GOVERNMENT OF WESTERN AUSTRALIA

Department of **the Premier and Cabinet**
Office of Digital Government

WA Government Artificial Intelligence Assurance Framework

Artificial Intelligence (AI) is an interdisciplinary field, dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning.

In WA, the scope of AI is wide and includes automated decision making and data driven tools.

This Framework is intended to be applied for development, training and use of AI or data driven tools. Apply the Framework **before** you use or deploy your AI system or data driven tools.

December 2023
Version 1.0

Contents

Throughout the Framework, the term Artificial Intelligence also refers to automated decision making, and data driven tools.

1. About the AI Assurance Framework

About the AI Assurance Framework

What is it?

The AI Assurance Framework (the Framework) will help you design, build and use AI or data driven tools technology appropriately. The Framework contains questions that you will need to answer at every stage of your project and while you are operating an AI system. If you cannot answer the questions, the Framework will let you know how to get help.

The aim of the Framework is to support the WA Government to innovate with AI technology, while making sure we use it safely and secure, with clear accountability for the design and use of our AI Systems.

Who should use it?

The Framework is intended to be used by:

- project teams who are using AI systems and data driven tools or digital in their solutions
- operational teams who are managing AI systems
- Senior Officers who are accountable for the design and use of AI or data driven tools and systems
- internal assessors conducting agency self-assessments

When should I use it?

All AI systems and projects based on data driven tools must be assessed against the Assurance Framework. You must use the Framework during all stages of an AI project from inception to handover.

Is applying the Framework everything I need to do?

The Framework is not a complete list of all requirements for AI projects. Project teams should comply with their agency-specific AI processes, policy requirements and governance mechanisms as well.

Before you start

You must read this before you start any AI project:

- WA Government Artificial Intelligence Policy

When you do not need to apply the Framework

You do not need to assess your product or service if:

- you are using AI systems and data driven tools or digital tools that are widely available commercial applications, and
- you are not explicitly training, prompting or customising the AI system, and
- you are not using the tool in any way the potentially creates Elevated Risk (see Elevated risk page)

Uses of Generative AI systems, Large Language Models and generic AI platforms must apply the Assurance Framework.

Exempt Examples: personal digital assistant, smart phones, smart watches, laptops, QR code reader, satnav system, smart card reader, smoke detector, digital thermometer.

Commitment to Human Rights

AI must not be used to make unilateral decisions that impact our citizens or their human rights

Questions to ask of any AI project

- Is the AI system likely to restrict human rights? If so, is any such restriction publicly justifiable?
- Have possible trade-offs between the different principles and rights been ascertained, documented and evaluated?
- Does the AI system suggest actions or decisions to make, or outline choices to human users?
- Could the AI system inadvertently impact human users' autonomy by influencing and obstructing their decision-making?
- Did you evaluate whether the AI system should inform users that its outputs, content, recommendations, or results arise from an algorithmic decision?

There are laws in WA that protect the human rights of all people. Examples include:
Disability Discrimination Act 1992 (Cth)
Equal Opportunity Act 1984 (WA)
International Covenant on Civil and Political Rights 1976 (OHCHR, UN)

Publicly available resources:
Australian Human Rights Commission
https://humanrights.gov.au/
Public Sector Guidance Sheets
https://www.ag.gov.au/rights-and-protections/human-rights-and-anti-discrimination/human-rights-scrutiny/public-sector-guidance-sheets

Do I need a Human Rights Impact Assessment (HRIA)?

An initial high level risk assessment should be made on all AI projects to indicate whether a more detailed HRIA would be required.

The parameters for an initial assessment should include the:

- Understanding the goals of the AI project
- Potential harms to people arising from use of the AI system
- Scale of any impact or potential harms
- Degree of transparency of the project or system
- Ethical risk severity (for example: financial, physical, mental)
- Quality of data to be used in the project

How to conduct an AI assurance assessment

Assess risk factors

Consider and determine the risk factors for your AI or data driven project using the risk metrics in the Framework

Answer questions & document reasons

Consider and capture your responses to the questions in the Framework
Make a decision about whether your project should:

- continue as-is
- continue with additional treatments
- Stop

Consider that any information you capture may be subject to Freedom of Information Act or public disclosure.

Self-assess and, if required, submit to the WA AI Advisory Board

Record your self-assessment.

See next page for when to submit to the WA AI Advisory Board

Responsible Officers:

- Executive sponsor for the project
- Project lead
- Officer responsible for the technical performance of the AI or data driven system:
- Officer response for data governance:

If additional space is required for responses, extra pages are provided at the end of the self assessment stage.

Responsible officers to complete the Framework:

This assessment is to be completed by (or the result confirmed with) the Responsible Officers. The roles cover the different elements of authorising Framework, project leadership and those responsible for technical performance and data governance.

These four roles have independent responsibilities and must not all be held by the same person.

The Responsible Officers should be appropriately senior, skilled and qualified for the role.

When to submit your project to the WA AI Advisory Board

Completing the assessment

Agencies will be required to submit their completed self-assessments to The Office of Digital Government for review by the WA AI Advisory Board (Advisory Board) if a project is utilising AI and is funded under the Digital Capability Fund, or exceeds a total cost threshold of $5 million. Additionally, AI projects that identify residual risks (after mitigations) during assessment against the Framework which are mid-range or higher, will also require review by the Advisory Board.

In all cases, the project assessment is to be completed by (or the result confirmed with) the Responsible Officers.

To submit your assessment to the WA AI Advisory Board, email ai-dgov@dpc.wa.gov.au .

Recommendations from the WA AI Advisory Board

The WA AI Advisory Board may provide feedback and recommendations to improve the AI or data driven project.

The Responsible Officers remain responsible for the impact and outcomes of the project.

2. Scope of application of the AI Assurance Framework

Completing the assessment – Classify the project

Classify the project

Determine whether the project is Elevated Risk use of AI or not.

Reflect on your risk factors

Consider how you answered your risk assessments.

Consider the residual risk after you have put treatments in place. If they are still high or very high, you will need more treatments

Risk summary

Complete your risk summary and determine how your project should proceed

Mitigations and next steps

Based on the risks identified and the data used, different “Control” levels are required.

Do I need to use the Framework? Most likely YES…

Yellow project "buy AI and use"

I am buying or using a product (or service) off the shelf. I will interact with it to generate results, but not change the data or algorithm. For example, using a generative AI tool in its native form (e.g. ChatGPT, Bard, Llama), or as built into a commercial browser tool (e.g. Chrome, Edge, Firefox).

Green Project "embed AI and / or co-train"

I am buying an AI component and will build my own product (or service) offering around it, or, I am buying a generic AI platform and will augment the training of this product with my own data. For example, I am building AI based biometrics into my system, or I am building a chatbot based on a Large Language model and I will add to its training with my own data.

Blue Project "develop AI and / or train"

I am building (co-developing) a custom tool. Even if based on a standard platform, I am developing algorithms and / or supplying the training data. For example, I am developing an analytics tool to generate insights, or I am training a large language model only with my own data.

Orange Project "rules-based automation"

I am building a rules based engine which does not learn or adapt but does automate decisions. For example, I am automating a decision tree, or I am automating an administrative process.

Do I need to use the Framework?

If your project looks like those described on the left side of this page, the answer is YES.

The scope of use of data driven, intelligent algorithms varies widely. Even if you are not changing the tool through custom training on your own data, or not modifying the underlaying algorithms, you still need to apply the Assurance Framework.

The range of considerations to assure appropriate use of AI or data driven tools or data driven tools will vary considerably for different project types.

The Transparency principle is particularly important for yellow and green projects.

If the Answer is NO, you can still apply the Framework. Let the WA AI Advisory Board know if you have a new use of AI or data driven tools not covered here!
ai-dgov@dpc.wa.gov.au

Is my use of AI potentially an “Elevated Risk” use?

Operational Impact

Does your project / system make administrative decisions? For example, does it make recommendations, outline choices, suggest actions, or prompt human users?

No

\+ **Operational Impact**

Does your project / system automate decisions or take actions that affect your organisation, members of the public or physical systems?

No

\+ **Privacy and Security**

Will the project / system involve the collection, use or disclosure of personal information? For example: Was personal information used as an input to train an algorithm? Will personal information be processed in regular operation of the system?

No

\+ **Possible Harms**

Is there any way your technology could fail, be misused or inappropriately deployed so as to cause harm to an individual, group or society at large?

No

\+ **Explainability and transparency**

Is the decision-making process of the algorithm opaque or not open to scrutiny by the user (e.g. Generative AI)? Is the data contributing to the algorithm output not practical to audit (e.g.. large language models)?

No

If you answer “yes” to any of these questions, then your use is potentially “Elevated Risk” and extra precautions should be considered.

Can I still use AI or data driven tools for an application which is “Elevated Risk”?

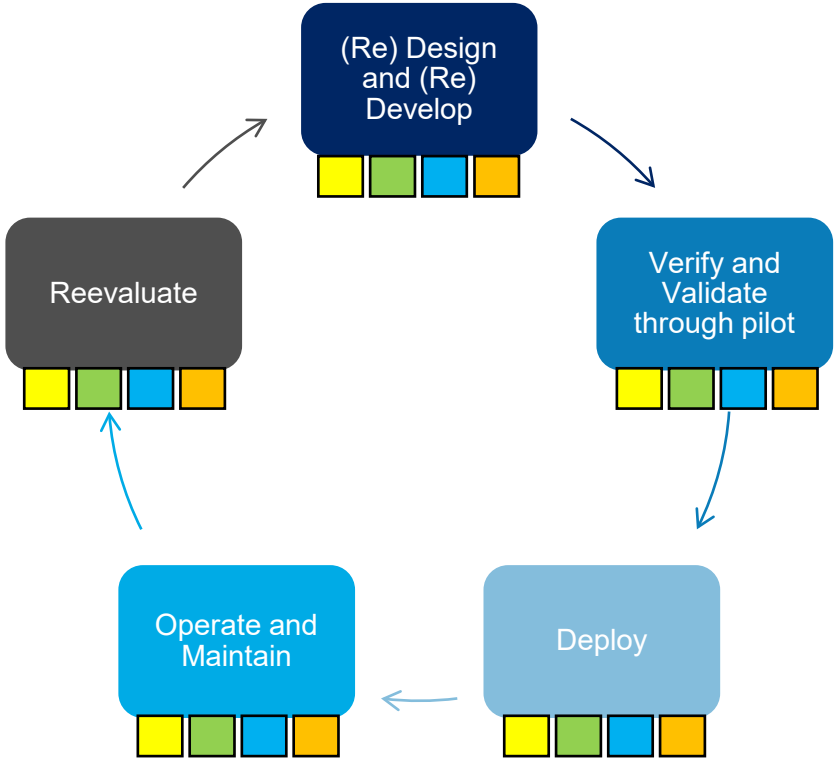
Look at the project uses descriptions on the left side of this page. Answering “yes” to any of these means your project is potentially Elevated Risk.

The range of considerations to assure appropriate use of AI or data driven tools or data driven tools will vary considerably for different project types.

Extra care should be applied to ensure increasingly independent evaluation and monitoring for harms at different stages of the project lifecycle.

Large language models and generative AI require special care associated with output validation and appropriate human decision making associated with use of algorithm outputs.

At what stage(s) should I apply the Framework?



Application of the Assurance Framework over the project lifecycle.

The simplified lifecycle shown gives multiple points where the Assurance Framework should be applied.

The design and develop phases will be different for Yellow, Green, Blue and Orange projects depending on the “buy and use”, “buy and train”, “build and train”, or just “build and automate” approaches.

All uses of AI and data driven tools should be piloted to verify correct operation and then evaluated before being deployed at scale. Greater focus should be given for uses with potentially Elevated Risk.

Similarly, all projects should have ongoing monitoring and periodic re-evaluation which may lead to redesign

3. Engaging with benefits and risks

Completing the assessment – Reflect on risks

Classify the project

Determine whether the project is Elevated Risk use of AI or not.

Reflect on your risk factors

Consider how you answered your risk assessments.

Consider the residual risk after you have put treatments in place. If they are still high or very high, you will need more treatments

Risk summary

Complete your Risk summary and determine how your project should proceed

Mitigations and next steps

Based on the risks identified and the data used, different “Control” levels are required.

Evaluating AI benefits and risks

Benefits and risks

WA Government has a strong commitment to the responsible use of technology.

This means you need to evaluate the potential risks of harms from deployment and operation of AI, as well as its benefits.

Currently, we use AI or data driven tools tools to:

- deliver insights that improve services and lives
- help agencies work more quickly and accurately

While there are many areas where AI can benefit the work we do, we need to engage with risks early and throughout the life lifecycle of the technology.

Evaluating and engaging with risk

This AI Assurance Framework is structured in sections that align to five identified AI Ethics Principles.

Each section starts with a page that prompts you to consider the types of risk that your project may carry and helps shape your response to questions in that section with risk in mind.

At the end of the self-assessment, you will assign a risk rating (highest risk and total number of risks ranked medium or higher) to the different Ethics Principles in your AI project. This rating will determine if your project should:

- proceed as is
- proceed, with additional risk mitigations
- stop.

Cannot answer some questions?

It is important to make a note of questions you cannot answer as you progress through the assessment. It may be because information is not available or can only be answered once a pilot is undertaken.

If the project proceeds, treat these unanswered questions as representing Midrange risk, commence with a pilot phase and closely monitor for harms and establish controls.

Understanding the balance of benefits and risks

Some projects carry real risk (for example within Health), but are undertaken to improve existing processes, or because of a clear benefit to community.

Identifying and managing of these risks during the life of the project is an essential requirement, as is clarifying the benefits of the project.

“Elevated Risk” uses of AI and data driven tools

Elevated risk uses of AI and data driven tools

Elevated risk uses of AI and data driven tools include those that have a real-world effect. The purpose is to generate a recommendation, an alert, alarm, decision or action and so either prompting a human to act, or the system acting by itself. Elevated risk uses of AI and data driven tools often work in real time (or near real time) using a live environment for their source data.

AI systems and data driven tools that have been trained on personal information as inputs have potentially Elevated Risk associated with bias, inclusion (or not) of data on minorities as well as added requirements for management of the data (and products created from the data) from a privacy and security perspective.

AI tools which have been trained by third parties on very large data sets and / or for which the algorithmic process cannot be effectively explained require additional mitigations around the ability to validate the accuracy of outputs and ensure appropriate and empowered human decision making.

Start assessment questions with

Community benefit

AI should deliver the best outcome for the citizen, and key insights into decision-making.

Extra care should be applied to ensure increasingly independent evaluation and monitoring for harms at different stages of the project lifecycle.

Other uses of AI and data driven tools

Other uses of AI and data driven tools which do not have an operational impact not use a live environment for their source data. Most frequently, they produce analysis and insight from historical data.

Irrespective of these general characteristics, the risk level needs to be carefully and consciously determined, especially where there is a possibility that AI insights and outputs may be used to influence important future policy positions.

Start assessment questions with

Fairness

Use of AI or data driven tools will include safeguards to manage data bias or data quality risks, following best practice and Australian or International Standards

Benefits identification

For all AI systems, the benefits of the AI project should be captured in a Benefits Realisation Management Plan before commencement.

AI risk factors exist on a spectrum

The key factor that determines risk is how the AI system is used, including whether it has an operational impact, or if it is not transparent or not explainable (page 11).

Very low risk or N/A

AI generates insights for **non-operational** human use from non-sensitive data

(example: analytics package reporting on historical non-sensitive data)

Low

AI generates insights or alerts for **operational** human use with minimal potential for harm

(example: anomaly detection software; alarm system)

--

AI makes **operational** actions, decisions or recommendations with no routine human oversight with minimal potential for harm

(example: automated door; biometric login with alternative login methods; automated phone menu, smart sign showing driver speed)

Midrange

AI generates **operational** insights / decisions / recommendations for human to action with some potential for harm

(example: public facing chatbot; red light camera, intruder alert system)

--

AI generates insights for **non-operational** human use analysing sensitive data

(example: analytics package operating on data of vulnerable individuals)

High

AI makes and implements **operational** decisions autonomously of human input in the interests of human safety and wellbeing

(example: anti-lock braking system)

--

AI makes and implements **operational** decisions within a specified range, and refers exceptions for human to review and action

(example: loan application system, autonomous tram)

Very high risk

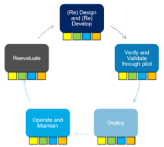
AI makes and implements **operational** decisions that can negatively affect human wellbeing autonomously of human input

(example: autonomous benefits eligibility reviews; judicial custodial sentence recommendations, unconstrained autonomous system, self driving car)

4. Begin self-assessment stage

Basic project information

Project Name:	
How is the project being delivered?	• Yellow project “buy AI and use” • Green Project “embed AI and / or co-train” • Blue Project “develop AI and / or train” • Orange Project “rules-based automation” * Other (please explain)
What is the current phase of the project?	• Design and Develop • Verify and Validate through pilot • Deploy • Operate and Maintain • Reevaluate
Who is Executive sponsor for the project?	
Who is the Project Lead in your organisation?	



The Assurance Framework is a living document.

The responses to the Framework questions should be updated at each stage of the project / use lifecycle.

Basic project information

Who is responsible for data governance for this project within your organisation?	
Who is responsible for the technical performance of the AI or data driven system?	
Which other team members are directly involved in the project?	
Who has contributed to this current assessment and on what date?	
What is the next date/milestone that will trigger the next review of the project?	

Ethics Principles

The below 5 ethics principles must be applied when conducting an AI assurance assessment.

Community benefit

AI should deliver the best outcome for the citizen, and key insights into decision-making.

Start here for Elevated Risk projects. Continue to end of Framework.

Fairness

Use of AI or data driven tools will include safeguards to manage data bias or data quality risks, following best practice and Australian or International

Start here for all other projects. Continue to end of Framework.

Privacy and security

AI will include the highest levels of assurance. Ensure projects adhere to the WA interim privacy position.

Transparency

Review mechanisms will ensure citizens can question and challenge AI-based outcomes. Ensure projects adhere to the Freedom of Information Act 1992

Accountability

Decision-making remains the responsibility of organisations and Responsible Officers. Responsible Officers will adhere to the State Records Act 2000 and maintain appropriate records to provide accountability for their actions and decisions.

More information

More information can be found in the WA Government Artificial Intelligence Policy. You must consider and apply this Policy when designing, implementing or running an AI System.

General benefits assessment

Consider the benefits associated with the AI project …	Very low or N/A	Low	Midrange	High	Very high
Delivering a better quality *existing* service or outcome (e.g.. accuracy or client satisfaction)	○	○	○	○	○
Reducing processing or delivery times	○	○	○	○	○
Generating financial efficiencies or savings	○	○	○	○	○
Providing an AI capability that could be used or adapted by other agencies	○	○	○	○	○
Delivering a *new* service or outcome (particularly if it cannot be done without using AI)	○	○	○	○	○
Enabling *future* innovations to existing services, or new services or outcomes	○	○	○	○	○

Benefits realisation management is essential for AI projects

Think about the *potential* benefits of your AI project and the likelihood of these benefits being *realised* in practice; as well the strength of available *evidence* supporting your assessment. Indicate the overall level of *confidence* in your assessment (e.g.. low, midrange, high, very high) and any major variation in the level of confidence between different types of benefit.

Comments:

General risk factor assessment

Consider the risks associated with …	Very low risk or N/A	Low	Midrange	High	Very high risk
Whether this AI system is delivering a new or existing service	○	○	○	○	○
The potential to cause discrimination from unintended bias	○	○	○	○	○
Whether the AI system is a single point of failure for your service or policy	○	○	○	○	○
If there is sufficient *experienced* human oversight of the AI system	○	○	○	○	○
Over-reliance on the AI system or ignoring the system due to high rates of false alert	○	○	○	○	○
Whether the linkage between operating the AI system and the policy outcome is clear	○	○	○	○	○

Is a new service or policy automatically high risk?

There are always risks associated with a new service or policy simply because it has not been implemented before. To address the risks of a new service, think ahead about the potential harms, their likelihood and how readily they can be reversed. Also think about the role of human oversight of the new service. It is important to document your responses to identified risks and provide evidence of controls enacted to mitigate risks.

Comments:

Community benefit

1. Will the AI system improve on existing approaches to deliver the outcomes aligned to:

- Government priorities, objectives and strategies
- Agency Key Efficiency Indicators and Outcomes
- Your Agency strategic plans and/or
- another relevant WA Outcomes Framework?

Response:

- ○ yes _____ document your reasons, then go to next question
- ○ partially ____ after your pilot, you must conduct a formal benefits review before scaling the project. Document your reasons and go to the next question
- ○ not sure ____ pause the project and prepare a Benefits Realisation Management Plan
- ○ no _____ do not proceed any further. Discuss this project with the policy or service owner

Benefits

All AI projects should have a benefits register that is kept up to date throughout the project.

The benefits register should be handed over to the service owner at the end of the project.

Community benefit

2. Were other, non-AI systems and data driven tools considered?

Response:

- yes _____ document your reasons, then go to next question
- informally___ after your pilot, you must conduct a formal benefits review before scaling the project. Document your reasons and go to the next question
- no _____ do not proceed any further. Discuss this project with the policy or service owner

Alternatives

For an AI project to be viable, AI must be the most appropriate system for your service delivery or policy problem.

AI systems and data driven toolscan come with more risk and cost than traditional tools. You should use an AI system when it the best system to maximise the benefit for the customer and for government.

Alignment with legal frameworks

3. Does this project and the use of data align with relevant legislation?

You must make sure your data use aligns with:

- Equal Opportunity Act 1984
- Freedom of Information Act 1992
- State Records Act 2000
- WA Interim Privacy Position

Other relevant WA or Commonwealth Acts including:

- WA Public Health Act 2016
- Public Interest Disclosure Act 2003
- Relevant Acts for your Agency

Response:

- yes _____ document your reasons, then go to next question
- unclear ____ pause the project. Seek advice from the State Solicitor's Office or an appropriate WA legal professional. You may need to redesign your project
- no _____ do not proceed any further unless you receive clear legal advice that allows the project to proceed. Consider redesigning your project.

More information

You must comply with laws for records, privacy, information access at all times, including when you are developing and using AI Systems.

AI Projects: Risk factors for individuals or communities

Consider the risks of…	None, negligible, N/A	Reversible with negligible consequences	Reversible with moderate consequences	Reversible with significant consequences	Significant or irreversible
Physical harms	○	○	○	○	○
Psychological harms	○	○	○	○	○
Environmental harms or harms to the broader community	○	○	○	○	○
Unauthorised use of health or other sensitive information	○	○	○	○	○
Impact on right, privilege or entitlement	○	○	○	○	○
Unintended identification or misidentification of an individual	○	○	○	○	○
Misapplication of a fine or penalty	○	○	○	○	○
Other financial or commercial impact	○	○	○	○	○
Incorrect advice or guidance	○	○	○	○	○
Inconvenience or delay	○	○	○	○	○
Other harms	○	○	○	○	○
	Very low risk or N/A	**Low**	**Midrange**	**High**	**Very high risk**

Comments: *these responses should be considered as residual risks after mitigations are in place. Use additional pages if space is insufficient.*

Possible harms

4. Considering planned mitigations, could the AI system cause significant or irreversible harms?

If there is a residual risk of significant or irreversible harms and the project proceeds, you must pilot the project first, then conduct a formal benefits review before scaling the project.

For more information on when a Human Rights Impact Assessment is required see https://humanrights.gov.au/

Response:

- no — document your reasons, then go to next question
- yes, but it's better than existing systems — you must seek approval from an ethics committee. You must have clear legal advice that allows this project to proceed. Consult with all relevant stakeholders. Consider a Human Rights Impact Assessment.
- yes — do not proceed any further unless you receive clear legal advice that allows the project to proceed. If you have legal approval: discuss the project with all relevant stakeholders, seek approval from an ethics committee, consider a Human Rights Impact Assessment.
- unclear — pause the project and prepare a Benefits Realisation Management Plan

Monitoring for possible harms

You must monitor your AI system closely for harms that it may cause. This includes monitoring outputs and testing results to ensure there are no unintended consequences.

You should be able to quantify unintended consequences, secondary harms or benefits, and long-term impacts to the community, even during testing and pilot phases. Testing can still do real harm if the system is making consequential decisions. You must consider and account for this possibility even if human testers are willing volunteers.

Changing the context or environment in which the AI system is used can lead to unintended consequences. Planned changes in how the AI is used should be carefully considered and monitoring undertaken.

Possible harms

5. Considering planned mitigations, could the AI system cause reversible harms?

If there is a residual risk of mid-range (or higher) harms and the project proceeds, you must pilot the project first, then conduct a formal benefits review before scaling the project.

Response:

- no — document your reasons, then go to next question
- yes, but it's better than existing systems — you may need to seek advice from an ethics committee. You should clearly demonstrate that you have consulted with all relevant stakeholders before proceeding to pilot phase. Consider a Human Rights Impact Assessment.
- yes — If the risk of harms identified are mid-range or higher, do not proceed any further unless you receive clear legal advice that allows the project to proceed. If you have legal approval: discuss the project with all relevant stakeholders, you may need ethics approval, consider a Human Rights Impact Assessment.
- yes — If the risk of harms identified are low or very low, document your reasons, then go to next question
- unclear — pause the project and prepare a Benefits Realisation Management Plan

Irreversible harms vs reversible harms

An irreversible harm occurs when it is impossible to change back to a previous condition. For example, if an AI system makes an incorrect decision to deny somebody a pension without an option to have that overturned.

You should consider how outcomes can be overturned in the event there is harm caused or the AI system leads to an incorrect decision.

Possible secondary or cumulative harms

6. Considering planned mitigations, could the AI System result in secondary harms, or result in a cumulative harm from repeated application of the AI System?

If there is a residual risk of mid-range (or higher) harms and the project proceeds, you must pilot the project first, then conduct a formal benefits review before scaling the project.

Response:

- ○ no — document your reasons, then go to next question
- ○ yes, but it's better than existing systems — you may need to seek advice from an ethics committee. You should clearly demonstrate that you have consulted with all relevant stakeholders before proceeding to pilot phase. Consider a Human Rights Impact Assessment.
- ○ yes — If the risk of harms identified are mid-range or higher, do not proceed any further unless you receive clear legal advice that allows the project to proceed. If you have legal approval: discuss the project with all relevant stakeholders, you may need ethics approval, consider a Human Rights Impact Assessment.
- ○ yes — If the risk of harms identified are low or very low, document your reasons, then go to next question
- ○ unclear — pause the project and prepare a Benefits Realisation Management Plan

Secondary harms

Sometimes harms are felt by people who are not direct recipients of the product of service. We refer to these as secondary harms. Secondary harms include things like a loss of trust.

You need to think deeply about everyone who might be impacted, well beyond the obvious end user.

Fairness: risk factors for AI projects

Consider the risks associated with…	Very low risk or N/A	Low	Midrange	High	Very high risk
Using incomplete or inaccurate data	○	○	○	○	○
Having poorly defined descriptions and indicators of “Fairness”	○	○	○	○	○
Not ensuring ongoing monitoring of “Fairness indicators”	○	○	○	○	○
Decisions to exclude outlier data	○	○	○	○	○
Informal or inconsistent data cleansing and repair protocols and processes	○	○	○	○	○
Using informal bias detection methods (best practice includes automated testing)	○	○	○	○	○
The likelihood that re-running scenarios could produce different results (reproducibility)	○	○	○	○	○
Inadvertently creating new associations when linking data and/or metadata	○	○	○	○	○
Differences in the data used for training compared to the data for intended use	○	○	○	○	○

Comments: *these responses should be considered as residual risks after mitigations are in place. Use additional page if space is insufficient.*

Fairness

7. Can you explain why you selected this data for your project and not others?

Response:

- yes — document your reasons, then go to next question
- unclear — consult with relevant stakeholders to identify alternative data sources or implement a data improvement strategy or redesign the project
- it's better than existing systems — document your reasons. You should clearly demonstrate that you have consulted with all relevant stakeholders before proceeding to pilot phase.
- no — pause the project and consider how absent data or poor quality data will impact your system.

Data relevance and permission

Your AI system may draw in multiple datasets from different sources to find new patterns and insights.

You need to determine you can and should use the data for the AI system. This can be challenging for historical data that may have been collected for a different purpose.

Fairness

8. Is the data that you need for this project available and of appropriate quality given the potential harms identified?

If your AI project is a data creation or data cleansing application, answer according to the availability of any existing data that is needed for the project to succeed, for example, training datasets.

Response:

- ○ yes — document your reasons, then go to next question
- ○ unclear — consult with relevant stakeholders to identify alternative data sources or implement a data improvement strategy or redesign the project
- ○ it's better than existing systems — document your reasons. You should clearly demonstrate that you have consulted with all relevant stakeholders before proceeding to pilot phase.
- ○ no — pause the project and consider how absent data or poor quality data will impact your system.

Data quality

Data quality is often described in terms of minimum requirements for accuracy, timeliness, completeness, and consistency.

Your AI system may be significantly impacted by poor quality data. It is important to understand how significant the impact is before relying on insights or decisions generated by the AI system.

Absence of data may lead to unintended biases impacting insights generated by the AI system. Unbalanced data is a common problem when training AI systems.

Fairness

9. Does your data reflect the population that will be impacted by your project or service?

- yes — document your reasons, then go to next question
- it's better than existing systems — you may need to seek advice from an ethics committee. You should clearly demonstrate that you have consulted with all relevant stakeholders before proceeding to pilot phase. Consider a Human Rights Impact Assessment
- no or unclear — pause the project and address the gaps in your solution design
- N/A — document your reasons as to why this does not apply, then go to next question

Response:

Fairness

10. Have you considered how your AI system will address issues of diversity and inclusion (including geographic diversity)?

11. Have you considered the impact with regard to gender and on minority groups including how the solution might impact different individuals in minority groups when developing this AI system?

Minority groups may include:

- those with a disability
- LBGQIT+ and gender fluid communities
- people from CALD backgrounds
- Aboriginal and Torres Strait Islanders
- children and young people

- yes _____ document your reasons, then go to next question
- it's better than existing systems ____ you may need to seek advice from an ethics committee. You should clearly demonstrate that you have consulted with all relevant stakeholders before proceeding to pilot phase. Consider a Human Rights Impact Assessment
- no or unclear ____ pause the project and address the gaps in your solution design
- N/A _____ document your reasons as to why this does not apply, then go to next question

Diversity and inclusion, and the impact on minorities

Services or decisions can impact different members of the relevant community in different ways.

Whether due to cultural sensitivities, or underrepresentation in training data sets. It is important to think deeply about everyone who might be impacted by AI Systems.

Response:

Fairness

12. Do you have appropriate performance measures and targets (including fairness ones) for your AI system, given the potential harms?

Aspects of accuracy and precision are readily quantifiable for most systems which predict or classify outcomes. This performance can be absolute, or relative to existing systems.

How would you characterise "Fairness" such as equity, respect, justice, in outcomes from an AI system? Which of these relate to, or are impacted by the use of AI?

Response:

- yes _____ document your reasons, then go to next question
- no or unclear ____ for Elevated Risk AI systems, pause the project until you have established performance measures and targets. for non-operational systems, results should be treated as indicative and not relied on.
- N/A _____ document your reasons as to why this does not apply, then go to next question

Measuring AI system performance

At the scoping stage, you will need to make important choices about what you measure. You should measure:

- Accuracy: how close an answer is to the correct value
- Precision: how specific or detailed an answer is
- Sensitivity: the measure of how many actually positive results are correctly identified as such
- Specificity: the measure of how many actually negative results are correctly identified by the AI system
- Fairness objectives: whether the system is meeting the fairness objectives defined for the system (which could include for example that there aren't more prediction errors on some cohorts than others)

Fairness

13. Do you have a way to monitor and calibrate the performance (including fairness) of your AI system?

Elevated Risk AI systems and data driven tools which are continuously updated / trained can quickly move outside of performance thresholds. Supervisory systems can monitor system performance and alert when calibration is needed.

Response:

- yes _____ document your reasons, then go to next question
- no or unclear ____ for Elevated Risk AI systems, pause the project until you have established performance measures and targets. for non-operational systems, results should be treated as indicative and not relied on.
- N/A _____ document your reasons as to why this does not apply, then go to next question

Measuring AI system performance

Elevated Risk AI systems and data driven tools should have clear performance monitoring and calibration schedules.

For Elevated Risk AI systems and data driven tools which are continuously training and adapting with moderate residual risks, weekly performance monitoring and calibration is recommended. For low risk, monthly evaluation and calibration is recommended.

For operational systems with high risk or very high risk, a custom evaluation and calibration will be required.

Sensitive data considerations for AI projects

Sensitive data including information on:	Identifiable cohort >50 or N/A	Identifiable cohort >20 and <50	Identifiable cohort >10 and <20	Identifiable cohort >5 and <10	Identifiable cohort <5
Children	○	○	○	○	○
Religious individuals	○	○	○	○	○
Racially or ethnically diverse individuals	○	○	○	○	○
Individuals with political opinions or associations	○	○	○	○	○
Individuals with trade union memberships or associations	○	○	○	○	○
Gender and/or sexually diverse individuals	○	○	○	○	○
Individuals with a criminal record	○	○	○	○	○
Specific health or genetic information	○	○	○	○	○
Personal biometric information	○	○	○	○	○
Other sensitive person-centred data	○	○	○	○	○
	Very low risk or N/A	**Low**	**Midrange**	**High**	**Very high risk**

Comments: *these responses should be considered as residual risks after mitigations are in place. Use a additional page if space is insufficient.*

Privacy and security

14. Have you applied the “Privacy by Design” and “Security by Design” principles in your project?

Response:

- yes — document your reasons, then go to next question
- partially — pause the project, consult with your stakeholders and determine how you will improve your data or practices
- no or unclear — pause the project until you have received appropriate advice including from the Office of Digital Government. You may need to re-design your project..

Privacy by Design and Security by Design

Even small AI projects may have privacy or security vulnerabilities. For example, an analytics project which stores commercially sensitive data in a non-secure environment unbeknown to the user.

The NSW Information Privacy Commissioner has prepared 7 Privacy by Design principles. These principles should be applied to your AI project. If you are unsure how to apply these principles, you seek help from the Office of Digital Government.

The WA Government Cyber Security Policy should also be applied to all digital projects. Be sure to always refer to the NSW Information and Privacy Commission (NSW IPC).

Privacy and security

15. Have you completed a privacy impact assessment (either third party or self-assessed)?

Response:

- yes _____ document your reasons, then go to next question
- no _____ pause the project until you have completed a privacy impact assessment.

Privacy Impact Assessment (PIA)

Even projects not focussed on person-centred data may reveal information about a person, their relationships or preferences. For example analysis of environmental or spatial data may reveal information about a land-holder's interaction with the local environment.

A PIA can help you to identify and minimise privacy risks; implement 'Privacy by Design' and demonstrate compliance with privacy laws.

The Office of the Australian Information Commissioner (OAIC) has published a helpful guide and e-learning course.

Privacy and security

16. If you are using information about individuals who are reasonably identifiable, have you sought consent from citizens about using their data for this particular purpose?

If you are using information about individuals who are reasonably identifiable, have you sought their consent to use their personal information for this particular purpose?

See the Freedom of Information Act 1992 for a definition of Personal Information. See also the OAIC Guidelines (B94-B97) on the meaning of "reasonably identifiable".

Response:

- yes — document your reasons, then go to next question
- Authorised use — for AI systems and data driven tools intended to operate under legislation which allows use of Identifiable Information, do not proceed unless you receive clear legal / independent privacy advice that allows this project to proceed. The project should be carefully monitored for harms during the pilot phase.
- partially — pause the project until you have consent, or redesign your project
- no — pause the project until you have either consent or clear legal advice authorising use of this information
- N/A — document your reasons as to why this does not apply, then go to next question

Exceptions

For AI systems and data driven tools intended to operate under legislation which allows use identifiable information, the public benefits must be clear before proceeding to pilot phase.

Governing Use of Personally Identifiable Information

You must apply higher governance standards if you are managing Personally Identifiable Information.

Privacy and security

17. Does your AI System adhere to the mandatory requirements in the WA Cyber Security Policy?

Have you considered end-to-end Security Principles for your project?

Response:

- yes — document your reasons, then go to next question
- no or partially — pause the project until these requirements can be met
- N/A — document your reasons as to why this does not apply, then go to next question

Cyber security

As with any emerging technology, AI can pose new cyber security risks and so it is important to be vigilant.

You must comply with the mandatory requirements in the WA Cyber Security Policy

The WA Government Cyber Security Unit has responsibility for leading a coordinated government response to cyber security failures including malware and ransomware attacks.

Privacy and security

18 Does your data set include sensitive information as defined in section 9 of the Privacy Act 1988 (Cth)?

Response:

- no — document your reasons, then go to next question
- yes — seek explicit approval from the Responsible Senior Officer to proceed with this risk. Consider seeking approval from an ethics committee.
- unclear — pause the project and clarify the nature of the data, address any inadvertent use of sensitive data in your system

Sensitive Information

The OAIC Guidelines (B141-144) provide further information about "sensitive information". The WA Information Classification Policy and supplementary guidance have been developed to help agencies correctly assess the sensitivity or security of information, so that the information can be labelled, used, handled, stored and disposed of correctly.

Governing Use of Sensitive Information

You must apply higher governance standards if you are managing Sensitive Information. Refer to the page addressing Governance Requirements.

Transparency: risk factors for AI projects

Consider the risks associated with…	Very low risk or N/A	Low	Midrange	High	Very high risk
Incomplete documentation of AI or data driven tools system design, or implementation, or operation	○	○	○	○	○
No or limited access to model's internal workings or source code ("Black Box")	○	○	○	○	○
Being unable to explain the output of a complex model	○	○	○	○	○
A member of the public being unaware that they are interacting with an AI system	○	○	○	○	○
No or low ability to incorporate user feedback into an AI system or model	○	○	○	○	○

Is a 'black box' system use, such as a large language model or generative AI, automatically high risk?

The inner workings of commercial AI systems and data driven tools are not always accessible and even if they are, they can be very complex to interpret. To address the risks this poses, think proactively about the role of human judgement in use of an "unexplainable" insight or decision. If you cannot explain the ways in which insights are outputted from an AI system, what are the potential harms that may arise? What's the likelihood of these harms and how readily they can be reversed? It is important that these considerations are documented. This is particularly important if midrange or higher risks are identified.

Comments: *these responses should be considered as residual risks after mitigations are in place. Use additional page if space is insufficient.*

Transparency

19. Have you consulted with the relevant community that will benefit from (or be impacted by) the AI system?

Response:

- yes _____ document your reasons, then go to next question
- Authorised use ____ for AI systems and data driven tools intended to operate under legislation which allows use without community consultation, do not proceed unless you receive clear legal advice that allows this project to proceed. The project should be carefully monitored for harms during the pilot phase.
- it's better than existing systems ____ you may need to seek advice from an ethics committee. Document your reasons. You should clearly demonstrate that you have consulted with all relevant stakeholders before proceeding to pilot phase.
- no _____ pause the project, develop a Community Engagement Plan and consult with the relevant community
- N/A _____ document your reasons as to why this does not apply, then go to next question

Consultation

You must consult with the relevant community when you design your AI system. This is particularly important for Elevated Risk AI systems.

Communities have the right to influence government decision-making where those decisions, and the data on which they are based, will have an impact on them.

For AI systems and data driven tools intended to operate under legislation which allows use without community consultation, the public benefits must be clear before proceeding to pilot phase.

Transparency

20. Are the scope and goals of the project publicly available?

Response:

- yes _____ document your reasons, then go to next question
- no _____ make sure you communicate the scope and goals of the project to relevant stakeholders and the relevant community who are impacted before proceeding beyond pilot
- N/A _____ document your reasons as to why this does not apply, then go to next question

Sharing project goals

The WA Government recognises we have important work to do to encourage public trust in AI, by ensuring Government is transparent and accountable, and that AI delivers positive outcomes to citizens.

Transparency

21. Is there an easy and cost-effective way for people to appeal a decision that has been informed by your AI system?

Response:

- yes _____ document your reasons, then go to next question
- no _____ pause your project, consult with relevant stakeholders and establish an appeals process
- N/A _____ document your reasons as to why this does not apply, then go to next question

Right to appeal

No person should ever lose a right, privilege or entitlement without right of appeal.

A basic requirement of Transparency is for an individual affected by a relevant decision to understand the basis of the decision, and to be able to effectively challenge it on the merits and/or if the decision was unlawful.

When planning your project, you must make sure no person could lose a right, privilege or entitlement without access to a review process or an effective way to challenge an AI generated or informed decision.

Transparency

22. Does the system using the AI allow for transparent explanation of the factors leading to the AI decision or insight?

Response:

- yes _____ document your reasons, then go to next question
- no, but a person makes the final decision ___ consult with relevant stakeholders and establish a process to readily reverse any decision or action made by the AI system. Actively monitor for potential harms during pilot phase..
- no _____ pause your project, consult with relevant stakeholders and establish a process to readily reverse any decision or action made by the AI system
- N/A _____ document your reasons as to why this does not apply, then go to next question

Clear explanations

As far as possible, you must have a way to clearly explain how a decision or outcome has been informed by AI.

If the system is a “black box” due to lack of access to the inner workings, or is too complex to reasonably explain the factors leading to the insight generation, it is essential to consider the role of human judgement in intervening before an AI generated insight is acted on. It is important to formalise and document this human oversight process.

In low (or very low) risk environments, it may be sufficient to identify and document mechanisms to readily reverse any action arising from such an insight (e.g.. a person overriding an automated barrier).

Accountability: risk factors for AI projects

Consider the risks associated with …	Very low risk or N/A	Low	Midrange	High	Very high risk
Insufficient training of AI or data driven tools system operators	○	○	○	○	○
Insufficient awareness of system limitations of Responsible Officers	○	○	○	○	○
No or low documentation of performance targets or “Fairness” principles trade-offs	○	○	○	○	○
No or limited mechanisms to record insight / AI System decision history	○	○	○	○	○
The inability of third parties to accurately audit AI system insights / decisions	○	○	○	○	○

The skill and training of the operators of the AI system are the most important elements

With all automated systems, there is always the risk of over-reliance on result. It is important that the operators of the system, including any person who exercises judgement over the use of insights, or responses to alerts, is appropriately trained on the use of the AI system. Training must include the ability to critically query insights generated, and to understand the limitations of the AI system.

For Elevated Risk AI systems, the users must be confident they can readily reverse any harms resulting from the use of an AI generated insight or decision, or ensure a Responsible Officer is empowered to make a decision on the use of an AI generated insight. For non-Elevated Risk AI systems, the users must be skilled in the interpretation and critiquing of AI or data driven tools generated insights if the insight is to be relied upon.

Comments: *these responses should be considered as residual risks after mitigations are in place. Use additional page if space is insufficient.*

Accountability

23. Have you established who is responsible for:

- use of the AI insights and decisions
- policy/outcomes associated with the AI system
- monitoring the performance of the AI system
- data governance and recordkeeping

Response:

- yes — document your reasons, then go to next question
- no or unclear — pause the project while you identify who is responsible and make sure they are aware and capable of undertaking their responsibilities
- N/A — document your reasons as to why this does not apply, then go to next question

Responsible officers:

This assessment is to be completed by or (the result confirmed with) the Responsible Officers. These include the Officer who is responsible for:

- use of the AI insights / decisions;
- the outcomes from the project;
- the technical performance of the AI system;
- data governance.

These four roles must not be held by the same person. The Responsible Officer should be appropriately senior, skilled and qualified for the role.

Accountability

24. Have you established a clear processes to:

- intervene if a relevant stakeholder finds concerns with insights or decisions?
- ensure you do not get overconfident or over reliant on the AI system?

Response:

- ○ yes _____ document your reasons, then go to next question
- ○ no _____ pause your project, consult with relevant stakeholders and establish appropriate processes
- ○ N/A _____ document your reasons as to why this does not apply, then go to next question

Human intervention and accountability

For Elevated Risk AI systems, you must make sure that humans are accountable and can intervene. This may also be relevant for non-Elevated Risk AI systems and data driven tools

This will help you to build public confidence and Control in your AI system.

5. Risk summary

Completing the assessment – Risk summary

Classify the project

Determine whether the project is Elevated Risk use of AI or not.

Reflect on your risk factors

Consider how you answered your risk assessments.

Consider the residual risk after you have put treatments in place. If they are still high or very high, you will need more treatments

Risk summary

Complete your Risk summary and determine how your project should proceed

Mitigations and next steps

Based on the risks identified and the data used, different “Control” levels are required.

Risks Identified

Community benefit	Fairness	Privacy and security	Transparency	Accountability
AI should deliver the best outcome for the citizen, and key insights into decision-making.	Use of AI or data driven tools will include safeguards to manage data bias or data quality risks, following best practice and Australian or International	AI will include the highest levels of assurance. Ensure projects adhere to the WA interim privacy position.	Review mechanisms will ensure citizens can question and challenge AI-based outcomes. Ensure projects adhere to Freedom of Information Act 1992	Decision-making remains the responsibility of organisations and Responsible Officers. Responsible Officers will adhere to the State Records Act 2000 and maintain appropriate records to provide accountability for their actions and decisions.

Write in

Highest risk:	Highest risk:	Highest risk:	Highest risk:	Highest risk:

From N/A / Low / Mid-range / High / Very High - these responses should be considered as residual risks after mitigations are place.

Monitoring ongoing risks

Operational AI projects which progress with high and very high risks must plan for regular external risk audits to cover

- the examination and documentation of the effectiveness of risk responses in dealing with identified risk and their root causes,
- the effectiveness of the risk management process

Top risks

Highest risk refers to the most significant risk identified in each of the five principle areas (e.g.., “Community Benefit” or “Fairness).

All Mid-range, High and Very High risks must have effective mitigations.

Projects which progress with medium, high and very high risks must have project-specific legal advice.

Risk summary

Is this actually an Elevated Risk AI system?

After all mitigations are considered if residual risk(s) of Mid-range or above remain, this is an Elevated Risk use of AI.

yes, and the decisions it makes or informs include high or very high risk factors	do not proceed without project-specific legal advice. If the project proceeds, pilot first with ongoing controls and monitoring. A formal review should be conducted after pilot phase. Use of an external review committee is recommended.
yes, and the decisions it makes or informs include medium risk factors	do not proceed without project-specific legal advice. Pilot first with ongoing controls and monitoring required once pilot commences.
yes, and the decisions it makes or informs include low or negligible risk factors	your project can proceed with appropriate ongoing controls and monitoring. Pilot the project first.
no, however its outputs may be used to inform policy and other important decisions	your project can proceed, but you need to review your risk treatments and make sure there are sufficient controls in place
no, it only uses historical data for reporting or informing purposes only	your project can proceed with appropriate ongoing controls and monitoring

Monitoring ongoing performance

For Elevated Risk AI systems, ongoing performance monitoring is essential. Even low risk systems such as an automated barrier, could rapidly change to operate outside of normal parameters. Mechanisms to monitor calibrate system performance should be identified before scaling beyond pilot phase.

The importance of documenting your assessment

You must make sure your answers, explanations and risk mitigating controls are recorded in your document management system.

For Elevated Risk AI systems and data driven tools which include medium risks or higher, the public benefits must be clear and documented before proceeding to pilot phase. Project specific legal advice is required. Projects should be actively monitored for potential harms and remedies identified.

6. Risk mitigations and next steps

Completing the assessment
Mitigations and next steps

Classify the project

Determine whether the project is Elevated Risk use of AI or not.

Reflect on your risk factors

Consider how you answered your risk assessments.

Consider the residual risk after you have put treatments in place. If they are still high or very high, you will need more treatments

Risk summary

Complete your Risk summary and determine how your project should proceed

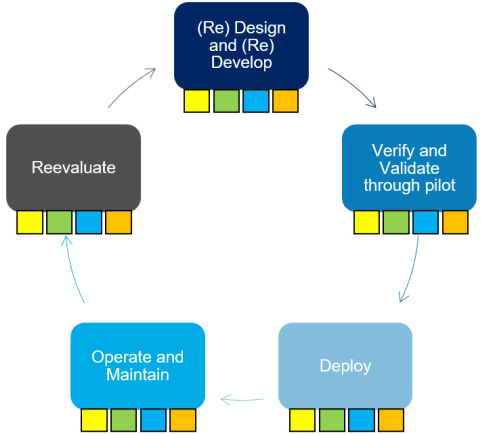
Mitigations and next steps

Based on the risks identified and the data used, different “Control” levels are required.

Mitigations and next steps

25. What are the areas which require special mitigations?

Response:



Monitoring ongoing performance

For Elevated Risk AI systems, ongoing performance monitoring and evaluation is essential.

All automated systems should have ongoing monitoring. Even low risk systems could rapidly change to operate outside of normal parameters.

Mechanisms to monitor and calibrate system performance should be identified before scaling beyond pilot phase.

Procurement considerations

26. If you are procuring all or part of an AI system, have you satisfied the requirements for:

- transparency?
- privacy and security ?
- fairness?
- accountability?

Response:

- yes _____ document your reasons
- no _____ pause your project. Make sure you can meet the requirements before you continue.

How do I work with procurement?

The scope of use of data driven, intelligent algorithms varies widely. Even if you are not changing the tool through custom training on your own data, or not modifying the underlaying algorithms, you still need to consider how the AI is used and ensure that the AI Ethics principles outlined in this document, and the WA Government AI Policy are adhered to.
Work with ICT procurement to ensure that risks identified are appropriately reflected through the procurement / build / train phases.

7. End of self assessment stage

Additional space

Additional space cont.

Glossary

Artificial Intelligence (AI) – is an interdisciplinary field, dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning. In WA, the scope of AI is wide and includes automated decision making and data driven tools.

Bias – in data, this means a systematic distortion in the sampled data that compromises its representativeness, in algorithms it describes systematic and repeatable errors in a computer system that create unfair outcomes, such as privileging one arbitrary group of users over others.

Data Governance – refers to a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods

Data Lifecycle –refers to the entire period of time that data exists in your system. This life cycle encompasses all the stages that your data goes through, from first capture onward.

Data Quality – is a term used to describe a documented agreement on the representation, format, and definition for data.

Data use sensitivity – means risks or considerations associated with data subjects themselves or use of data.

Generative AI – is artificial intelligence capable of generating text, images, or other media, using generative models. Generative AI models learn the patterns and structure of their input training data and then generate new data that has similar characteristics.

Harm – means any adverse effects experienced by an individual (or organisation) including those which are socially, physically, or financially damaging.

Human Rights – are rights inherent to all human beings, regardless of race, sex, nationality, ethnicity, language, religion, or any other status. Human rights include the right to life and liberty, freedom from slavery and torture, freedom of opinion and expression, the right to work and education, and many more. Everyone is entitled to these rights, without discrimination.

Glossary

Large language model (LLM) – a specialized type of artificial intelligence that has been trained on vast amounts of text to understand existing content and generate original content.

Non-operational AI – systems do not use a live environment for their source data. Most frequently, they produce analysis and insight from historical data.

Operational AI – are those that have a real-world effect. The purpose is to generate an action, either prompting a human to act, or the system acting by itself. Elevated Risk AI systems and data driven tools often work in real time (or near real time) using a live environment for their source data.

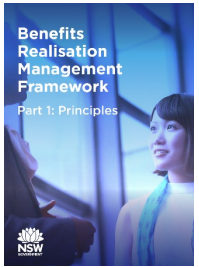
Responsible Officer – These include the Officer who is responsible for: use of the AI insights / decisions; the outcomes from the project; the technical performance of the AI system; data governance.

Serious harm: always context specific, a harm which would result in a serious threat to the life, health, safety or welfare of any individual.

8. Useful resources

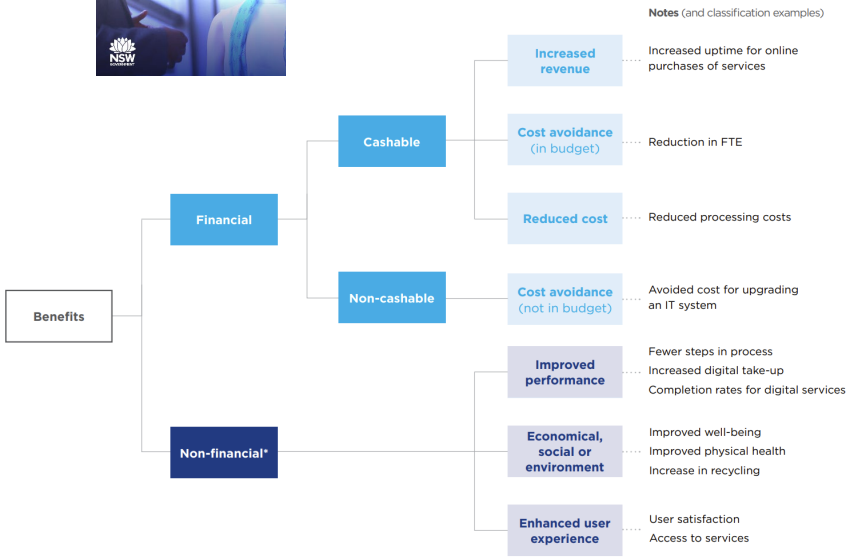
Resource – Benefits Realisation Framework

Community Benefit from the Use of AI or data driven tools Systems



Governance is key to implementing benefits management, as benefits need to be owned by appropriate sponsors and managers from within the organisation. To support active program sponsorship at the senior leadership and executive level:

- develop a program vision statement, to be promoted by senior leadership, to assist with the transformational change required to realise the program benefits.
- review the underlining principles of benefits realisation management
- use benefits management deliverables to clearly articulate the program outcomes and intended benefits
- when possible, manage, report and approve benefit deliverables within existing governance meetings, noting that the size, complexity, priority and risk of a program and its benefits will affect the level of governance required to control its delivery and benefit realisation
- when possible, integrate benefits management processes with other business processes

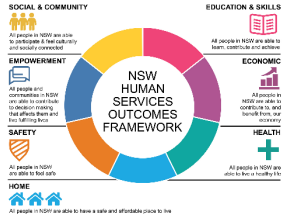


*These are examples and should be tailored to the program/project environment

https://www.nsw.gov.au/customer-service/publications-and-reports/benefits-realisation-management-Framework

Resource - Lean Canvas

Community Benefit from the Use of AI or data driven tools Systems



Community benefit

Overall costs and benefits for the project likely to be established by the business case.

Community benefit in the use of AI or data driven tools to be set out:

- Were alternatives to AI considered and why were they discounted?
- How will the use of AI or data driven tools result in improved customer and service delivery outcomes and efficiencies?

Lean Business Canvas: TITLE OF PROJECT

PROJECT SPONSOR NAME

Hypothesis	Stakeholders	Desired Outcomes	Benefits
Key Questions	**Data Available**	**Current Metrics**	**Value derived from project**
Background/Problem			

Resource – Co-design Example

Community Benefit from the Use of AI or data driven tools Systems

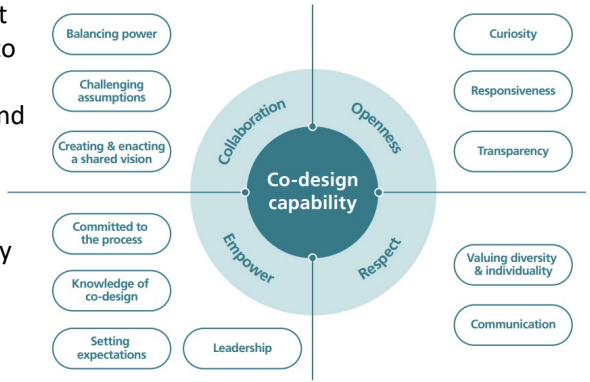
Co-design is a way of bringing major stakeholders together to improve services. It creates an equal and reciprocal relationship between all stakeholders, enabling them to design and deliver services in partnership with each other.

Planning, designing and producing services with people that have experience of the problem or service means the final system is more likely to meet their needs.

This way of working demonstrates a shift from seeking involvement or participation after an agenda has already been set, to seeking consumer leadership from the outset so that consumers are involved in defining the problem and designing the system.

Co-design typically uses a staged process that adopts participatory and narrative methods to understand the experiences of receiving and delivering services, followed by consumers and health professionals co-designing improvements collaboratively.

An example is available from the NSW Agency for Clinical Innovation via the link below.



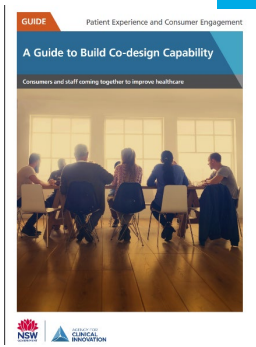
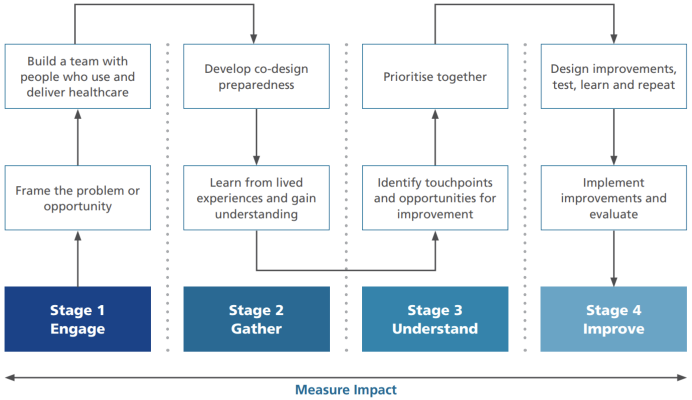


Figure 1. The co-design process



https://aci.health.nsw.gov.au/__data/assets/pdf_file/0013/502240/Guide-Build-Codesign-Capability.pdf

Resource - Recommended Harm Mitigation Approaches

Harm Type	Ethics Expert Review of AI or data driven tools System	Policy Domain Expert Review of AI or data driven tools System	Data Governance / Cyber Security Focus	Analytical Expert Review of AI or data driven tools System	Co-Design of project / actions
Physical	X	X			X
Psychological	X	X			X
Unauthorised Use of Health / Sensitive Information			X	X	X
Unauthorised Use of Personal Information			X	X	X
Impact on Right, Privilege or Entitlement	X	X			X
Misidentification of Individual				X	X
Misapplication of Penalty / Fine		X		X	X
Other Financial Impact		X		X	X
Incorrect guidance / advice				X	X
Inconvenience, Delay				X	X
Other Harms	X	X	X	X	X

9. Some relevant standards

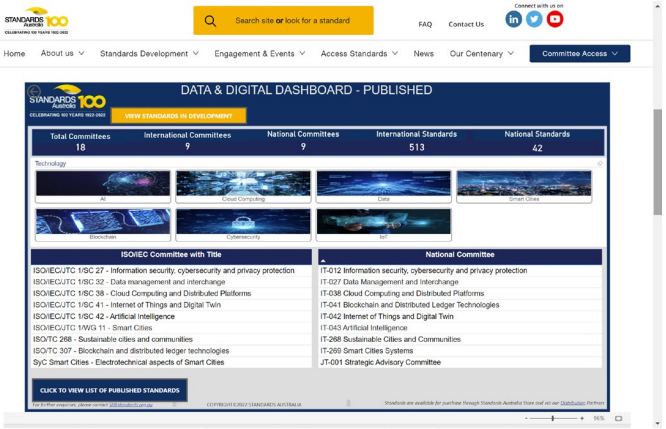
Existing and Developing Standards Families

https://www.standards.org.au/getmedia/f132c974-1ecb-4601-884d-f1e10610fbf3/Data-Digital-Standards-Landscape.pdf.aspx

The most relevant groups within the IEC/ISO/JTC1 family include subcommittees (SC) for data sharing and use include:

SC 27 - Information Security, Cybersecurity and Privacy Protection

- SC 32 - Data Management and Interchange
  - Within SC 32, Working Group 6 (WG6) on Data Usage
- SC 38 - Cloud Computing and Distributed Platforms
- SC 40 - IT Service Management and IT Governance
- SC 41 – Internet of Things and Digital Twin
- SC 42 - Artificial Intelligence



https://www.standards.org.au/engagement-events/strategic-initiatives/critical-and-emerging-technologies/data-digital-dashboard

Recent Standards for Data Quality for AI and ML

https://www.standards.org.au/getmedia/f132c974-1ecb-4601-884d-f1e10610fbf3/Data-Digital-Standards-Landscape.pdf.aspx

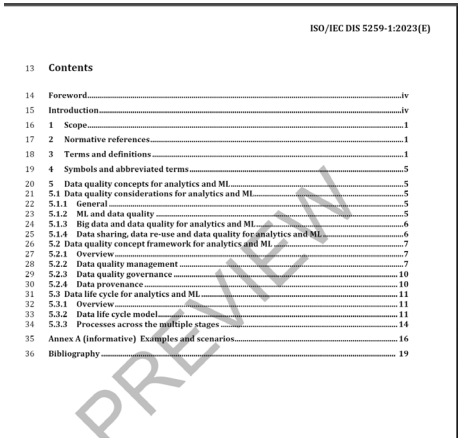
ISO/IEC AWI 5259-1, Data quality for analytics and ML — Part 1: Overview, terminology, and examples

ISO/IEC AWI 5259-2, Data quality for analytics and ML — Part 2: Data quality measures

ISO/IEC AWI 5259-3, Data quality for analytics and ML — Part 3: Data quality management requirements and guidelines

ISO/IEC AWI 5259-4, Data quality for analytics and ML — Part 4: Data quality process framework

ISO/IEC AWI 5259-1, Data quality for analytics and ML Part 1: Overview, terminology, and examples

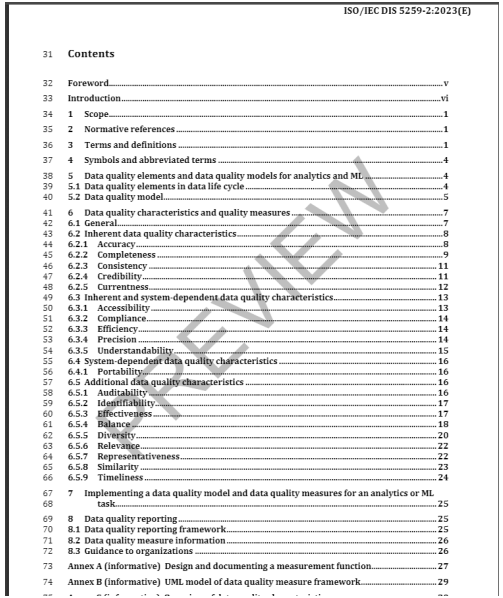


ISO/IEC DIS 5259-1:2023(E)

Contents

Terms and definitions
Symbols and abbreviated terms
Data quality concepts for analytics and ML
Data quality considerations for analytics and ML
ML and data quality
Big data and data quality for analytics and ML
Data sharing, data re-use and data quality for analytics and ML
Data quality concept framework for analytics and ML
Data quality management
Data quality governance
Data provenance
Data life cycle for analytics and ML
Data life cycle model
Processes across the multiple stages

ISO/IEC AWI 5259-2, Data quality for analytics and ML Part 2: Data quality measures



ISO/IEC DIS 5259-2:2023(E)

Contents

Data quality elements and data quality models for analytics and ML.
Data quality elements in data life cycle
Data quality model
Data quality characteristics and quality measures
Inherent data quality characteristics:
Accuracy, Completeness, Consistency, Credibility, Currentness
Inherent and system-dependent data quality characteristics:
Accessibility, Compliance, Efficiency, Precision, Understandability
System-dependent data quality characteristics:
Portability
Additional data quality characteristics:
Auditability, Identifiability, Effectiveness, Balance, Diversity, Relevance, Representativeness, Similarity, Timeliness
Implementing a data quality model and data quality measures for an analytics or ML task
Data quality reporting:
Data quality reporting framework
Data quality measure information

Data Quality – ISO 8000

The ISO 8000 series provides frameworks for improving data quality for specific kinds of data (asset intensive industries).

The series defines which characteristics of data are relevant to data quality, specifies requirements applicable to those characteristics, and provides guidelines for improving data quality.

The series is applicable within all stages of the data life cycle.

The ISO 8000 series can be used either in conjunction with or independently of standards for quality management systems.

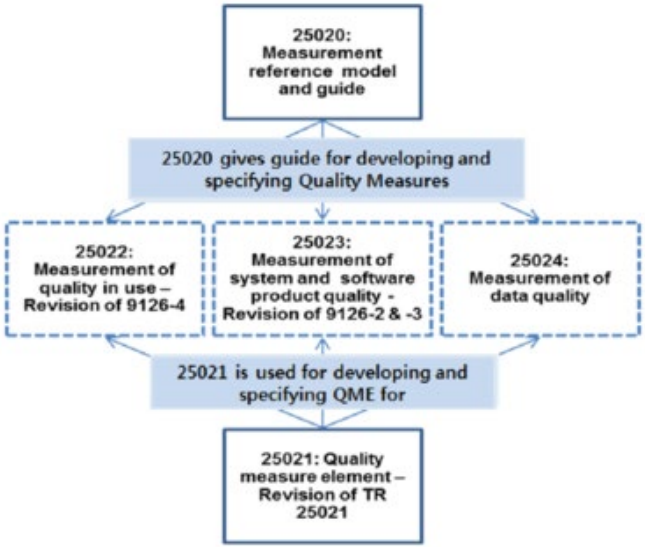
The following are within the scope of the ISO 8000 series:

- general aspects of data quality, including principles, vocabulary and measurement of information and data quality;
- data governance;
- data quality management, including processes, roles, responsibilities and maturity assessment;
- data quality assessment, including profiling and data rules;
- quality of master data, including exchange of characteristic data and identifiers;
- quality of industrial data, including product shape data.

Outside the scope of the ISO 8000 series:

- quality of the things represented by data;

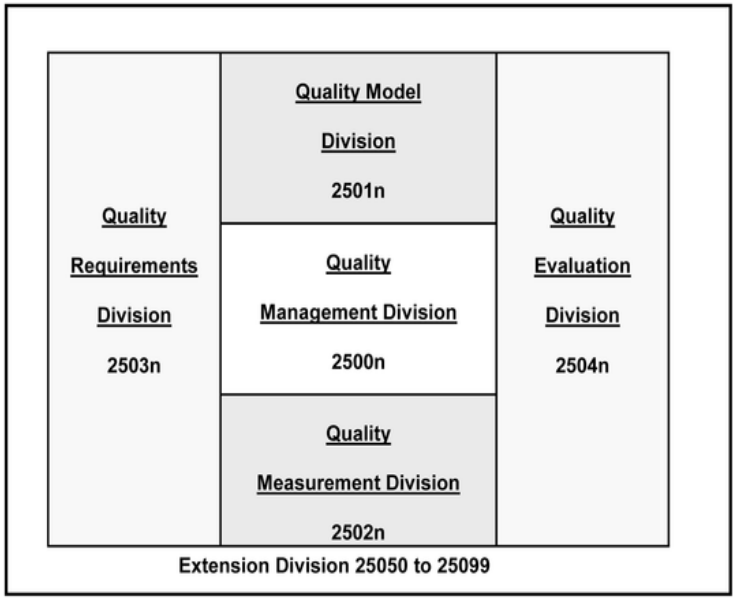
ISO/IEC 25000 Series
- Systems and Software Data Quality



- 25012 on software products data quality model (definitions and framework)
- 25024 on systems and software data quality requirements
- 20252 re data representativeness – primarily the material in the annexes on quality criteria, though referring to market, opinion and social research

ISO/IEC 25012:2008 - Software engineering

Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model



ISO/IEC 25012:2008 defines a general data quality model for data retained in a structured format within a computer system.

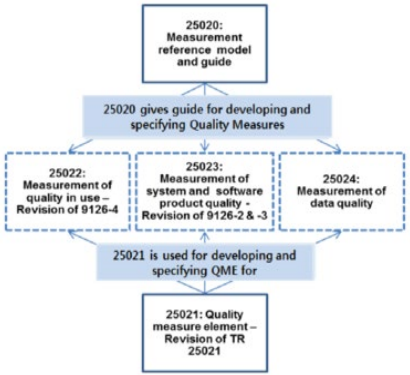
ISO/IEC 25012:2008 can be used to establish data quality requirements, define data quality measures, or plan and perform data quality evaluations:

- to define and evaluate data quality requirements in data production, acquisition and integration processes,
- to identify data quality assurance criteria, also useful for re-engineering, assessment and improvement of data,
- to evaluate the compliance of data with legislation and/or requirements.

ISO/IEC 25012:2008 categorizes quality attributes into fifteen characteristics considered by two points of view: inherent and system dependent. Data quality characteristics will be of varying importance and priority to different stakeholders.

ISO/IEC 25012:2008 is intended to be used in conjunction with the other parts of the SQuaRE series of International Standards, and with ISO/IEC 9126-1 until superseded by ISO/IEC 25010.

ISO/IEC 25000 Series
Systems and Software Data Quality



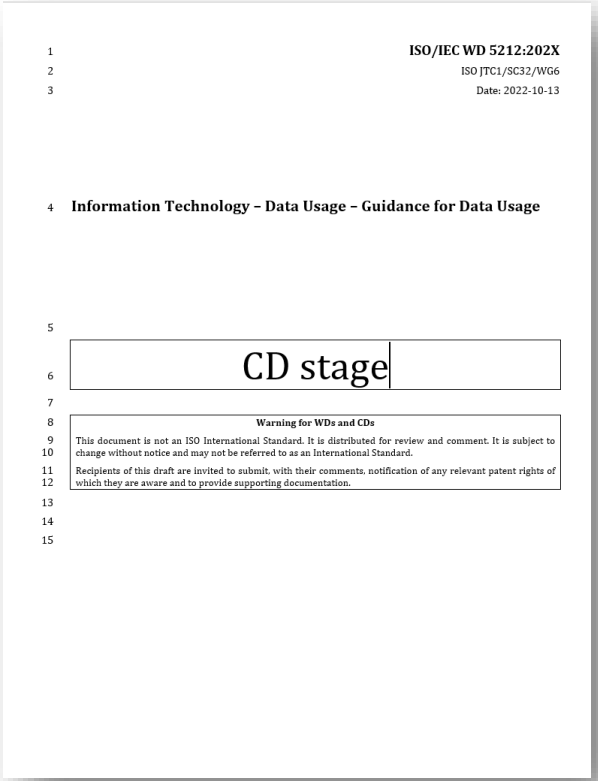
ISO/IEC 25024:2015 defines data quality measures for quantitatively measuring the data quality in terms of characteristics defined in ISO/IEC 25012.

ISO/IEC 25024:2015 contains the following:
- a basic set of data quality measures for each characteristic;
- a basic set of target entities to which the quality measures are applied during the data-life-cycle;
- an explanation of how to apply data quality measures;
- a guidance for organizations defining their own measures for data quality requirements and evaluation.

It includes, as informative annexes, a synoptic table of quality measure elements defined in this International standard (Annex A), a table of quality measures associated to each quality measure element and target entity (Annex B), considerations about specific quality measure elements (Annex C), a list of quality measures in alphabetic order (Annex D), and a table of quality measures grouped by characteristics and target entities (Annex E).

This International Standard does not define ranges of values of these quality measures to rate levels or grades because these values are defined for each system by its nature depending on the system context and users' needs.

Guidance for Data Use – JTC1 SC 32 WG6 (expected 2Q 2024)



ISO/IEC WD 5212:202X
ISO JTC1/SC32/WG6
Date: 2022-10-13

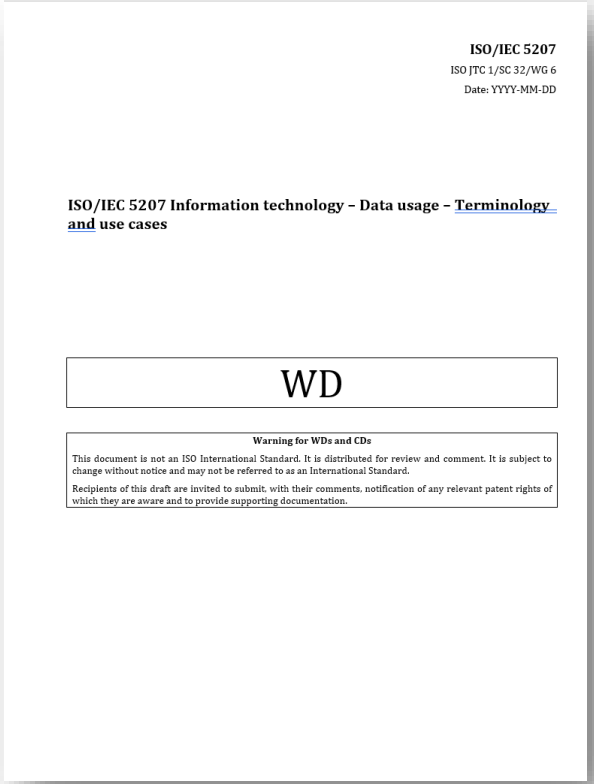
Information Technology – Data Usage – Guidance for Data Usage

CD stage

Warning for WDs and CDs

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.



ISO/IEC 5207
ISO JTC 1/SC 32/WG 6
Date: YYYY-MM-DD

ISO/IEC 5207 Information technology – Data usage – Terminology and use cases

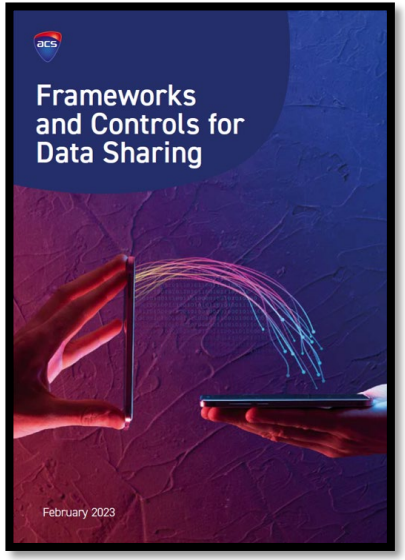
WD

Warning for WDs and CDs

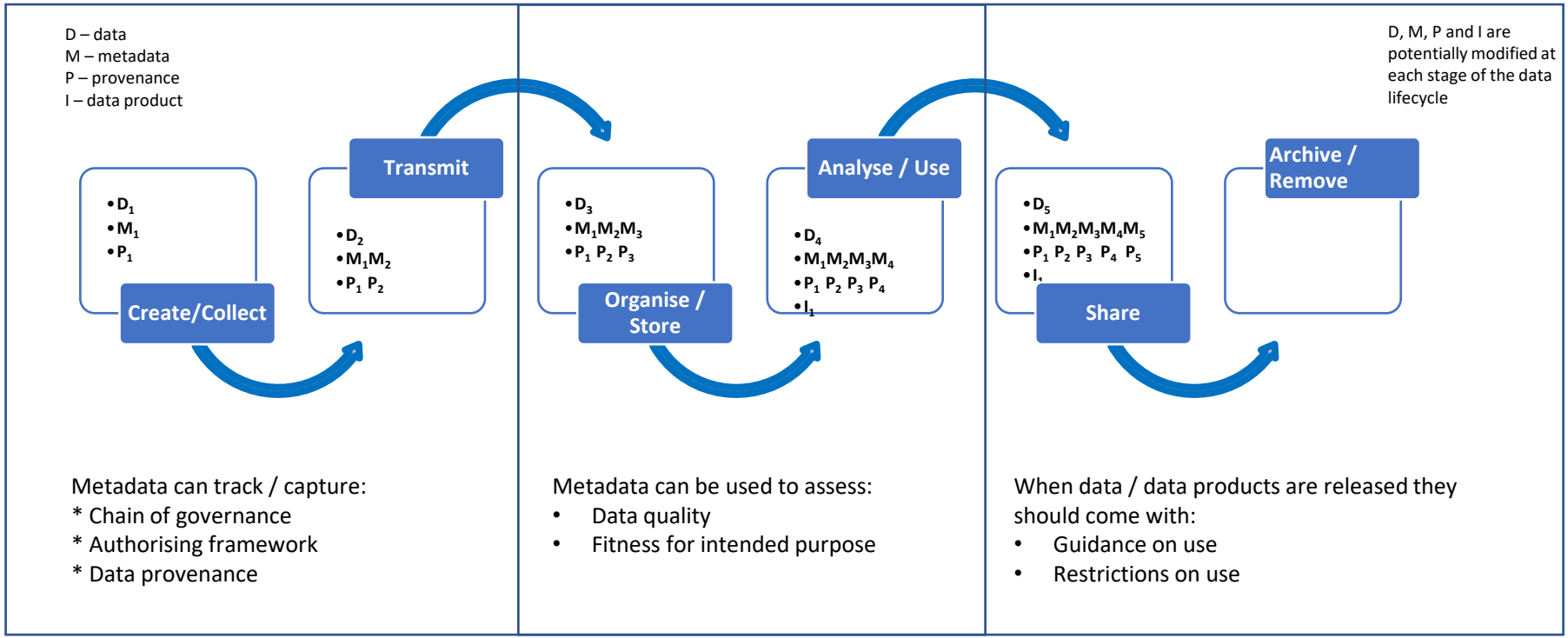
This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

10. Example data sharing frameworks



Data Lens 1 - Simplified data lifecycle



Data Lens 2 - Considerations (Risk Factors) for Data Use

Sensitivities about data itself:

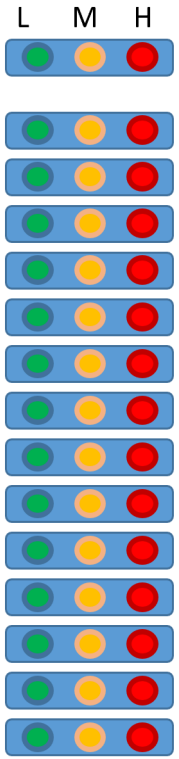
1. Concerns that data contains high levels of personal information
2. Concerns that data contains uniquely identifiable individuals
3. Concerns that sensitive subjects are captured in data (culturally subjective but often described e.g. religion)
4. Concerns about data quality (accuracy, timeliness, completeness, and consistency)
5. Concerns about fitness-for-purpose of data for analysis

Sensitivities about capability and governance:

6. Concerns that context is not captured with data (metadata, provenance, consent)
7. Concerns about authority to share data for analysis
8. Concerns about poor governance or accidental release of data or insights (outputs)
9. Concerns that expert knowledge / context is required to appropriately interpret data and results of analysis
10. Concerns about authority to release results of analysis

Sensitivities about use of insights:

11. Concerns about the level confidence in outputs (accuracy, precision, consistency, explainability, bias)
12. Concerns about unintended consequences from how outputs (insights / data driven decisions) will be used
13. concerns about whether human judgement will be applied before an insight becomes a decision
14. Concerns possible harms resulting from use of outputs (reversable, reversable with cost, irreversible)
15. Concerns that results from analysis may lead to negative surprises (especially for data not analysed before)
16. Concerns that commercial value may be degraded if insights are shared

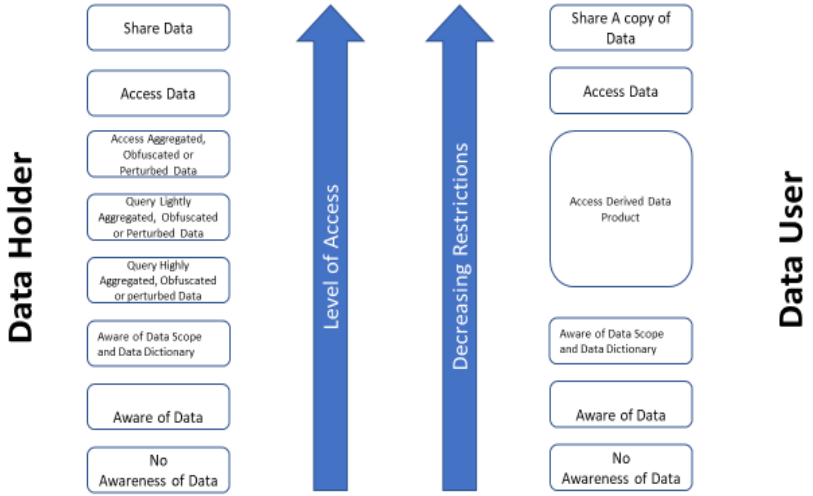


PIF: L M H

Inherent Sensitivity: L M H

Data Lens 3 – Different levels of sharing or accessing data

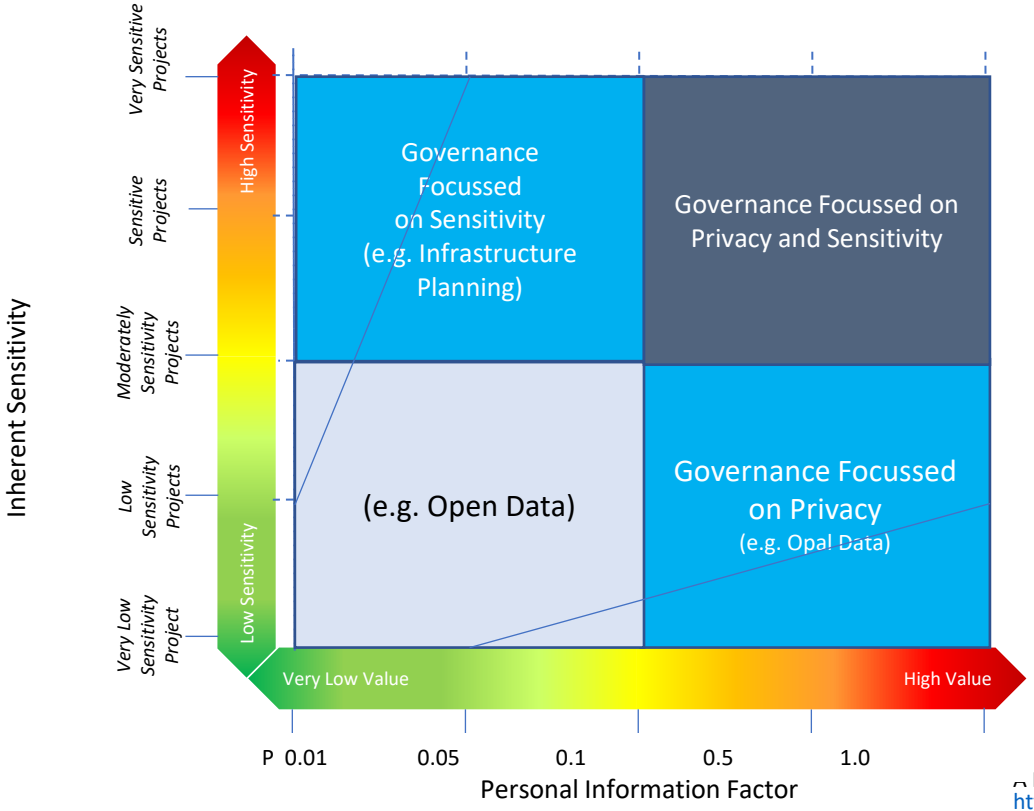
Data sharing and use can involve more than taking a copy of data and using or analysing without oversight. Different degrees of access can be provided, from none (most extreme), allowing access to prepared data products (including insights or aggregations), limited analysis access, to providing a copy of the data without restriction. These various modes of sharing allow increasing (or decreasing) levels of control depending on the sensitivities or risks associated with the data.



Data products are created from data. They can be aggregated versions, subsets of original data, perturbed data, an insight, chart, dashboard or any other result of use of data.

Data products may have different levels of inherent sensitivity and different levels of personal information compared to the original data asset.

Governance Lens 1 – Governance focusses on sensitivity versus personal information content



This governance framework should be applied to think about where mitigations should be applied across the data lifecycle based on

- The level of personal information in a data set
- The inherent sensitivity of the data itself or the use of that data
- Levels of protection which must be applied to data products produced at the “Use” step of the data lifecycle.

A PIF tool demonstration video is available at https://www.youtube.com/watch?v=wrD6FI2U4Rs.

An open source PIF tool is available at https://github.com/PIFtools/piflib.

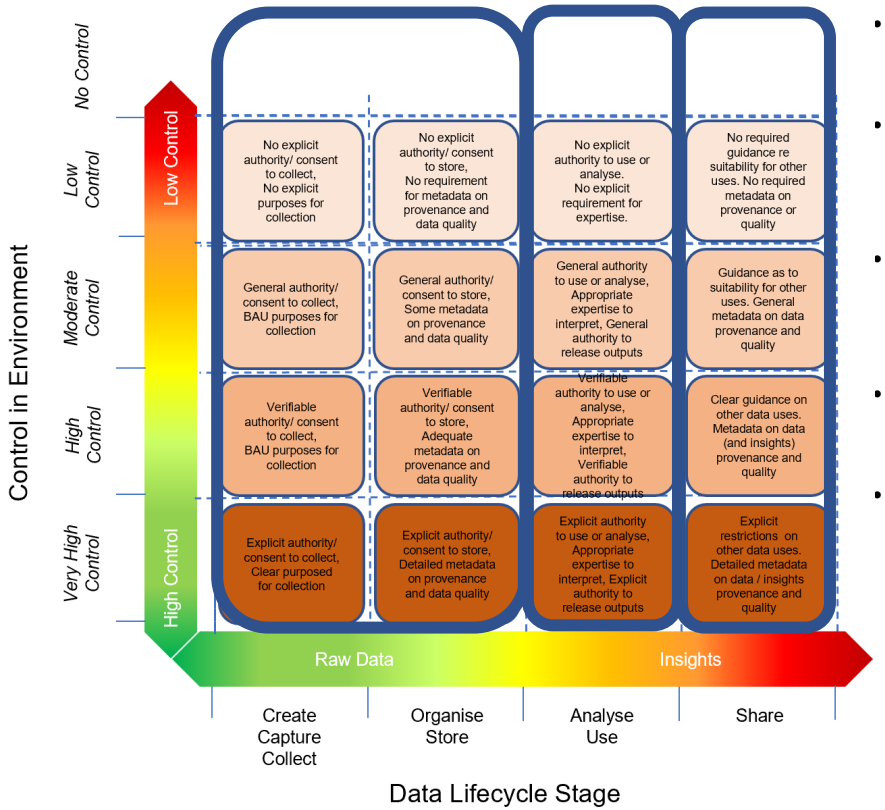
Governance Lens 2
Level of control required over data lifecycle

May have assumed authority to collect, use, and Use data. May have metadata on data provenance and quality. **Data** - low PIF.

Must have understanding of data quality and provenance, capable analysists and domain experts, adequate governance / security at each stage. **May have** broad authority to collect, use, and Use data. **Data** - moderately sensitive / moderate PIF.

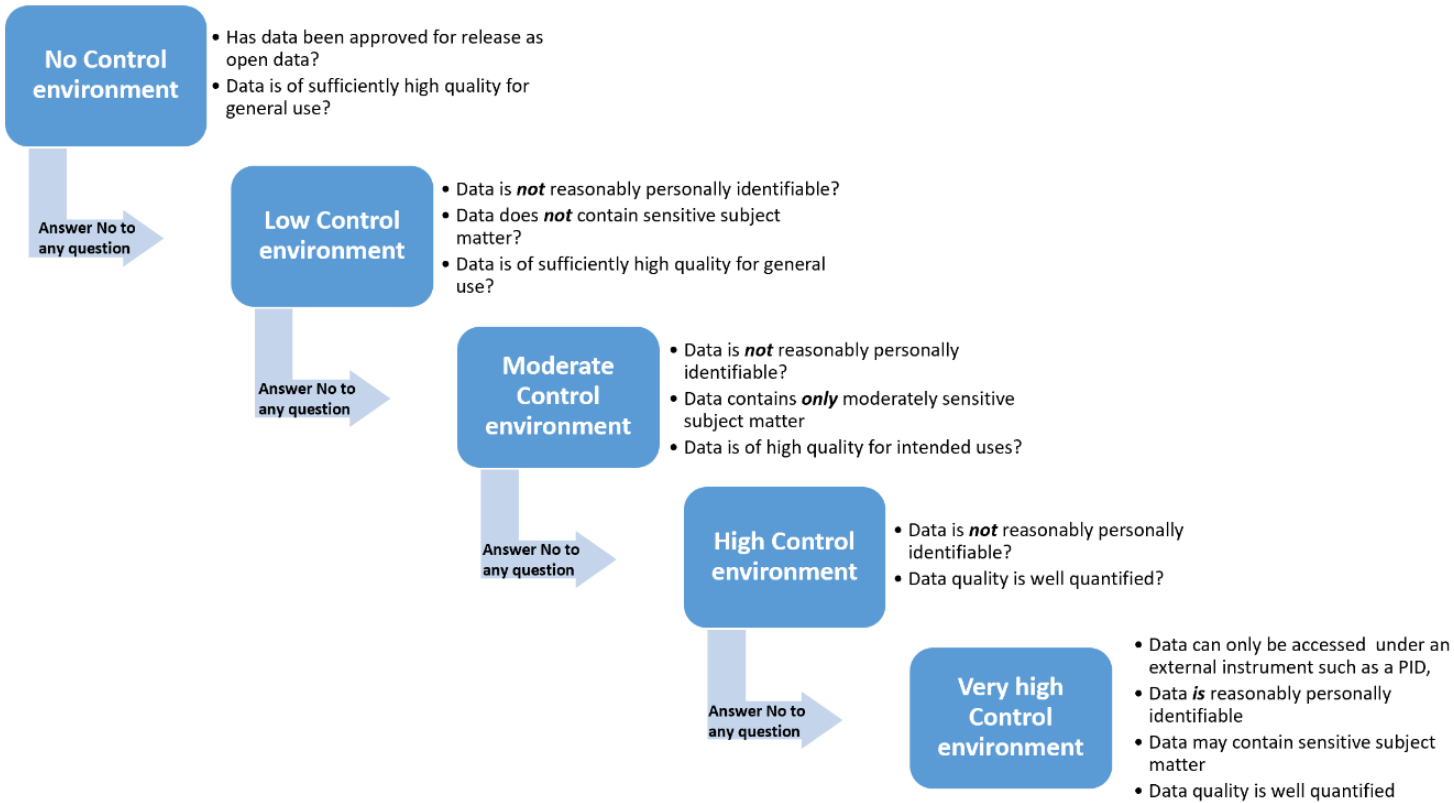
Must have understanding of data quality and provenance, highly skilled analysists and domain experts, strong governance / security at each stage. **May have** general authority to collect, use, and Use data. **Data** - high sensitivity / high PIF.

Must have explicit purpose and authority, high quality data and metadata, expert analysists and domain experts, strong governance / security at each stage. Explicit restrictions on secondary use of data and insights. **Data** - very high sensitivity and very high PIF

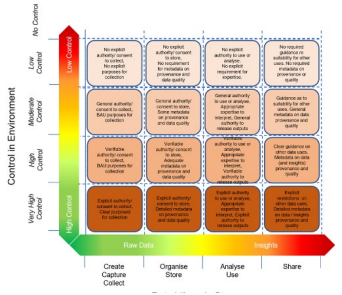


- Control = (proven) capability * (assessable) governance * (verifiable) purpose
- Capability includes skill in all stages of Data Lifecycle - data analysis, data provenance, governance, security
- High Control = skilled people working in strong governance environment with clearly authorised purpose
- No Control environment = no assessments or no restriction on people accessing or utilising data
- Requires an objective, repeatable, standardised assessment of
  - capability,
  - governance,
  - purpose,
  - data quality and provenance
  - sensitivity of data
  - degree of personal information contained in datasets

Determining the level of control required



Levels of control



A Very High Control environment: **Must have**

- explicit purpose and authority to access and use data,
- expert users experienced with the data of the quality provided and with associated metadata,
- expert analytical capability and domain expertise,
- strong governance and security at each stage of the lifecycle,
- explicit restrictions on release of data and insights, or secondary use of data and insights.
- People have met General expertise requirements as well as Project specific requirements for a “Safe Person”, and who agree to be bound by limitations on data access and use.

Suitable for:

- Data which can only be accessed under an external instrument such as a Public Interest Direction (PID),
- Data which is reasonably personally identifiable
- Data which contains sensitive subject matter
- Data which has a well quantified quality (need not be high quality)

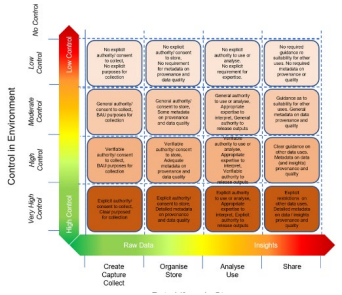
A High Control environment **Must have**

- explicit purpose and authority to access and use data (although may not have Project specific requirements),
- expert users experienced with the data of the quality provided and with associated metadata,
- very skilled analytical capability and domain expertise,
- strong governance and security at each stage of the lifecycle,
- explicit restrictions on release of data and insights, or secondary use of data and insights,
- People with access have met General expertise requirements for a “Safe Person” and who agree to be bound by limitations on data access and use.

Suitable for:

- Data which is not reasonably personally identifiable
- Data which contains sensitive subject matter
- Data which has a well quantified quality

Levels of control



A Moderate Control environment **Must have**

- general purpose and authority to access and use data (such as an authorising regulatory framework),
- experienced users dealing with the data of quality provided and with associated metadata,
- skilled analytical capability and domain expertise,
- strong governance and security at each stage of the lifecycle,
- general restrictions on release of data and insights, or secondary use of data and insights,
- People with access have met General requirements for a “Safe Person” and agree to general conditions on data access and use.

Suitable for:

- Data which is not reasonably personally identifiable
- Data which contains some sensitive subject matter
- Data which is of sufficiently high quality for the intended use

A Low Control environment **May have**

- no explicit authority to collect and use data, but no known restrictions to use data,
- users with some experience dealing with data of the quality provided,
- users with some analytical capability and domain expertise,
- appropriate governance and security at each stage of the lifecycle.
- May not have restrictions on release of data and insights, or secondary use of data and insights

Suitable for:

- Data which is not reasonably personally identifiable
- Data does not contain sensitive subject matter
- Data which is of sufficiently high quality for general use

No Control environment suitable for:

- **May have** no controls in place.
- **suitable for:**
- Data which has been approved for release as open data
- Data which is of sufficiently high quality for general use