## Data and Implementation

## Why it matters

Data is essential for enabling AI. Without high-quality, well-governed data, some of the best-designed AI system can fail. Implementation planning matters just as much. Public service AI systems must be phased, careful, and evidence-based.

In New Zealand’s public service, agencies are expected to:

* Use data responsibly, ethically, and with respect for privacy.
* Uphold the principles of data stewardship and mana whakahaere (self-determination over data).
* Implement new technologies through pilots, monitoring, and ongoing improvement.
* Consider the whole-of-life cost, including implementation, data storage, compute requirements and ongoing operational oversight.

Strong data and implementation planning ensures that AI systems are accurate, ethical, adaptable, trustworthy, and safe for public use.

## How to plan for data and implementation

When preparing the Data and Implementation section, your team may consider the following:

* Describe the data needed:
* Specify what types of data will be required to train, test, and operate the AI system, depending on what approach you are taking to use, adapt or build AI.
* For example:
	+ Text data (e.g., applications, reports)
	+ Image data (e.g., scanned documents)
	+ Operational data (e.g., processing times, error rates).
* Assess (or plan to assess) data quality and suitability: Good AI needs high-quality data that is accurate, relevant, complete, representative, up-to-date
* Respect data sovereignty and ethics: Be clear about how you will manage Māori data, personal information, and sensitive records.
* Consider the implementation plan - Perhaps a phased approach:
	+ Prototype: Build a small-scale version to test viability.
	+ Pilot: Trial the system with real users in a controlled setting.
	+ Scale: Expand carefully based on lessons learned.
* Build in monitoring considerations from day one: Implementation must include plans for ongoing monitoring, evaluation, and adaptation, not just a launch date.
* Research what the cost of the system is at the outset and try to understand the total cost over time.

## Example of Strong Data and Implementation Planning

Data needs:

* Historical hardship grant application forms.
* Applicant demographic data (anonymised and consented).
* Outcome records (approval, rejection, reasons).

Data quality checks we will consider:

* Audit of historical datasets for completeness and bias.
* Engagement with Māori data experts to review data relevance and fairness.

Privacy and sovereignty:

* All personal identifiers removed.
* Māori applicants consulted on governance of their data use.

Implementation plan:

* Prototype: Build and test summarisation AI model on a 5% sample.
* Pilot: Deploy in one regional office for 6 months.
* Scale: Expand nationally if pilot success criteria met.

Monitoring:

* Regular reviews of system accuracy, fairness, and user feedback.

## Example of Poor Data and Implementation Planning

**Data and Implementation:** We will use whatever existing data we can find and launch the AI system across the country in one go.

Why this is poor:

* No assessment of data quality, ethics, or fitness for purpose.
* No safeguards for privacy or data sovereignty.
* No phased testing or learning.
* High risk of failure, public harm, or loss of trust.

**Common Traps**

* Underestimating data preparation effort: Data cleaning, quality assessment, and governance take time and expertise, far more than many projects initially plan for.
* Ignoring data bias: If historical data reflects inequities, the AI system will learn and repeat them unless corrected.
* Assuming pilot success guarantees full-scale success: Scaling up introduces new risks (e.g., different user groups, different operating environments).

## Summary Checklist

1. Question: What data is needed for which purpose, and where is it coming from?

Purpose: Defines inputs and risks early

1. Question: Is the data good enough?

Purpose: Ensures AI will perform ethically and accurately

1. Question: How will data governance, sovereignty and privacy requirements be met?

Purpose: Upholds public trust and obligations

1. Question: What are the phases of implementation?

Purpose: Ensures safe, incremental rollout

1. Question: Have you considered the whole-of-life cost for your AI system?

Purpose: Ensures the system is cost-effective and will endure over time