**RESPONSIBLE ARTIFICIAL INTELLIGENCE**

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. Often, bias can be introduced into AI by the data that's used to train machine learning models. When the training data is biased, it naturally follows that decisions made by the programming are also biased. Resolving ambiguity for where responsibility lies if something goes wrong is an important driver for responsible AI initiatives.

Some instances of when AI has failed include:

* **AI to fight cancer could kill patients:** The AI system by IBM, according to a doctor at Jupiter Hospital in Florida, was a complete failure. Dangerous and erroneous therapy suggestions were reported by medical experts and customers.
* The first revolutionary Henn-na Hotel opened its doors to visitors in Japan in 2015. All of the hotel’s employees were robots, including the front desk, cleaners, porters, and in-room helpers. However, the bots quickly accumulated consumer complaints: they regularly broke down, we’re unable to offer adequate responses to visitor questions, and in-room helpers frightened guests at night by misinterpreting snoring as a wake command. The hotel group that owned the hotel finally got rid of the last of its unreliable, costly, and irritating bots, replacing them with human staff after years of effort.
* **AI-driven cart malfunctioned on the tarmac:** On the tarmac, an AI-driven food cart malfunctioned, circling out of control and ever-closer to a vulnerable Aeroplane parked at a gate. Finally, a yellow-vest worker was able to stop the cart by hitting it with another vehicle and knocking it down.

### Strategies for responsible AI

* Ensure data is explainable in a way that a human can interpret
* Ensure design and decision-making processes are documented to the point where if a mistake occurs, it can be [reverse-engineered](https://searchsoftwarequality.techtarget.com/definition/reverse-engineering) to determine what transpired.
* Build a diverse work culture and promote constructive discussions to help mitigate bias
* Create a rigorous development process that values visibility into each application's latent features.

### Best practices for responsible AI

The steps required to prevent discrimination and ensure transparency vary from company to company. In a large enterprise, the chief analytics officer ([CAO](https://searchcio.techtarget.com/definition/Chief-analytics-officer)) is typically tasked with developing, implementing and monitoring the organization's responsible AI framework. When designing responsible AI, governance processes need to be systematic and repeatable. Some methods for best practices include:

* Implement [machine learning best practices](https://internetofthingsagenda.techtarget.com/tip/3-machine-learning-best-practices-to-use-in-IoT-projects).
* Creating a diverse culture of support. This includes creating gender and racially diverse teams that work on creating responsible AI standards.
* Make a best effort towards transparency and make the work as measurable as possible so that any decisions made by AI are explainable.
* Use responsible AI tools to inspect AI models. Options such as explainable AI and the TensorFlow toolkit are available. In addition, perform tests such as bias testing or [predictive maintenance](https://whatis.techtarget.com/definition/predictive-maintenance-PdM).
* Stay mindful and learn from the process.