Appendix A: Data Governance for AI Checklist

Teams can use this checklist to prepare an AI use request, even if they do not have all the answers. It helps identify what is known, what remains unresolved, and who can help. Where information already exists in governance systems or metadata, teams should reuse it rather than recreate it.

Use case: _____ Project or contract: _____

Prepared by: _____ Date: _____

What will the AI actually do? The task, how much we rely on it, and the risk if it is wrong

	The use case is bounded: the task, the project, and the documents.	
	The review method and the reviewer are identified.	Reviewer:
	The timing of review, before or after use, is set.	Before or after use:
	The impact of a wrong result is understood.	
	The check that stops a wrong result from spreading or standing is identified.	Which QC step catches it:

What data will it touch? The data, who can approve its use, and the rules it carries

	The data needed for the use case is clearly identified.	Data, and where it lives:
	The person or organization responsible for the data, and the person who can approve this use, are identified.	Owner and approver, and where it is recorded:
	The data is approved for this use, and the approval is recorded.	Where the record lives:
	The AI's access is appropriate for the approved use.	Own account, or a service account?
	The AI returns to each user only what that user is allowed to see.	How user permissions are enforced:
	Any FCI in the data is identified.	If yes: contract and clause
	Any CUI in the data is identified.	If yes: category, marking, clause
	Any PII in the data is identified.	If yes: is a privacy review needed?
	The governing version of the data is identified.	Governing version:
	The data suits the proposed use case.	Accuracy, completeness, currency checked by:
	The data's source, change history, and last update are known.	
	The AI can read the data accurately enough for the task.	
	Important data preparation and transformation decisions are recorded.	Where recorded:

Where will the data be processed?
Each environment the data touches and who controls it

	Each environment is authorized for this data and this AI use case.	Basis: FedRAMP status, Impact Level, CMMC scope
	Every environment the data touches, and its operator, is listed.	
	Each point where the data leaves the team's control is marked.	
	The model, its version, and its provider are identified.	Model and version:
	A way to learn of model or service changes is in place.	Who watches, and where:
	Controls inherited from each provider, and those retained by the team, are identified.	Inherited from:

Can you show how the AI behaved with this data?
The evidence that shows what the AI did and what people did

	The request is recorded: what was asked, by whom, when, and under which approved use case.	
	The data the AI read or retrieved, and its designations, are recorded.	
	The model or service that did the work, and the actions the AI took on its own, are recorded.	
	The AI's output and where it went are recorded, and the records can be tied together to reconstruct one use.	What ties them together:
	Decisions about what evidence to keep, and why, are recorded.	
	The evidence is protected at the level required for its most sensitive content.	Most restrictive content:
	Only authorized people can view the evidence or manage the logging function.	
	Unauthorized changes to or deletion of the evidence are prevented or detected.	
	The retention period for the evidence is known.	Retention period:
	Disposal of the evidence at the end of the retention period is planned.	