AT&T

AT&T, Inc.

Artificial Intelligence (AI)
Risk Self-Assessment Guidance and Recommended Questions

As outlined in the **AT&T Artificial Intelligence (AI) Policy**, for each AI Solution (initial or repurposed), an AI Risk Self-Assessment must be completed. AI Solutions must be categorized as no, low, medium, or high risk.

- AI Risk Self-Assessments will focus on Key Risk Areas (listed below).
- If the AI Risk Self-Assessment identifies no risk or low risk, the proposed use of the AI Solution may proceed without further review.
- If the AI Risk Self-Assessment identifies a medium risk or higher, the proposed use of the AI Solution may not proceed absent further review and approval by the Data & AI Governance Review Board (US) or legal team (non-US).
- All AI Risk Self-Assessments must be in writing and be available for audit or review for 5 years, per AT&T's retention policy.

Here are the suggested sample questions of an AI Risk Self-Assessment.

1. Will the AI Solution be used in a regulated environment? e.g. FCRA, CPNI, EEOC
2. [For the Unfair Bias Key Risk Area] Is there a potential adverse impact to protected classes from the use of the output of the AI Solution?
3. [For the Consequential Use Case Key Risk Area] Could this use of an AI Solution potentially deny or diminish access to consequential services or employment, with meaningful and significant impact (e.g., hiring, promotions, denied Mobility services or devices, denying a discounted service, digital divide/redlining, network prioritization)
4. [For the Health & Safety Risk Area] Is there a risk to the health or safety of individuals, posed by use of the AI Solution?
5. [For the Privacy Key Risk Area] Is use of personal information consistent with AT&T's privacy policies? Is additional consent required? Does the use of the AI Solution process or collect personal or sensitive information or address a sensitive topic? The more sensitive the personal information (e.g., facial recognition), the higher the risk.
6. [For Net Risk Ratings] If the AI Solution doesn't work properly or fails to work, what are the consequences? Is there a fallback plan proposed? Example: self-driving car that doesn't work properly puts human lives in danger, a self-cleaning vacuum has low consequences/risks if it fails.
7. Are the information security measures protecting the use of the AI Solution proportionate to the potential consequences of it being hacked? Example: a

hacked coffee machine is far less consequential than a hacked communications network for a hospital.

8. What ongoing monitoring and controls are in place to prevent, detect and remediate AI Solution outcomes that are inaccurate or unfairly biased or otherwise harmful?
9. Does the AI Solution follow target architecture?
10. Does the AI Solution contribute back to AT&T's knowledge base or is there a reason the AI Solution's knowledge should be siloed?
11. Does the AI Solution require access controls to be used?

Can you mitigate the risk(s) identified? If no, can an escalation team get together to solve before moving forward?
Consider the potential for the risk(s) to occur. Identify what you have done to mitigate those risks.
Where high risk(s) are identified, escalate your initiative to your business unit sponsor, who will engage the appropriate parties including Legal, to get signoff of accepted risk(s) and help mitigate risk.

Artificial Intelligence (AI) Key Risk Areas

Key Risk Areas	
Law	Risks that an AI Solution output may affect rights under domain-specific laws, rules and regulations (e.g., EEOC, FCRA, GDPR), intellectual property concerns, or AI-specific laws, rules and regulations.
Unfair Bias	Risk that an AI Solution may generate outputs that are discriminatory in a manner that is either unlawful, unfair or both. Bias is an inclination of prejudice towards or against a person, object, or position. Bias can be good or bad, intentional or unintentional. Biased outputs can arise from several causes including bias in the learning or operational data sets (e.g., exclusion of data elements about a particular protected class or inclusion of data elements that effectively serve as a proxy for an ethnicity), the data selection, in the machine learning process, or in the algorithms themselves unbeknownst to the developer. *Harm from outcomes that are unfairly biased is not limited to negative monetary or financial impacts and can include other types of actions or omissions, such as a denial of service.* Domestic US only: SIFT must be completed on every new model or significant changes to the model.
Privacy	Risk that the lifecycle or output of an AI Solution may adversely affect interests under applicable privacy laws and policies. Such risks can arise in a few AI contexts including data collection, selection and usage of data, as well as AI Solution outputs.
Accessibility	Risk that use of an AI Solution may pose difficulties for those with disabilities, including when an individual interacts directly with an AI Solution, with no human intermediary.

Consequential Use Cases	Risks associated with AI Solutions that affect access to services or employment which, if denied or diminished, could have meaningful and significant impact on an individual, or group of individuals (e.g. employment, digital divide/redlining, network prioritization).
Health & Safety	Risk that an AI Solution output could diminish health or physical safety.
Cybersecurity	Risk that an AI produces, creates, or generates content that could potentially or actually affect the confidentiality, integrity, or availability of AT&T assets including but not limited to data, systems, applications, or people.

Document Owner

Chief Privacy Office, Director-Privacy
Mary Kay Thurlkill, MR2369

Date:

Effective: January 2021
Updated: July 2023