Project: Supporting Implementation of Information Classifications

Monash University, University of Tasmania

1. Information Classification Schemes and the Institutional Underpinnings

1.1 Background

In the Active Data Management (ADM) working group outputs, insights and recommendations revolved around enabling researchers to make informed decisions about appropriate infrastructure for their needs in line with the soft infrastructure of an institution. Similarly, in the Sensitive Data (SD) working group discussions, there was acknowledgement that to make informed decisions about how to manage sensitive data, there first needed to be the development and implementation of a Data Classification Framework. Both working groups highlighted that due to the collaborative nature of research, classifying data to enable decision requires consideration of other Institutions and Call to Actions in both groups called for greater sharing between Universities.

Leading into this project the University of Tasmania was in the process of developing and implementing a new data classification framework (to replace a draft version for research developed in 2019 but not implemented), while both Monash University and the University of Melbourne were looking at mechanisms to further implement their existing frameworks into the research data ecosystems. In 2020 Monash revised the Electronic Information Security – Information Classification Procedure, a single procedure that encompassed both research and enterprise data. The revision of the existing procedure was initiated following a need to highlight the unique differences between institutional and research data to enable clear directions on ICT platforms available for use at the University following the implementation of research capabilities such as the Monash Secure eResearch Platform (SeRP). As a result of the procedure updates, additional materials such as an internal enterprise managed approved services list and public facing security classification scheme for research data (managed by the University Library) were developed.

This situation led to the development of a joint IU project between Monash and UTAS (with input from Melbourne) on the development and implementation of Data Classification Frameworks, testing various aspects of the Institutional Underpinnings framework, especially in ADM and SD.

1.2 Data Classification Framework development:

As part of this project, UTAS developed a Data Classification Framework, using the learnings from stage 1 of the Institutional Underpinnings project, and the draft elements, to inform the process. We then compared our draft framework with the existing frameworks used by Monash and Melbourne, to ensure consistency across sites and to identify any deviations. This document covers the development of the UTAS and Monash Framework as well as some key intersections with the Elements.

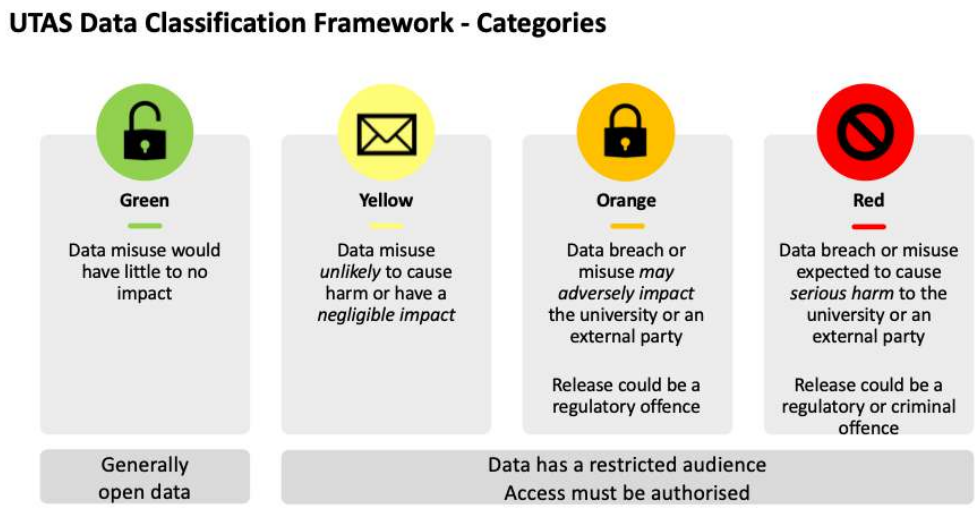
1.3 Local Governance

	University of Tasmania	Monash University (at the time of framework update)
Project Lead	Associate Director Digital Research	Senior Manager Cyber Operations and Architecture (Cyber Risk and Resilience Team)

Policy/Procedure Owner	Chief Information Office (CIO)	Chief Information Office (CIO)
Relevant committees	**Data, Information Management and Cybersecurity Committee** – chaired by the Chief Counsel and including the DVCR, CIO, CPO, VP Finance and Marketing and the Associate Directors of each of the ITS portfolios	N/A Individual consultation with specific groups within the University were conducted throughout development. Examples include; ● Data Protection and Privacy Office (DPPO) ● eSoutions (enterprise IT) ● University Library ● Pro-vice Chancellor (Research Infrastructure) portfolio including Director, Research Platform Data Strategy and platform leads
Review and Approval	**Strategic ICT Working Group** (main ITS governance committee) and then endorsed by the **University Executive Team** including the Vice Chancellor	Procedural updates require approval from the procedure owner, in this instance, the CIO.
Scope	● All UTAS data (excluding Defence Contracts)	● All Monash data (excluding Defence Contracts)

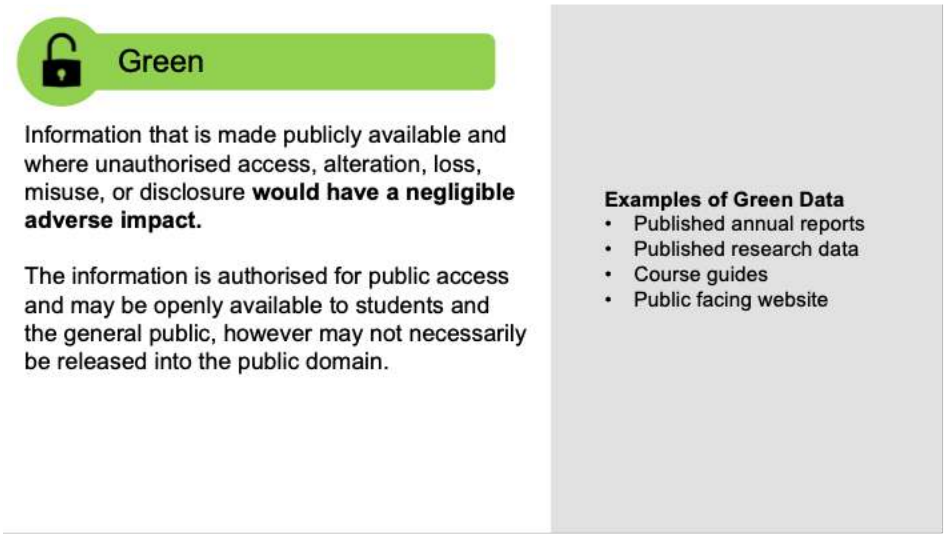
The specifics of the Framework that UTAS and Monash developed and adopted are shown (below), followed by some commentary on the development process and integrations with the Sensitive Data Element recommendations. Relevant points taken from the SD Element are shown within borders.

1.4 Final UTAS Data Classification Framework



UTAS Data Classification Framework - Categories

Green	Yellow	Orange	Red
Data misuse would have little to no impact	Data misuse *unlikely* to cause harm or have a *negligible impact*	Data breach or misuse *may adversely impact* the university or an external party Release could be a regulatory offence	Data breach or misuse expected to cause *serious harm* to the university or an external party Release could be a regulatory or criminal offence
Generally open data	Data has a restricted audience Access must be authorised		



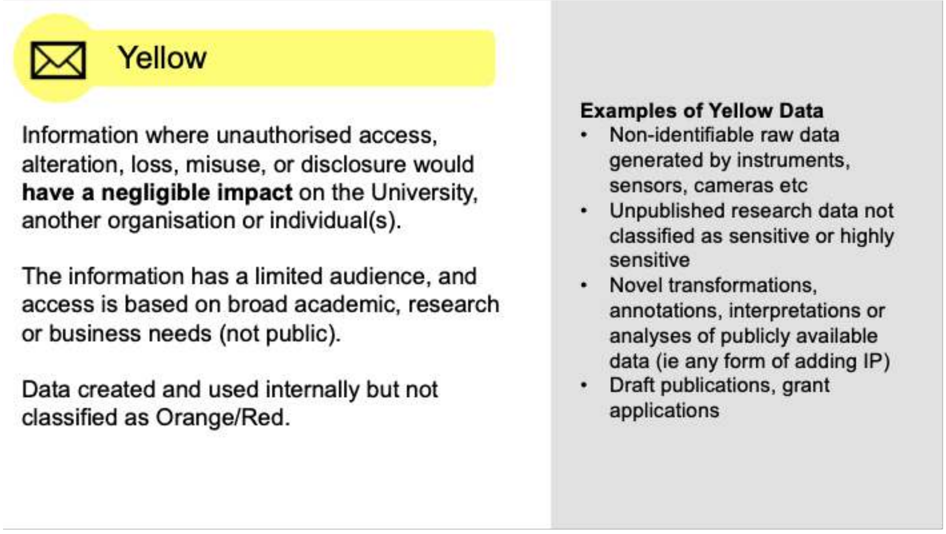
Green

Information that is made publicly available and where unauthorised access, alteration, loss, misuse, or disclosure **would have a negligible adverse impact.**

The information is authorised for public access and may be openly available to students and the general public, however may not necessarily be released into the public domain.

Examples of Green Data

- Published annual reports
- Published research data
- Course guides
- Public facing website



Yellow

Information where unauthorised access, alteration, loss, misuse, or disclosure would **have a negligible impact** on the University, another organisation or individual(s).

The information has a limited audience, and access is based on broad academic, research or business needs (not public).

Data created and used internally but not classified as Orange/Red.

Examples of Yellow Data

- Non-identifiable raw data generated by instruments, sensors, cameras etc
- Unpublished research data not classified as sensitive or highly sensitive
- Novel transformations, annotations, interpretations or analyses of publicly available data (ie any form of adding IP)
- Draft publications, grant applications

Orange

Information where unauthorised access, alteration, loss, misuse, or disclosure **may adversely impact** the University, another organisation or individual(s).

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

Compromise **may** constitute a breach of legal or regulatory responsibilities.

Disclosure to third parties is for specific purposes and requires authorisation.

Examples of Orange Data

- Personally identifiable data as defined by the Privacy Act
- Re-identifiable sensitive data
- Culturally, ecologically or commercially sensitive data
- Data with 'commercial in confidence' or other contractual restrictions
- Financial data – budget forecasting
- HR data (home contact details, bank records)
- Student records
- Sensitive infrastructure/physical records – locations of chemical stores etc

Red

Information where unauthorised access, alteration, loss, misuse, or disclosure could reasonably be **expected to seriously and adversely impact** the University, another organisation or individual(s).

The information has a restricted audience, and access may be subject to regulatory obligations.

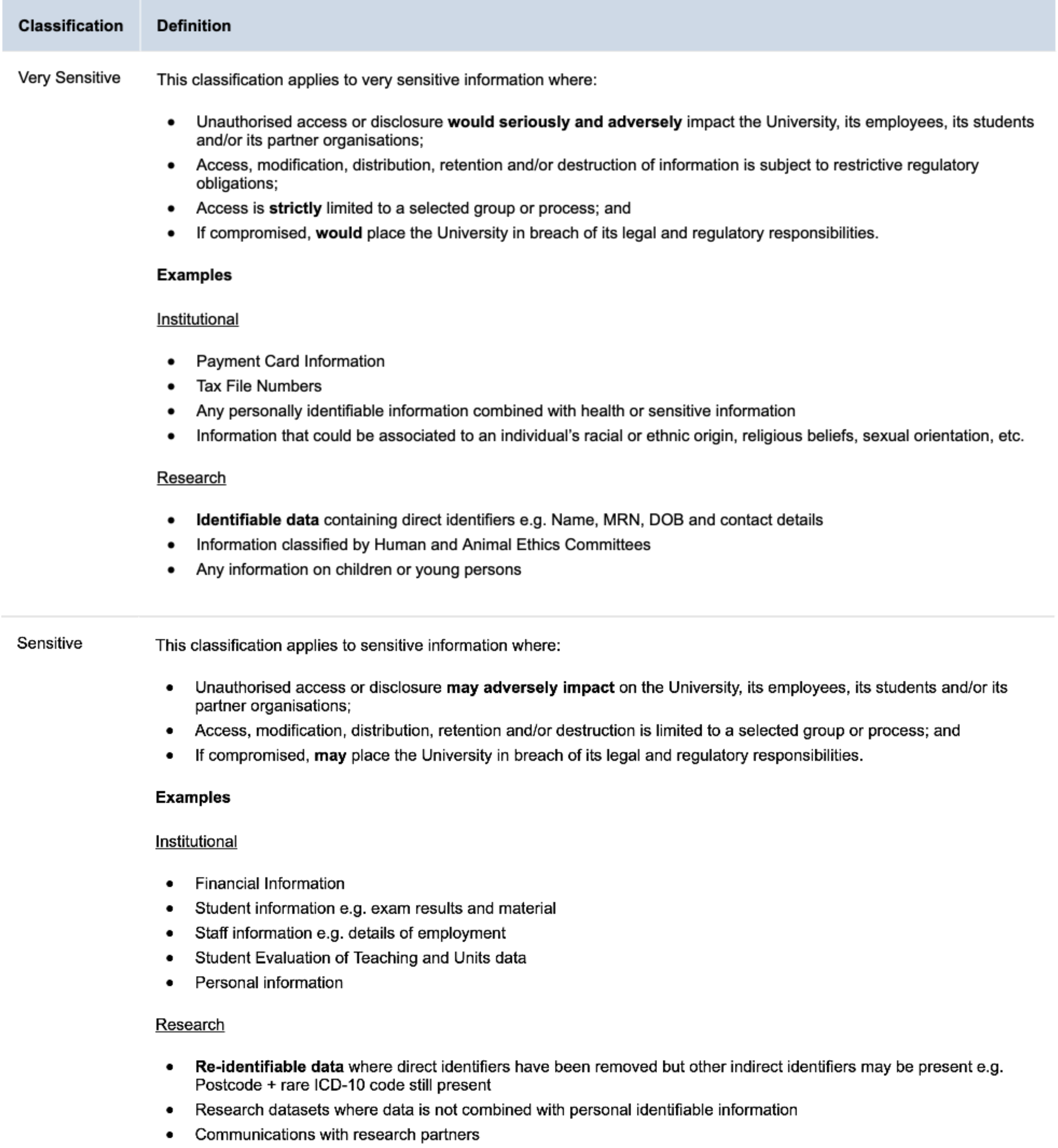
Compromise would constitute a breach of legal or regulatory responsibilities.

Disclosure to third parties is for specific purposes and requires authorisation.

Examples of Red Data

- Information classified as *highly sensitive* or requiring additional protections by a governance committee or commercial agreement
- Data subject to regulatory control
- Data with extreme commercial or strategic sensitivity
- Operational records that are strategic or highly confidential

1.5 Final Monash Data Classification Framework



Classification	Definition
Very Sensitive	This classification applies to very sensitive information where: • Unauthorised access or disclosure **would seriously and adversely** impact the University, its employees, its students and/or its partner organisations; • Access, modification, distribution, retention and/or destruction of information is subject to restrictive regulatory obligations; • Access is **strictly** limited to a selected group or process; and • If compromised, **would** place the University in breach of its legal and regulatory responsibilities. **Examples** Institutional • Payment Card Information • Tax File Numbers • Any personally identifiable information combined with health or sensitive information • Information that could be associated to an individual's racial or ethnic origin, religious beliefs, sexual orientation, etc. Research • **Identifiable data** containing direct identifiers e.g. Name, MRN, DOB and contact details • Information classified by Human and Animal Ethics Committees • Any information on children or young persons
Sensitive	This classification applies to sensitive information where: • Unauthorised access or disclosure **may adversely impact** on the University, its employees, its students and/or its partner organisations; • Access, modification, distribution, retention and/or destruction is limited to a selected group or process; and • If compromised, **may** place the University in breach of its legal and regulatory responsibilities. **Examples** Institutional • Financial Information • Student information e.g. exam results and material • Staff information e.g. details of employment • Student Evaluation of Teaching and Units data • Personal information Research • **Re-identifiable data** where direct identifiers have been removed but other indirect identifiers may be present e.g. Postcode + rare ICD-10 code still present • Research datasets where data is not combined with personal identifiable information • Communications with research partners

Restricted

This classification applies to restricted information where:

- Unauthorised access, modification, distribution, retention and/or destruction or disclosure **may have a negligible** impact on the University, its employees, its students and/or its partner organisations;
- Does **not** include very sensitive or sensitive information, but is created or received within the University (including by students) and used internally;
- Disclosure **would not** cause damage to the University, its employees, its students and/or its partner organisations;

Examples

Institutional

- Course materials and content
- Educational resources
- Training material
- Building plans and associated information
- Internal processes and procedures

Research

- **De-identified data** that is aggregate data with no identifying information included e.g. Counts of patient admissions to ICU ward per month
- Drafts of research publications
- Data from instruments and imaging systems (excluding those linked to an MRN or patient ID)
- Data from sensors, cameras, recorders etc. that do not contain identifiers (e.g. faces)

Public

This classification applies to publicly available information where:

- It Is made available, or released to the general public; and
- No adverse effects are expected to result from the wide circulation of this information.

Examples

Institutional

- The Monash University home page (www.monash.edu.edu) and web presence
- Faculty course lists and the University Handbook
- Monash research achievements and broadcast events
- Information in the public domain
- General institutional and business information

Research

- Monash research achievements and broadcast events
- Publicly released annual reports (e.g. clinical study reports)
- Published research data/information in Bridges or discipline repositories

1.6 Regulatory landscape considered

University of Tasmania (particularly for Red Data)	Monash University
• Australian Code for Responsible Conduct in Research, 2018 • National Statement on Ethical Conduct in Research • Foreign Influence Transparency Scheme • Foreign Arrangements Scheme • Autonomous Sanctions Act 2011 • Security Legislation Amendment (Critical Infrastructure Protection) Bill, 2020 • Privacy Act, 1988 (Cth) • Personal Information Protection Act 2004 (Tas) • Defence Trade Controls Act, 2012 • Environment Protection and Biodiversity Conservation Act, 1999 • National Health Security Act, 2007 and National Health Security Regulations, 2018 • Security Sensitive Biological Agents standards • AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research, 2020 • European Union General Data Protection Regulation 2016/679 (GDPR) • Royal commissions	• Information Privacy Act 2000 (Vic)- note Information Privacy Principles within the Act (Section 14 and Schedule 1) • Privacy Act 1988 (Commonwealth) • Privacy Amendment (Enhancing Privacy Protection) Act 2012 • (Commonwealth) • Health Records Act 2001 (Vic)- note Health Privacy Principles within Act (Section 19 and Schedule 1) • Higher Education Support Act 2003 (Commonwealth)- note Part 5-4 Management of Information, and specifically section 179-10 Use of Personal Information • Education Services for Overseas Students Act 2000 (Commonwealth) – specifically The National Code 2007, Standard 3.1(d) • Epidemiological Studies (Confidentiality) Act 1981 (Commonwealth) - where relevant to a research project (needed) • Public records Act 1973 (VIC) • Monash University (Council) Regulations Part 7 Monash University (Vice-Chancellor) Regulations Part 5 Monash University Statute

2. The UTAS and Monash Frameworks - their reflection of the ADM and SD working Groups

2.1 One Framework to rule them all

> We also recommend that this [data classification] scheme be specific to research data (although it may be beneficial to align it with an existing enterprise information management classification scheme). Research data is less uniform and has a different risk profile and user requirements than enterprise data. A classification scheme that was developed specifically for enterprise data is unlikely to be fit for purpose when applied to research data.

Initially UTAS was planning two Classification Frameworks, one for research and one for all other university data, however it quickly became apparent that this was going to cause confusion and lead to a disconnect in procedures. Instead, it was agreed one framework would be developed that should be flexible enough to allow a research or enterprise lens to be applied. This was a similar experience in Monash, where only a single procedure exists for the purposes of both enterprise and research data. While there had been consideration into creating a separate classification scheme, it was ultimately decided to ensure alignment and streamline the process, the update of the CIO Electronic

Information Security - Information Classification Procedure would also include research. The process for refreshing the Monash procedure was hence initiated by the enterprise IT side of the University with consultation on how the language could be flexible enough to represent application to research scenarios.

In both universities, the decision to incorporate both enterprise and research data into a single classification scheme was largely related to resourcing and effort levels required for individual schemes to be developed. Including enterprise data in the Framework improved the ability to leverage funding for this process, including external consultants on the legal aspects, cybersecurity platform assessments and even for updates to research infrastructure underpinning data management.

This is somewhat conflicting with the recommendation in the SD Element; however it does align with the underlying message. We did not retrofit a Framework designed for enterprise data but co-designed a framework for both*.

- As we were using a risk/consequence lens in the development we removed any dimensionality to the data that would make research different to enterprise – noting this only works if the risk matrix being used to determine consequence has research relevant elements. In both information classification schemes, examples for both a research and enterprise context were included as part of the classification descriptions.
- Including enterprise data meant this process was initiated and supported by Audit and Risk.
- Including enterprise data in the framework greatly increased the priority of this project at a university leadership level and allowed it to progress through the approval processes faster than if it had been a research-only consideration.
- The joint Framework provided a common understanding for implementation; including allowing ITS and Cyber Security to streamline the assessment of platforms storing both research and enterprise data.

2.2 Crosswalk analysis

The working group encourages institutions to continue the development of the draft sensitive data classification level crosswalk

and

The second call to action in the ADRM element states that ‘institutions are encouraged to share with one another examples of the standards in use for their data classification’

Part of the reasoning for the establishment of this project between UTAS and Monash for the second phase of the Institutional Underpinnings was to contribute to the crosswalk and sharing of learnings for an information classification framework. The intention was to provide insights into different approaches on how to approach to develop a framework that may be of use to a University beginning their journey in this area.

The UTAS Framework development started with the Protective Security Policy Framework (PSPF) used by the government. It was then reduced to four categories spanning the equivalent of *Unofficial* to *Protected* in the PSPF, since anything with a classification higher than *Protected* would likely be bound by the PSPF. We then referred to the work of Owen Griffiths from Griffiths University who provided Institutional Underpinnings with an excellent spreadsheet comparing Data Classification schemes across Australian universities. Although this crosswalk primarily referred to Research data classifications it provided some valuable insights such as:

- Most universities use a four-tiered classification

- The default for unpublished research data was not always consistently in the second or third tier
- Separation between the top two tiers (the most sensitive or high-risk data) were not clearly defined and mostly referred to broad categories of risk
- The various data classification frameworks, in their raw form, would be hard for researchers to interpret and apply
- The naming conventions used between the top two tiers sometimes make it difficult to distinguish which is the highest risk (ie confidential/protected or sensitive/protected)

In the case of Monash, as explained below in section 2.4, development of the classification framework was built upon an existing classification scheme. This meant that updating this classification largely related to ensuring that the language and examples were relevant given current understandings.

2.3 Risk/Consequence

Recommendation 3: Sensitive Data Classification levels should be defined by the severity of risk/consequences of mishandling or exposure of the data
The SD Element recommends the adoption of a “risk and consequence” data classification scheme – where “the degree of sensitivity reflects the severity of the risk posed by mishandling or exposure of the data”

Both UTAS’ and Monash’s frameworks define their classifications in terms of the risk to the University. In the case of UTAS, development using a risk/consequence approach (in alignment with the SD element) involved classifying data primarily by risk rather than other potential dimensions such as size, access requirements etc. These components would be considered in the practical implementation of the Framework (i.e. what systems might be used to hold the data) but would not impact the actual classification of the data. In Monash’s development, feedback from research stakeholders in the process resulted in changes to language about how the risk should be communicated to ensure it better reflected the research environment. For example, feedback from the research representatives (University Library and PVC-Research Infrastructure portfolios) resulted in the inclusion of “unauthorised **access** or disclosure” instead of just “unauthorised disclosure”. The feedback highlights that while one classification is more efficient and effective for development, careful consideration into how the language of the information classification scheme translates across enterprise and research contexts is important. Ensuring relevant examples for both enterprise and research contexts were included as part of each classification level was another commonality between each of the Universities' frameworks.

2.4 Classification Scheme language

Use an ordinal structure, from the lowest to highest risk/consequences/protections

and,

Keep it simple, with relatively few levels

and,

Consider how all levels of sensitivity (from open to extreme risk) will be approached

and,

Avoid labelling or defining the levels using language that matches or clashes exactly with other classifications to reduce potential confusion

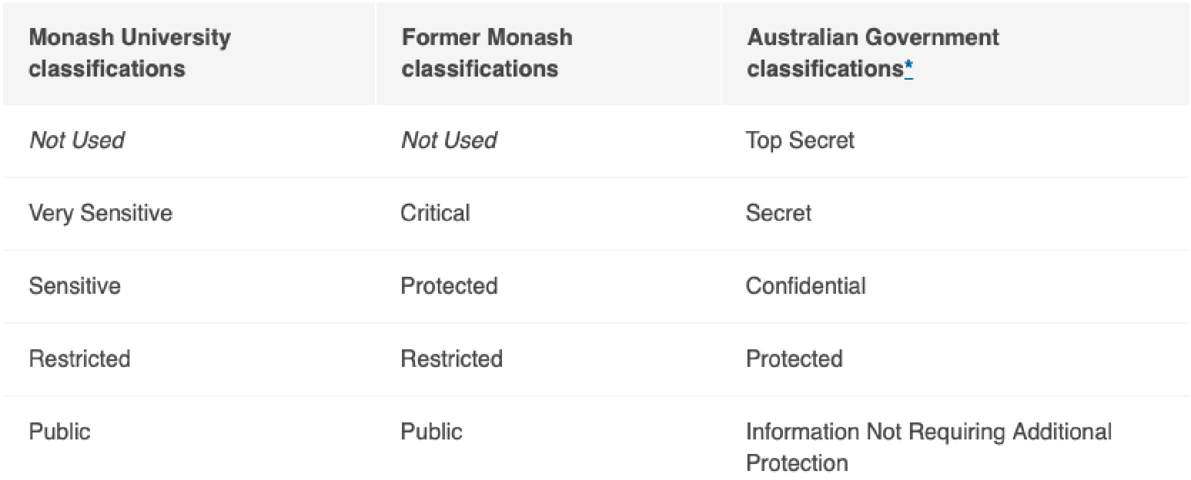
In UTAS' information classification framework, it was decided that colours over written labels would be used as classification categories. Green to Red is well recognised within the risk/consequence landscape (although might have unintentional consequences in the diversity space for colourblind individuals which may warrant investigation). Green, yellow, orange, red also aligns directly with the UTAS Risk Matrix, Risk Appetite statements and other risk/consequence related evaluation tools without requiring any additional mapping or integration work.

UTAS opted for a four-tiered classification in line with *unofficial* to *protected* from the PSPF. Anything higher than the top tier would be bound by the PSPF or equivalent external classification.

As above, the classification ranged from Green (open data will have no risk consequence if misused) through to Red with potentially severe to catastrophic risk to an individual or institution if breached or misused. Considerations for handling sensitivity higher than Red would be stipulated by the PSPF or equivalent external classification relating to those data.

Based on the crosswalk analysis, the UTAS Data, Information Management and Cybersecurity Committee determined that any text base labels for the levels were open to confusion – particularly in distinguishing between the two highest risk tiers (ie Sensitive vs Protected vs confidential - and which one would be the most extreme). Instead UTAS opted to follow Melbourne University's lead and label the levels by colour – as this was the easiest to interpret and immediately matched with the University Risk Matrix and Risk Appetite scoring system.

While the final framework at UTAS was similar in many ways to Monash's framework, including the 4 tired classification scheme, the colours of the categories and the description of the categories, the major difference was the labeling of the categories. Monash decided on using worded labels for their classification categories in conjunction with colors instead of pure colours to indicate classification levels. This decision was largely based on a need to ensure that communication and understanding of the classifications was as straightforward as possible. Monash's previous information classification scheme (pictured below) was determined to be ambiguous and easily misinterpreted in its meaning.



Monash University classifications	**Former Monash classifications**	**Australian Government classifications***
Not Used	*Not Used*	Top Secret
Very Sensitive	Critical	Secret
Sensitive	Protected	Confidential
Restricted	Restricted	Protected
Public	Public	Information Not Requiring Additional Protection

For example, particularly with researchers, the term "Critical" or "Protected" data was associated with importance of their data rather than its sensitivity or classification level. Hence, the decision during the updating of the classification scheme was to revise these to be "Sensitive" and "Very Sensitive" to ensure clear communication on the purpose of these classifications. This enables a more informed decision and discussion about appropriate services. With this in mind, Monash did not decide to use colours as representations of classification on their own because of the potential for miscommunication when applying the classification scheme. While this is the case, it was also important to consider how the language could potentially be interpreted.

It should be noted that at the time of Monash's classification scheme revision, the Defence Industry Security Programme information classifications were not implemented and hence these classification levels are not considered in the Electronic Information Security - Information Classification Procedure. A separate defence related classification has been developed and implemented at Monash to reflect this difference in sensitivity classifications. A consideration for other Universities developing or looking to improve their classification schemes would be to consider classes of data that cannot fit into a single or centralised information classification scheme. These classes may require separate consideration and evaluations to meet different requirements.

3. Next steps and Final Insights

The ADRM Element notes that:

> Designing active data management systems requires an understanding of the types of data that must be managed and the requirements for those data types. **We recommend the creation or adoption of a clear research data classification scheme** to help with this task. Platforms can then be assessed against these requirements, and appropriate procedures developed to ensure that platforms are used appropriately.

Following the update to the Electronic Information Security - Information Classification Procedure at Monash there have been a number of initiatives that align with the ADM working group's above recommendation.

Appropriate services for data classification schemes
The Cyber Security team (part of the CIO) has developed an internal website that maps mostly enterprise managed services to their appropriate data classification. Assessment of the appropriate information security classification for a service is conducted through an Information Security Risk Assessment (ISRA) process, established and conducted by the Cyber Security team. An ISRA takes into account largely technical controls of a service to determine which classification of data it is suitable for holding. This process is conducted for both new and established enterprise and research managed services.

A research-focussed approved services list for corresponding research services and data is awaiting approval and is currently led by the Monash University Library. The ADM working group specifically states that *"it is the institution's responsibility to provide (or endorse) the infrastructure for good active data management…"*. While work is still being undertaken, it is currently not clear who will own and endorse a research-focussed approved services list at Monash. Additionally, it has been identified through the feedback process of Monash's framework development that while the ISRA reflects technical controls, there is not an equivalent evaluation method for governance and process controls also required in the research context.

Specifically responding to Recommendation 4 of the ADM *"Clearly outline the appropriate use cases and terms for each platform or service, and disseminate information about endorsed platforms or services through multiple channels"*. Following the implementation of ISRAs and the acknowledgment of the need to communicate to researcher's the capability and appropriateness of services, Monash has developed Capability Statements for new services that have been developed in response to the Information Classification Scheme eg. SeRP. Capability statements are intended to be a single source of information on the purpose and intended use for a service following its assessment. This provides a mechanism for clear communication on the expectations for a service to stakeholders.