Social insurance number (SIN)
		  o Social security number (SSN)
		  o Individual tax number (ITN)
		  o Passport number
		  o Driver's license number
		• Cardholder information for students, staff, vendors, merchants and community members as defined under the Payment Card Industry Data Security Standard (PCI DSS)
		• Bank account number for students and staff
		• Personal health information (PHI) as defined by the *Personal Health Information Protection Act (PHIPA)*
		• Biometric data
		• Passwords and credentials (e.g., system credentials, password stores, personal identification numbers [PINs], encryption keys)
		• High-risk case files (e.g., files managed by the Office of Safety and High Risk or the Community Safety Office)
		  o Investigative reports related to workplace and sexual violence or special investigations

UNIVERSITY OF TORONTO

Level	**Definition**	**Explanation and examples**

UNIVERSITY OF TORONTO

Level 3	Level 3 data is non-public information that contains personal information, as defined by the *Freedom of Information and Protection of Privacy Act (FIPPA)*, where permission to disclose has not been granted. It also includes other data the university has designated as level 3.	Level 3 data includes many types of administrative information, such as general email and business paperwork in a typical university office. Administration of the university's teaching often involves handling personal information about students and sometimes about staff and faculty. In addition to level 1 and level 2 risks, *FIPPA*-related risks also apply. **Examples (not exhaustive):** • Personally identifiable information about applicants, students, faculty, staff or donors (can include data on its own or when combined with other information) o First and last name, email address, phone number, home address o Combination of identifiers such as student/employee ID and/or name with: ♣ Sensitive demographic information (e.g., gender identity, sexual orientation, disability, Indigenous identity, racial or ethnocultural identity, religious affiliation) ♣ Student record data (e.g., student advising data, financial aid data, grades, GPA) o University photo ID • Investigative reports related to Code of Student Conduct investigations • Security camera recordings • Security event logs • Security system vulnerabilities and risk records o Vulnerability scan results o Risk registers o Data Asset Inventory and Information Risk Self-Assessment (DAI-IRSA) data o Risk assessment reports • Location data that tracks an individual's movement (e.g., IP addresses) • Administrative health information (AHI) • University contracts/agreements with third parties • Budget or financial information including non-public financial statements • Legal advice – solicitor-client privileged information • Crisis and emergency preparedness plans • System and network architecture diagrams • Detailed building and facilities plans
Level	**Definition**	**Explanation and examples**

Level 2	Level 2 data is information the university has not chosen to make public and has not designated as belonging to another level.	Level 2 data is the default category. In addition to level 1 risks, this data should not be disclosed to the general public or to anyone other than those authorized by the data owner or steward, unless or until it is designated for public release. **Examples (not exhaustive):** • Non-public aggregated data • De-identified data (see definition below) • Most unpublished research • Most course materials (including assessment and assignment questions) • Building floor plans • Internally developed custom source code
Level 1	Level 1 data is information available for broad or general public use.	Level 1 data is publicly accessible. Privacy and confidentiality are not issues; the concern is authenticity and integrity, ensuring no unauthorized additions, modifications or deletions. **Examples (not exhaustive):** • Institutional and department policies and procedures • Directory information (staff and faculty) • Course information (e.g., curriculum, fees, learning outcomes, syllabi, class schedules, course catalogues) • Published research • Public website data (cannot include any level 2 to 4 data) • Press releases • News articles • Published annual reports • External job postings, distributed • Open-source code

Data Classification guidance

This section serves as the **data classification guidance** for the university's institutional data. It includes **examples of data elements** by classification level and supplementary considerations for classifying data. It can help you determine how and why to classify your data.

Methodology

Factors considered when classifying data include:

1) **Sensitivity**: Determined by potential harm to an individual or organization in case of a data breach
2) **Regulatory and legal requirements**: Determined by e.g., the Freedom of Information and Protection of Privacy Act *(FIPPA)*, the Personal Health Information Protection Act *(PHIPA)*

3) **Personal identifiability**: Determined by whether someone is *directly identifiable*, whether data was *de-identified* or *aggregated*, etc.
4) **Criticality**: Determined by the nature of the data, e.g., whether it is proprietary information, or constitutes trade secrets or intellectual property
5) **Impact**: Determined by e.g., financial or operational risk

How to use this guidance

All staff and researchers who work with institutional and research data should be aware of its classification and handle it appropriately.

Here are some tips on how to classify your data.

- **Combination of data elements**: If a file contains data with different classification levels, assign the highest level to the entire file.
    - **Example**: If a record is broadly categorized as "legal advice: solicitor-client privileged information" (level 3) but contains data elements like passport number (level 4), the whole record must be classified as level 4.
- **Contextual sensitivity**: Individual pieces of information that are not sensitive on their own may become sensitive when combined.
    - **Example**: A list of student IDs is not sensitive on its own but when combined with names and addresses it becomes sensitive.
- **Masking sensitive data**: Data can be classified at a lower level if sensitive elements are fully or partially masked. Data trustees and their delegates (individuals responsible for overseeing access to and protection of institutional data) must ensure that masked data is properly de-identified.
    - **Example**: A report can use only the first three digits of postal codes instead of full postal codes to reduce sensitivity from level 3 to level 2.
- **Additional protection protocols**: Data trustees can recommend extra protection protocols within a classification level. This may include level 3 data elements that need to meet additional compliance requirements.
    - **Example**: Credit card information (level 3) must comply with the Payment Card Information Data Security Standard (PCI DSS), requiring encryption and restricted access.

For questions about this guidance or for help classifying examples not listed above, contact the Institutional Research & Data Governance (IRDG) Office at data@utoronto.ca for institutional data and Research Information Security at research.infosec@utoronto.ca for research data.

UNIVERSITY OF TORONTO

Key definitions in data classification guidance

General terminology

Institutional data: Comprises all data held by the university to support its administrative operation.

Data trustee: An executive officer with policy-level accountability for managing a major area of the university's data assets. They champion the collection and use of their respective data assets. Data trustees typically hold senior leadership roles (e.g., vice-president, vice-provost, dean).

Data trustee delegate: A subject matter expert who understands the meaning, context and appropriate use of the data within their domain. They oversee key aspects of data governance, including metadata management, data quality, access control and overall data stewardship. Data trustee delegates typically hold leadership roles (e.g., chief administrative officer, director, manager).

Health data terminology

Personal health information (PHI): Under the *Personal Health Information Protection Act (PHIPA)*, PHI refers to identifiable information about individuals collected or maintained by health information custodians (HICs) in relation to healthcare provision (e.g., lab tests, medication prescriptions). HICs include, for example, hospitals, labs, nursing homes and related institutions or bodies.

Administrative health information (AHI): Identifiable health information about individuals collected by the university for administrative purposes (e.g., immunization records, accommodation requests).

What is the difference between PHI and AHI?
The key difference between PHI and AHI lies in the collector and intended use.
PHI is collected by a health information custodian (HIC) or its authorized agents specifically for delivering healthcare services. This includes clinical records such as lab results or prescriptions and is governed by *PHIPA.* AHI, by contrast, refers to identifiable health-related information collected by the university for non-clinical, administrative purposes – such as immunization records for enrolment or documentation supporting accommodation requests.
Important: Outside of university-operated health and wellness or specialized clinics, any health information collected should be treated as AHI. Even if the data originated as PHI, once voluntarily submitted to the university for administrative use, it is no longer governed by *PHIPA.*

How should health card information (e.g., Ontario Health Insurance Plan, University Health Insurance Plan) be classified?
The classification depends on the context in which the information is collected, stored and used:

- **Level 4 (PHI):** Applies when health card information is collected by a health information custodian or its authorized agents for the purpose of delivering healthcare services (e.g., clinical care at a university-operated health clinic).

- **Level 3 (AHI):** Applies when health card information is collected by the university for administrative purposes unrelated to clinical care (e.g., verifying eligibility for health coverage, enrollment requirements).

UNIVERSITY OF TORONTO

Key definitions in data classification guidance (cont'd)

Data minimization terminology
Aggregated data: Information that has been grouped together in a way that protects individual privacy. Instead of showing details about specific people, the data is presented as totals or averages. For example, if a group includes six or fewer individuals, or if a percentage is exactly 0% or 100%, the data is still shown in a way that prevents anyone from being identified. **De-identified data:** Data that has been modified to prevent identification of an individual – whether on its own or combined with other data sources. This process is essential to ensure that data cannot reasonably be linked back to an individual. For detailed guidance, refer to the Information and Privacy Commissioner of Ontario's De-identification Guidelines for Structured Data.