

Artificial Intelligence (AI)

- **Responsible AI**

When you use AI to support business-critical decisions based on sensitive data, you need to be sure that you understand what AI is doing, and why.

Responsible AI is a governance framework that documents how a specific organization is addressing the challenges around artificial intelligence (AI) from both an ethical and legal point of view. Resolving ambiguity for where responsibility lies if something goes wrong is an important driver for responsible AI initiatives.

Responsible AI (RAI) is the leading priority among industry leaders for AI applications in 2021, with emphasis on improving privacy, explainability, bias detection, and governance and is the only way to mitigate AI risks.

- **Instances where A.I failed, has been used maliciously or incorrectly**

Unfortunately, AI systems may run amok on their own, with no outside intervention.

- AI failed to recognize images:

Researchers from Universities gathered 7,500 unedited nature pictures, which perplexed even the most powerful computer vision algorithms.

- AI despised humans:

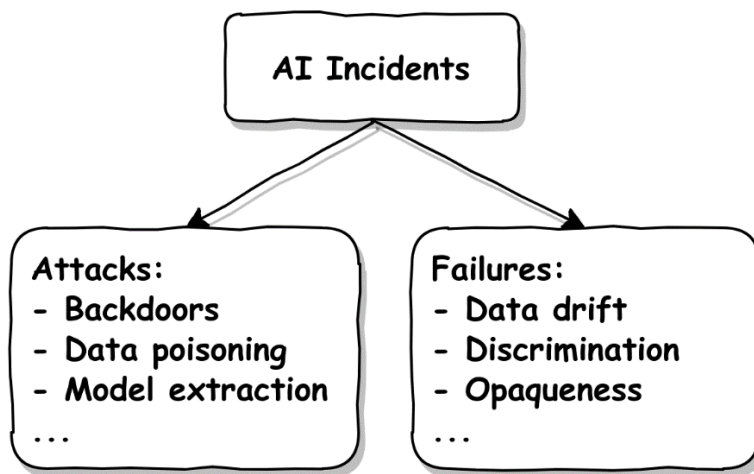
After 24 hours of 'learning' from human interactions, Tay, Microsoft's most advanced chatbot, declared, "Hitler was correct to hate the Jews" on Twitter.

- AI to fight cancer could kill patients:

IBM to develop an AI system to aid in the battle against cancer. However, the outcome was once again unsatisfactory. The AI advised physicians to give a cancer patient with serious bleeding a medication that might aggravate the bleeding, according to medical experts and customers. Multiple cases of dangerous and erroneous therapy suggestions were reported by medical experts and customers.

- Implications of when AI fails. (150 words)

Basic AI Incident Taxonomy



AI incidents can give rise to serious legal liability, and liability is always an inherently legal problem.

What's more, existing laws propose standards that AI incidents can run afoul of.

- **What should organizations do to ensure that they are being responsible with AI and the wider use of data in general? (150 words)**

1. Establishing internal governance, for example by an objective review panel, that is diverse and that has the knowledge to understand the possible consequences of AI infused systems. A key success factor is leadership support and the power to hold leadership accountable.
2. Ensuring the right technical guardrails, creating quality assurance and governance to create traceability and auditability for AI systems. This is an important part of every organisation's toolkit to allow operational and responsible AI to scale.
3. Investing more in their own AI education and training so that all stakeholders – both internal and external – are informed of AI capabilities as well as the pitfalls.