# Principle 5: Accountability

Let's see how this principle guides responsible AI practice.  We'll look at a key risk this principle is designed to address and explore a brief practical example of how this principle relates to an AI use case.

**Definition:** “AI use within the Public Service should be subject to oversight by accountable humans with appropriate authority and capability at every stage. This should include the application of relevant

regulatory and governance frameworks, reporting, auditing and/or independent reviews. Agency AI capabilities need to keep pace with technological changes, to maintain a strong understanding of AI systems and their limitations.”

**Let’s focus on: Human oversight:**

While AI opens an increasing range of activities to automation, it is critical that human oversight be maintained over any AI system. This human oversight relates to both the:

* AI system directly, e.g. a human reviews output before it is passed on to the next step in the process, and the
* Overall environment that the system operates in, e.g. a business expert is involved in use case selection and risk rating

**Practical example:**

A help centre is looking to automate tax advice provided to the New Zealand public but is concerned that advice will be given incorrectly. It is decided that automated advice will only be generated for workflows / tax types in which vary little variation exist, whereas other advice will be pre-drafted for human review but not sent without approval.

An expert within the agency provides input on the types of advice generally given, and which workflows have strong automation potential vs which need more human review. The expert co-designs an ongoing process for sample review of the automated workflows to ensure they keep performing as expected.

*To achieve trust in the AI systems that they implement, agencies need effective AI governance that encapsulates the principles, guidance and regulation they aim to adhere to*

So what: Key takeaways for evaluating use cases.

**Does the use case...**

* Ensure human oversight at every stage of AI system operation?
* Include humans in reviewing AI outputs before they progress in processes?
* Engage business experts in use case selection and risk assessment?
* *Implement effective AI governance aligned with principles, guidance, and regulations?*
* *Facilitate co-design of processes for ongoing review and performance monitoring of AI workflows?*