THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

CREATE CHANGE

Data Governance Essentials Handbook

What is Data Governance?

Definitions of Data Governance are numerous. However, they agree on a key principle - that it's dedicated to the organisation of people, processes, and technology to enable effective data management.

Data Governance includes:

- Understanding the **roles and responsibilities** of those who use data.
- The **technologies, tools and systems** to support capabilities in the areas of data management, quality, security, integration, discovery, accessibility, and availability.
- **Processes and controls** implemented to support the consistency, integrity, usability and privacy of data.
- **Policies, procedures and standards** to support regulatory compliance, ethics, and provide clarity over responsibilities.
- **Data quality** - including metadata and information security - to ensure reliability of data.

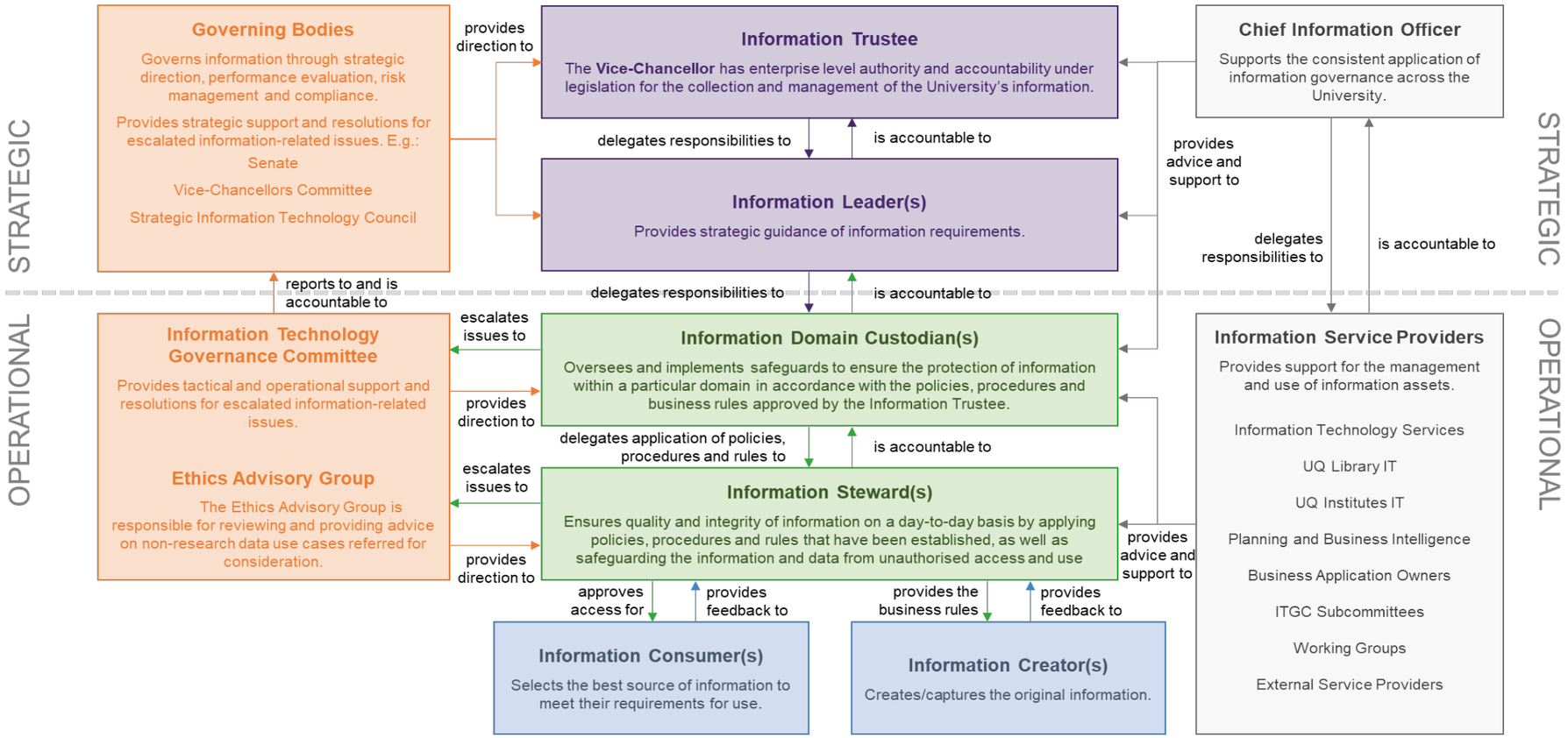
Roles and Responsibilities

Who owns data at UQ? We all have a role to play when it comes to data and information. However, we as an organisation or as individuals collecting data from our community, do not 'own' the data.

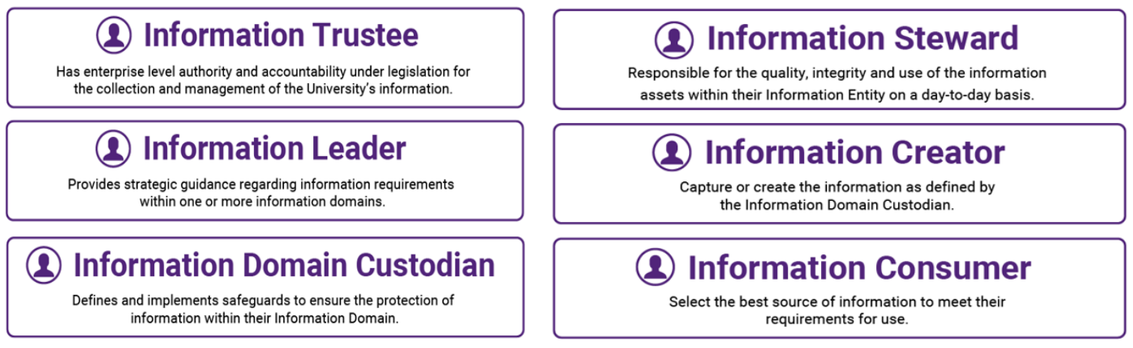
Legislation like the European Union's (EU) General Data Protection Regulation (GDPR) attempts to answer this question, as it's very clear about who owns the data: the person the data represents.

The organisation that collects the data must act as a steward of this data, but in reality, there is no ownership of personal information.

UQ has clearly defined roles and responsibilities for data, known as 'decision rights'.



Right: *Decision Rights Model, outlining the hierarchy of various data roles at UQ, and the relationships between them.* ***Cover:*** *Photo by fabio on Unsplash.*



Above: *Roles for data at UQ*

Decision Rights

The roles and responsibilities outlined previously are referred to as 'decision rights'. **Decision rights across the University is a key enabler of good data governance to support efficient decision making regarding the management of data and information through its lifecycle.**

The decision rights model outlines a hierarchy of relationships, describing levels of accountability and responsibilities, and provides a reference point for data governance decisions.

From this example, we can see that these roles work together to ensure decision-making about data management is conducted efficiently.

Decision Rights Model

The Decision Rights Model shows the relationship between different roles. On this diagram (above), you can see the various roles we have discussed, with the relationships between each clearly outlined.

Example of Decision Rights

We will work through an example of how this model is used. **Using student admissions data, let's understand the responsibilities of each individual acting in their role.**

At the bottom of this diagram, we have Information Creators. In this example, the Creators are Admissions staff, responsible for entering data into SI-net. Also at the bottom of this diagram are Consumers. This might be a School Manager, looking to access student admissions data to generate a report on enrolments.

At the next level in the hierarchy are Stewards. In this example, this is the Deputy Director, Academic Services Division. **They are responsible for reviewing and approving requests to access data within their Information Entity** which is 'student admissions'.

The Information Domain Custodian in this case is the Academic Registrar, and the Information Leader is the Deputy Vice-Chancellor (Academic). At the top of the hierarchy, the Vice-Chancellor is the Information Trustee, as for all Information Domains.

Finally, we can see ITS Enterprise Support Services is the **Information Service Provider, providing the necessary technical support.**

Definitions

Within Data Governance, there are a number of concepts, theories, and definitions. It is important to understand these, as they are key to various processes and procedures that you will encounter. Below, read some key terms within Data Governance, and their definitions.

Term	Definition
Data	Raw data describes data in its most basic digital format. Data is raw, individual facts that need to be processed. Data can be structured (eg. in a database or a spreadsheet, including student, research, or financial data), or unstructured (eg. audio, vide, unstructured text that does not have a pre-defined model or structure).
Information	Information is data with context. It includes, but is not limited to, physical (e.g. paper records) or digital files (e.g. email, voicemail, meeting minutes, video and audio recordings) in any format (e.g. PDF, .wav, .docx, or .jpeg) and data recorded by University applications (often in a database of some form).
Knowledge	Knowledge is synthesised information. This means that it is personal to us - our brains store information and use it to make judgements about the world (eg. where to eat lunch, based off how many good meals we've had there).
Records	Records aren't just collections of data—they comprise of the content, context, and structure necessary to provide sufficient evidence of a business activity. As a collection, records are fundamental to our institutional memory and contribute directly to our understanding of UQ in the past, present and future.
Data governance	Data governance is a collection of practices and processes, which helps to ensure the formal management of data assets within an organisation. Key elements of data governance: • Data categorisation • Data quality and integrity • Data ethics • Information security classification • Clearly defined roles and responsibilities
Data management	A collection of capabilities delivered through people, processes, and technology to ensure the confidentiality, integrity, availability, quality, and security of our information throughout the information lifecycle.
Personally Identifiable Information (PII)	Any information that can be used to identify a person. This includes: student numbers or staff IDs, date of birth, medical records. If shared with an unauthorised person, PII can be used to commit identity theft and access resources and credit under another person's name. As a result, it's important to ensure the principle of 'least privilege' is followed, with access granted only to those for whom it's essential.
Information management lifecycle	Data is managed in all phases of the information management lifecycle. The lifecycle sets out the process for consistent management of information, from creation to final disposition. The information management lifecycle at UQ includes the phases outlined below: • **Plan and design** information appropriately; • **Create, capture and classify** information adequately; • **Store and secure** information appropriately and securely; • **Manage and maintain** information in line with external and internal policies and expectations; • **Share and reuse** information where appropriate; • **Retain and archive** information for a minimum period; and • **Dispose of and destroy** information correctly. It's important to individually define these specific lifecycle phases, so we all have a shared understanding and speak a common language. It also means we can define controls to protect and best govern data for each stage, as the requirements may vary between stages.

Foundations of Data Governance

Data Governance is a collection of practices and processes that provide a framework for the methods, technologies, and behaviours that support the sound management of data, impacting all aspects of organisational processes, decision making and actions. However, as you can see below, improvements to Data Governance must be made holistically.

Challenges	Description
Often considered a technology problem	Effective data governance requires the use of good tools; however, the use of good tools does not guarantee effective data governance.
Value of data not understood	A good data governance and management practice can bring many benefits to an organisation. However, it is important that benefits are truly understood, otherwise, they may not be embedded in day to day processes.
Roles and accountability	Knowledge is synthesised information. This means that it is personal to us - our brains store information and use it to make judgements about the world (eg. where to eat lunch, based off how many good meals we've had there).
Silos of data and departments	A huge hurdle to data governance is how data and departments are structured. Often datasets are locked away, only accessible by certain teams. Different departments also operate in entirely different ways and may have no knowledge of the available data and potential value it holds.
System instead of data thinking	People tend to understand information systems and the business and technical responsibilities associated. But one information system could hold many datasets that may fall under the responsibility of more than one person.
Culture	People and/or teams that create data, often feel that they 'own' the data. This is a major hurdle to enable visibility and access to data enterprise-wide.
Lack of standardised processes	Many organisational units will already have good data governance practices in place, but the challenge can be the lack of consistency and they may not be formally documented.

Benefits	Description
Greater data quality	A significant focus of data governance is improving data quality. A data governance practice can introduce initiatives to mature community skills, processes and technologies to address data quality.
Data analytics & operational efficiency	UQ is investing in advanced analytics capabilities and sound data governance practices are foundational to the sustainability of these new and rapidly evolving capabilities. Increasing visibility and controlled access to more data enable our analytics communities to generate meaningful insights to support decision making or identify risks.
Discovery, findability and access to data	As the University relies more on data to improve user experiences and inform decision-making, knowing which data is available, understanding its meaning, and the ability to easily access this data in a controlled way is critical.
Better decision-making	One of the key benefits of data governance is better decision-making. This applies to both the decision-making process, as well as the decisions themselves. Well-governed data is more discoverable, making it easier for the relevant parties to find useful insights. It also means decisions will be based on the right data, ensuring greater accuracy and trust.
Responding to regulatory requirements	Regulatory scrutiny, both domestically and internationally, of protection of data is becoming more intense, and community expectations are rising. Being disciplined in the ways we manage our data, enables us to satisfy regulatory requirements related to data, ultimately reducing cost and risk to the University.
Information (Cyber) Security	Being the target of a cyber-attack is inevitable for a large university such as UQ. Data governance can support cybersecurity initiatives by identifying our critical data assets and determining minimum controls that should be in place to protect them.

Information Management Lifecycle

When managing data at UQ, we need to consider what is necessary at each of the seven stages of the lifecycle.

Plan and design

While Plan and Design is discussed first, you can see it sits in the centre of the lifecycle diagram. This is because the activities in this stage should also be revisited at each subsequent lifecycle stage.

At this stage, the following questions are asked:

- What is the purpose of the activity?
- How will data be collected?
- How will it be used?
- Who will have access?
- Will any interventions be implemented?

For research data, planning is facilitated by Data Management Plans. This is considered 'best practice'.

Create, capture, and classify

- When capturing personal data, consideration to consent and privacy requirements is needed.
- Information creators have a responsibility to ensure the information they capture is complete and accurate. To ensure data remains of high quality, it should be captured correctly to ensure the reliability for downstream uses.
- The correct security classification should be applied at the time data is captured or created. The responsibility for this task falls upon the Information Creators.

Store and secure

In this stage, we must decide upon an appropriate location to store the data. This will be influenced by the purpose for the data use (such as for research purposes, or a business need), the type of data, the Information Security Classification, and any relevant regulatory requirements.

- UQ has many storage options available. Most of our data is contained within specific systems, on networked drives, in cloud storage, and records management systems. UQRDM is the preferred storage option for research data.
- Storing data outside of UQ approved systems may be in breach of relevant privacy legislation, and unethical as exposes data to unnecessary increased risk and abuses the trust of data subjects.
- Any data that is not considered public must be stored in a way that allows for access to be controlled. Access should be granted judiciously and revoked when no longer required.

Maintain and manage

Once the information is stored, it then needs to be Maintained and Managed. All UQ staff need to manage and maintain information in line with external and internal policies and expectations.

- The Information Governance and Management Framework outlines responsibilities for Information Consumers, Information Creators, Information Stewards, Information Domain Custodians, Information Leaders, the Information Trustee, and other relevant actors.
- Information Service Providers are responsible for the maintenance of UQ's information systems.
- Information Stewards have a responsibility to ensure data quality and integrity requirements are met.

Share and reuse

- Information Stewards are responsible for approving or rejecting requests to access data.
- UQ Corporate Data Sharing Agreements are used to help manage sharing of non-research data.
- Access to research data depends on a number of factors. For information, you should generally contact the project's Lead Chief Investigator.
- Informal sharing of data between individuals presents an ethical concern as UQ has oversight to ensure the use of data is appropriate and just.

Disposal or Archival?

UQ has protocols for retaining data, including time requirements for certain records, which are outlined on the Records Management Service website (uq.edu.au/rams).

Archiving is when data is moved to a different location because it is no longer regularly required but must be kept for knowledge and/or legal reasons.

Rules apply for checking the record value and sensitivity level to determine if the data must be kept for a certain period to meet legal retention responsibilities.

Depending on the record value of the data, some platforms are recommended for keeping the information for its legally required retention time, so that it remains protected, yet accessible by the right people and readable for the time it needs to be kept.

If data can be disposed of, review the defined data storage and retention policy before you dispose of data using the file deletion tools in your storage platform.

Some platforms retain the deleted data for a specified time to allow for accidental deletion recovery, before permanently deleting it. Refer to the documentation for your storage platform for any retention periods.

Retain and archive

The Retain and Archive phase encompasses the period after you've used the data, where you may need to retain it for reference or as a record.

- Archival must balance the benefits of retaining a transactional record with the risks associated with storage.
- Archiving data at UQ is governed by the relevant legislation and policies. May also be historical and cultural justifications for archiving data – it may be unethical not to archive!
- The Information Stewards and Information Domain Custodians will respectively review and approve requests for archiving data within their information domain.

Dispose and destroy

The final lifecycle phase is Dispose and Destroy. When you dispose of data it needs to be done appropriately, and following the relevant legislation and policies.

- Disposal of data should be done appropriately and in accordance with relevant legislation and policies.
- Information Stewards and Information Domain Custodians are responsible for the review and approval of disposal requests.

Where do I store my files?

Recommended storage platforms are:

OneDrive: All UQ staff have access to 1TB of self-managed storage on OneDrive. Recommended for: working documents, shared documents, and records.

Research Data Manager (UQ RDM): UQ RDM is specifically designed for research projects and provides 1TB of storage per project, which can be increased on request. Recommended for: research data

AARNET CloudStor: Cloud storage, able to be accessed by external collaborators. Recommended for: shared files.

TRIM: Used for document control and full lifecycle management of UQ documents. Recommended for: records.

Network (Shared) Drives: These drives can be seen by all staff members in a work group who have access to the shared folders. Recommended for: shared documents.

Right:
Information Management Lifecycle

Use Cases

Below, read some real-world examples of situations that would've been aided with improved Data Governance.

Data Quality

What happened?

- The Mars Orbiter was supposed to be the first weather observer on another planet. However, as it approached Mars on 23 September 1999, it vanished in space.

Why did it happen?

- The software controlling the orbiter's thruster calculated the force the thrusters needed to exert in pounds. However, a separate piece of software took in the data assuming it was in the metric unit: newtons.
- This occurred as during the design phase, the propulsion engineers at Lockheed Martin in Colorado expressed force in pounds. While it was standard practice to convert to metric units for space missions, this did not occur. However, engineers at NASA's Jet Propulsion Lab assumed the conversion had happened.
- This navigation mishap pushed the spacecraft dangerously close to the planet's atmosphere where it presumably burned and broke into pieces, killing the mission on a day when engineers had expected to celebrate the craft's entry into Mars' orbit.

What were the consequences?

- Significant loss of money. According to NASA, the cost of the mission was $327.6 million total for the orbiter and lander, comprising $193.1 million for spacecraft development, $91.7 million for launching it, and $42.8 million for mission operations.
- World-wide news coverage.

How could better Data Governance Practices have helped avoid this?

- Data quality - consistency across fields, definitions, and units of measurement.
- People and communication - managing change and communication safeguards against costly and dangerous mistakes.

Improper Disposal of Sensitive Data

What happened?

- A researcher, who goes by d0tslash on Twitter, found unencrypted video of US Fort Huachuca military police officers at work on a used bodycam he purchased on eBay. He then posted this finding on Twitter.
- Used body cameras are commonly for sale online, and several other hackers followed suit by purchasing decommissioned police cameras from eBay to see what they could find.
- The type of footage these subsequent hackers found ranged from traffic stops, responding to calls at retail stores, calls to houses, etc.

Why did it happen?

- Data was not destroyed properly before retiring and on-selling equipment.
- There was no encryption on the devices.

What were the consequences?

- Reputation damage.
- Leaked footage of police at work.
- Sensitive military information/secrets could possibly be obtained from this.

How could better Data Governance Practices have helped avoid this?

- Better data literacy and understanding of the information lifecycle.
- Understanding of dispose and destroy needs.
- Understanding of device retirement implications.
- Having policies and procedures around data handling throughout the information lifecycle, specifically regarding disposing and destroying data, and retiring devices that once held data.
- Oversight and understanding of the data being collected, its Information Security Classification, and the protection/handling controls required based on the Information Security Classification.

Previous page: *Photo by markusspiske on Pixabay.* **This page:** *Photo by TheDigitalWay on Pixabay.*

Key Policies

The University is required to meet legislative and regulatory requirements that relate to the management of information across all departments. Below, we've listed some key federal and state acts and policies that we need to comply with.

Federal Policies

- Education Services for Overseas Students Regulations 2001
- National Code of Practice for Providers of Education and Training to Overseas Students 2018

Queensland Policies

- Information Governance Policy
- Records Governance Policy
- Queensland State Archives Retention and Disposal Schedule
- General Retention and Disposal Schedule (GRDS)
- Information Security Policy (IS18:2018)
- Metadata (IS34)
- Information Asset Custodianship Policy (IS44)

Information Management Policy

The Information Management Policy outlines the requirements for the governance and management of information at UQ, informing the development of supporting frameworks and procedures.

The Policy has three underpinning procedures.

It is also supported by the Information Governance and Management Framework, which describes our obligations throughout the information lifecycle and outlines the decision rights model in greater detail.

Data Handling Procedure

The Data Handling Procedure outlines handling requirements for all information in digital format at UQ. If you manage a UQ information system, please familiarise yourself with the procedure on the Policy and Procedures Library.

What Does the Procedure Contain?

The Data Handling Procedure outlines minimum handling requirements for data, through each Information Lifecycle phase. The procedure itself is technical, however key aspects you should be aware of are outlined below.

- **Plan and design:** All data must be assigned an appropriate Information Domain, Information Domain Custodian and Information Steward as per the Information Governance and Management Framework (for research data, a First-named Chief Investigator and Head of School or Institute). Considerations should be given to minimising the number and types of roles which require highly privileged access to classified data.
- **Create, capture and classify:** Data quality at the time of creation is the primary responsibility of the Information Creator and must be compliant with the overarching policies and procedures. The individual who created the data must be identified and recorded where possible.
- **Store and secure:** Systems should be designed and configured following the principle of "least privilege" – users and systems should only be given access to the data required to execute their responsibilities. Single Sign-On (SSO)/UQ Authenticate should be used where possible. Policies and procedures relating to employee terminations, resignations or changes in responsibilities should carefully and thoroughly consider issues of data access control and retention.
- **Manage and maintain:** Details the requirements for access control policy reviews, pen testing of systems, surveillance response times, vendor contract reviews, network audits, and information security classification reviews.
- **Share and reuse:** Outlines requirements for copyright, necessary protections, storage, and permissions, during sharing and transport.
- **Retain and archive:** Provides requirements for appropriate archival of data and information.
- **Dispose and destroy:** In tandem with retain and archive, details how to appropriately dispose of data and information.

Why is the Data Handling Procedure Important?

The Data Handling Procedure was developed in response to threats to UQ's information assets.

The Procedure defines how to protect and handle data and information at UQ, outlining handling requirements for data, information and records in digital/electronic format at UQ.

It also provides clear direction and a more effective approach to securing and safeguarding valuable University digital information assets.

The Procedure was informed by a comprehensive threat analysis of data and information at UQ.

Where Can I Access the Procedure?

The procedure is now available in the Policy and Procedures Library (6.40.03b).

It should be read in conjunction with the Information Management Policy and the Information Governance and Management Framework.

The Data Handling Procedure applies to all University staff and students, as well as any groups or individuals authorised by UQ to access University information.

How Can the Data Handling Procedure Protect UQ?

The last decade has seen a sharp rise in cyber security attacks on organisations.

According to The Office of the Australian Information Commissioner, between July and December 2019, there were 537 notified data breaches in Australia. Whils some of these breaches can be the result of human error (eg. sending an email to the wrong recipient, forgetting to use the BCC function in mass emails, unintentionally disclosing information), others, particularly those against large organisations, are from hacking attempts. These breaches can result in large, unauthorised breaches or losses of personal data, making staff and students vulnerable to identity theft and other hacking attempts.

To ensure that University information, systems and intellectual property remain safe, it's important to control and/or mitigate internal and external threats to information assets.

Although this is an enterprise-wide procedure, the audience is mainly the technical community.

It is important to note that the Data Handling Procedure outlines the 'ideal' state for controls at UQ and may include capabilities that are not currently implemented or available.

Where Can I Find Out More?

If you have further questions, or are unsure of how the Data Handling Procedure or Enterprise Data Ethics Framework impacts you or your work, please visit data.uq.edu.au, or contact the Enterprise Data Governance Program for advice at datagovernance@uq.edu.au.

Data Ethics

Data ethics is concerned with the moral standards applied, and assessments made, when working with data. It's balancing 'what can we do' with 'what should we do'.

Just because we have the data available or capability to leverage said data to gain insight, does not mean all applications are proper or justifiable.

Ethics help us navigate the shades of grey between what is allowable under legislation and policy, and what is considered appropriate among the wider UQ community.

Overlooking ethical considerations may lead unintended negative consequences for both individuals and the University's reputation.

Why is Data Ethics Important?

As data ethics is concerned around what organisations can and cannot do with the data they collect, good data ethics is important to ensure the UQ community can retain trust in how UQ is using their data.

Good data ethics facilitates privacy, transparency, fairness, and regulatory compliance, and reassures the UQ community that their data is safe.

Good data governance supports ethical practice – provides guidance, oversight and ensures the right people are making decisions.

Overlooking ethical considerations may lead unintended negative consequences for both individuals and the University's reputation. This is important to ensure the UQ community can retain trust in how UQ is using their data.

Data ethics is concerned with (but not limited to):

- Privacy: reasonable effort needs to be made to preserve privacy, especially if others are accessing this data (eg. do you have PII you need to de-identify?)
- Consent: is consent needed?
- How is the data you are working with being used/ interpreted by others (eg. downstream uses)?
- Identifying potential harm and bias
- Balancing benefits for stakeholders.

Enterprise Data Ethics Framework

There is an additional framework that covers enterprise-wide expectations for the ethical use of non-research data across UQ.

The Enterprise Data Ethics Framework (EDEF) is a helpful guide to assist you in your daily decision making.

How Will the Enterprise Data Ethics Framework Help Me?

The EDEF is not designed to limit UQ staff by introducing another set of complex rules to remember – rather, it is there to help you make better decisions, and navigate the complex and constantly shifting data security landscape with confidence. To help you apply the EDEF to your work, you can access an interactive Data Ethics Assessment Tool and Handbook on data.uq.edu.au.

Regardless of your role at the University – admin officer or data analyst – you might handle a range of confidential information. You may consider ethical concerns to be limited to personally identifiable information (PII), however, they may also be relevant (but not limited) to financial, commercial, and contractual information.

The EDEF aims to assist you to consider potential ethical issues during the collection, use, sharing, archival, and disposal of information (the Information Lifecycle). It will help you navigate the space between legislation and policy, and what you should ethically do.

The result is a risk mitigation strategy for both individuals and the University, helping you consider how you work with data, and the risks and impacts of your work.

This is part of a continuing commitment to make UQ a dynamic community that responds to the shifting challenges of our world.

Visit the data ethics hub for more information.

Information Security Procedure

The Information Security Procedure provides requirements for classifying digital and physical information at UQ.

Information Security Classifications

Information security classifications are designed to categorise UQ's information assets (physical or digital) based on its confidentiality, availability, and integrity needs. A holistic approach will consider the impact a compromise to the information asset might have on the University's broader profile. **There are five Classifications:**

- **UNOFFICIAL** For personal communications that are unrelated to the University. For example, a staff member is emailing their partner about their dinner plans. Because the information is not associated with the University and does not contain confidential personal information, it does not pose a potential threat. Therefore, it is classed as UNOFFICIAL.
- **OFFICIAL - PUBLIC** Information about the University that is available to the public, without a UQ login. For example, a course coordinator emails a tutor the next semester's course outline, which has updated assessment due dates. Because the rest of the information is already published online, available to the general public (not requiring a UQ login), it is considered to have an insignificant impact in the event of an accidental leak or malicious breach. Therefore, it is classed as OFFICIAL – PUBLIC..
- **OFFICIAL - INTERNAL** Information that is only available to those within the UQ community, but would be unlikely to cause harm if accidentally released publicly. For example, a supervisor is sending a manager the team leave calendar for the next three months. The information is of a private nature, and will not be accessible to the general public, with access restricted by business (or academic, or research) need. However, it doesn't contain higher-level human resources information (eg. Tax file numbers, bank account details), and if it were breached, would be unlikely to harm the individual or the University. Therefore, it is classed as OFFICIAL – INTERNAL.
- **SENSITIVE** Information that would cause harm, either to UQ or an individual, if accidentally released publicly, and is of a private nature. For example, a researcher is sending findings to a colleague. The research data is not yet published. The data is potentially being used for an important publication, therefore it needs to be appropriately secured. To maintain the safety of their work, it is classed as SENSITIVE.
- **PROTECTED** Information that would cause serious harm if accidentally released publicly, and is of a private nature. For example, someone wishes to email medical records to a colleague, to set up up support mechanisms for a student. This information could cause serious harm (breach of privacy) to the student if released publicly, and access is only for specific purposes, so it needs to be appropriately secured. As a result, it is classed as PROTECTED.

To assist in deciding which classification to apply, you can also use the decision-making tree (opposite page).

Why Are Information Security Classifications Useful?

Information security classifications inform the implementation of appropriate security and other mechanisms to control the information from being leaked, manipulated, or becoming unavailable. They're an important feature in the University's ongoing commitment to better Data Governance practices, embedding protocols for improved decision making in daily work.

All information and data at UQ must have a classification applied, typically applied during the 'create, capture, and classify' stage of the information lifecycle.

Where Can I Read More?

The purpose, process, requirements, roles, responsibilities, and accountabilities of the information security classifications can be found on the UQ Policies and Procedures Library (6.40.02).

Office 365 Sensitivity Labels

Data security is an ongoing concern at both UQ and other institutions around the world, particularly as threats of phishing and other scams and hacking attempts become more complex. As UQ staff handle confidential information in a variety of respects - including personal information regarding students, and research data - it is vital we continually work to prevent this, protecting the confidentiality of the information and the integrity of the University.

This is something that Office 365 Sensitivity Labels can help achieve. **In line with the information security classifications, there are five labels:**

- UNOFFICIAL
- OFFICIAL - PUBLIC
- OFFICIAL - INTERNAL
- SENSITIVE
- PROTECTED

Once Office 365 Sensitivity Labels are enabled for you, **every time you create an Office 365 document or email, a label will be applied**. The default label is OFFICIAL - INTERNAL. If this is incorrect for the information you are dealing with, you will need to assign a label.

Additional controls to protect the information (e.g. encryption, restriction on access/ sharing) are also applied to some labels. This helps protect information and enables better clarity over the confidentiality of documents.

To find out more about Office 365 Sensitivity Labels, including use cases, visit data.uq.edu.au.

WHAT **INFORMATION SECURITY CLASSIFICATION** SHOULD YOU APPLY?

Was the information created, sent or received as part of your work for UQ?

- NO → **UNOFFICIAL**
- YES →

Would unauthorised disclosure, modification or destruction of the information reasonably be expected to cause ***serious harm*** to UQ, another organisation or an individual?

- YES → **PROTECTED**

  Information that if breached owing to accidental or malicious activity could reasonably be expected to cause serious harm to UQ, another organisation or an individual if released publicly.

  The information has a restricted audience, and access must only be authorised based on very strict academic, research or business need.

  Examples: confidential legal documents; medical records; work cover forms.
- NO →

Would unauthorised disclosure, modification or destruction of the information reasonably be expected to cause ***harm*** to UQ, another organisation or an individual?

- YES → **SENSITIVE**

  Information that if breached owing to accidental or malicious activity could reasonably be expected to cause harm to UQ, another organisation or an individual if released publicly.

  The information has a restricted audience, and access must only be authorised based on strict academic, research or business need.

  Examples: business documents such as budgets and business cases; unpublished research data; personal identifiable information (e.g. human resources data and tax file number); student and staff records.
- NO →

Is this information intended to be publicly available?

- NO → **OFFICIAL - INTERNAL**

  Information that if breached owing to accidental or malicious activity would be unlikely to cause harm to UQ, another organisation or an individual if released publicly.

  The information has a restricted audience, and access must only be authorised based on academic, research or business need.

  Examples: OMC design templates; internal correspondence; administrative documents (e.g. organisation structure documents or leave calendars).
- YES → **OFFICIAL - PUBLIC**

  Information that if breached owing to accidental or malicious activity would have an insignificant impact. The information is authorised for public access, however it may not be made available in the public domain.

  Examples: published course outlines; policies and procedures; research data made public; official publications, such as the UQ strategy.

For more information or help with classifying your data, contact **datagovernance@uq.edu.au**

The University Of Queensland AUSTRALIA
CREATE CHANGE

Above: Information security classifications decision-making tree. This model assists in determining which classification you should apply to your data, information, and communications.

Data Categorisation

UQ collects data across three key categories:

- **Corporate:** Corporate data includes data that UQ collects, generates and uses as part of administrative activities. This includes data used for student and research management, finance, legal, properties and facilities, and other activities. Corporate data is mostly stored within systems and networked drives.
- **Teaching and Learning:** Data about UQ's teaching and learning activities – including course, program & curriculum information. Most Teaching and Learning data is contained within systems, including Blackboard, SI-net, and Jac.
- **Research:** UQ has custodianship over research data, but day-to-day responsibility lies with Lead Chief Investigator. There are significant variations in data types, size of datasets, storage options, and processes for controlling access and sharing data.

Information Domains

Data collected can be broken into 'information domains'.

An Domain is a broad category or theme under which UQ information can be identified and managed.

At UQ, we have many information domains. These include:

1. Curriculum
2. Teaching and Learning
3. Research
4. Research Management
5. Finance
6. Student Management
7. ICT
8. Legal
9. Strategy and Planning
10. Governance and Risk
11. Advancement
12. Services
13. Human Resources
14. Marketing
15. Facilities
16. Library
17. Organisation
18. Person
19. Common

Strategic guidance regarding information requirements within domains is provided by Information Leaders. As part of their duties, Leaders approve policies, procedures and rules to ensure the protection of information within their Information Domain. Each domain will have one Information Domain Custodian responsible for it; the Information Domain Custodian is responsible for defining and implementing these safeguards.

Data Quality and Integrity

Data quality refers to the quality of data. There are six dimensions of data quality. Considering and addressing these dimensions when creating and saving data will help improve data quality:

- **Accuracy.** How well does a piece of information reflect reality?
- **Completeness**. Is it comprehensive / are all required fields filled?
- **Consistency.** Does information stored in one place match relevant data stored elsewhere?
- **Timeliness.** Is it available when you need it?
- **Validity.** Is it in the right format / follow business rules?
- **Uniqueness.** Is there a single 'source of truth', or have you accidentally recorded the same data multiple times?

Data is valuable, however its value is heavily determined by its quality. Good quality data provides confidence in inferences, while poor quality data hampers opportunities to utilise it.

While data is valuable, its value is heavily determined by its quality. While good quality data allows for confidence in inferences, bad quality data can provide incorrect results or be less reliable.

By considering and addressing these dimensions when you create data, you can help increase overall data quality.

Relationship Between Data Quality and Integrity

We covered earlier the aims of data governance and management: a significant focus of data governance is improving data quality.

While we've just gone over quality, it's important to also understand **data integrity**. There is a strong relationship between data quality and integrity.

Data quality is used to describe the degree to which data is accurate, complete, timely, valid, unique and consistent with business requirement rules. Data quality can be improved by deciding what is wanted from it, implementing new processes, and using technology to the maximum possible advantage. A significant focus of data governance is improving UQ's data quality.

Data integrity relates to the reliability and trustworthiness of data. It requires maintenance of the data to ensure it remains accurate and consistent throughout the information lifecycle. Data integrity can be improved by change controls, regulat archiving and auditing, and striving for accuracy.

As a UQ Staff member, you have certain data roles and responsibilities. By using the frameworks and concepts outlined in this handbook, you can navigate them with confidence.

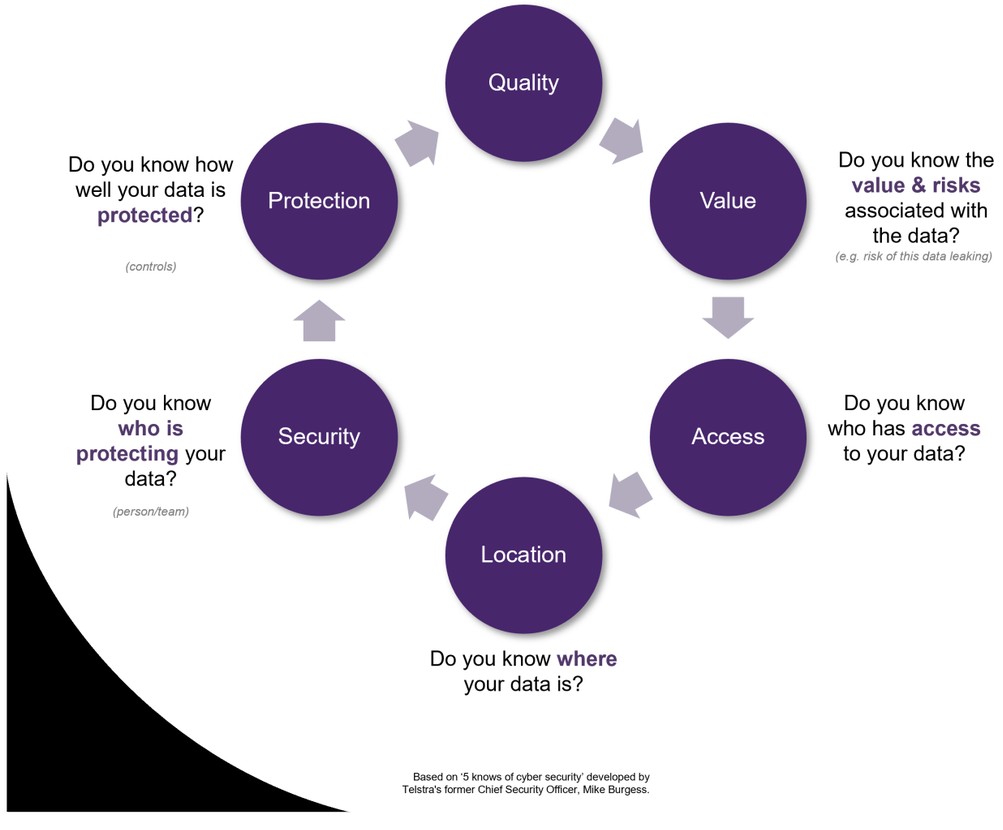
Photo: *The 'Six Knows'.*

The 'Six Knows'

What should you know about your data?

- Do you know if your data is of suitable **quality** to support decisions? Of course, it is hard to look back retrospectively at these things. However, we need to ensure that the above is adhered to with new data – if we are able to do that, we're one step towards better data governance.
- Do you know the **value & risks** associated with the data? For example, the risk of data leaking. Here, think about metrics involved, with continuous monitoring and evaluation; and whether how to understand and interpret your data? is clear to your users.
- Do you know who has **access** to your data? Here, think about whether you know who is currently accessing your data (could be through PBI Portal, Data Services, etc.). Do people have access to data they shouldn't?
- Do you know **where** your data is? Is your data saved in one location or across multiple systems? Do you understand the access and security around these locations? Do you follow a naming convention? What happens if you need to retire your data?
- Do you know **who is protecting** your data? It could be a person, or a team. Know who is protecting your data. ITS can protect the actual copy of the data (e.g. encrypting backups / database tables); however Information Stewards must protect access (be on top of access control). It's important to know the processes, procedures, and automated methods in place to ensure the security of data.
- Do you know how well your data is **protected**? What controls are in place? Here, think about security assessments: do people have access to data they shouldn't? Is the data that is supposed to be protected, appropriately protected? Have you classified your data? (using Information Security Classifications). The Classification will help determine what sort of security controls are required.

It's important to emphasise that improved data governance is a journey – these are ideas to work towards, and it will be easier to answer these questions, with both time and help from the Data Strategy and Governance team.



THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

CREATE CHANGE

Questions?

Data Strategy and Governance is here to help.
Email: datagovernance@uq.edu.au
Web: data.uq.edu.au