# Data Privacy: A Friend or Foe of Artificial Intelligence

**August 15, 2024**

*Pittsburgh, PA*

***TEQ Hub***

(by **Kristen Petrina**)

Artificial Intelligence (AI) is advancing at unprecedented speeds. AI relies on vast amounts of datasets for processing and model training, creating the challenge of balancing the benefits of AI, while protecting data privacy. As a result of improper data processing and usage, organizations are facing harsh penalties including AI usage prohibition, algorithm disgorgement, and multibillion dollar fines. When considering how to introduce AI into any organization, one of the first questions to consider is, "How can AI be utilized to drive innovation without violating privacy and misusing collected data?"

AI governance analysis should be under a privacy lens, as personal data is at the core of many opportunities that come with AI development. Privacy risks may result in societal and ethical impacts on individuals which speaks to the heart of responsible AI usage. Incorporating responsible AI practices is user specific to each organization and it is possible to protect privacy and drive innovation. In order to achieve both goals, organizations should consider data protection preventative measures before implementing AI into its processes.

- Data privacy should be addressed at the onset of AI implementation. Organizations should conduct risk assessments and consider data enablement through AI from the beginning before it becomes an issue. Generative AI in particular is self-learning, the more data fed into the model, the harder it will be to unwind or remove data if improperly used.
- AI and data privacy governance teams must work together from the beginning to address any risks that may arise. Organizations may consider forming an ethical AI committee engaging diversified team members to reduce potential bias in the development and design.
- Contemplate data inputs by asking questions such as what the existing and potential future data sources may be, what data will be collected, what are in the datasets, how to categorize the types of data, will the data modeling receive personal or sensitive data, should those things be included.
- Consider data outputs by asking questions such as what information will be displayed after processing, what is the impact of the processing, is there any potential for harm from the processing and results, what controls are needed at the data layer to mitigate the risks.
- Review regulatory and data privacy requirements that impact and influence AI to assess and address any privacy policy gaps as a result of the introduction of AI into the organization's processes. Policies can include but are not limited to addressing transparency into training data origins, acceptable use policies, data quality, validation of algorithms to confirm the AI model meets the organization's AI and data policies, and sharing or transfer of data with third parties.
- What consent did the data owner give, particularly what purpose did the owner agree to? When implementing AI, organizations can get ahead of consent issues by educating the data owner of the intended purpose and use of the data.
- Build privacy measures into the system's architecture to guarantee alignment with purpose consent given by the data owner and careful treatment of the data.
- Is the value provided to the organization proportional to the data owners risk? If data is used, should it be minimized to strip identifying features, or is it essential to include information such as sensitive data to determine if the model is biased?

- Determine the permanence of the data, depending on how it is incorporated into the AI model, an enforcement action can result in the loss of years of data. Additionally, some states allow data owners to be forgotten. If data is not de-identified it is possible to remove it from the datasets, however, if the data is de-identified it will not be possible to determine how the data was used and for the data to be removed from the datasets, potentially requiring a retraining of the model.
- Implement data security and privacy controls for stored decommissioned AI systems and associated data.

Organizations must safeguard data and ensure privacy compliance within AI systems, however, that does not mean innovation cannot thrive. An organization that considers privacy from the onset of AI implementation can drive innovation while also protecting privacy and reducing the risk of future penalties.

To view the full article, **click here**.