

For US State & Local Government CIOs

Dark data & risk reduction worksheet

What's hiding in your archives, and what it's costing you.



What you can't see in your archives becomes your next vulnerability

"Without robust governance controls on unstructured content, sensitive information flows into AI applications, creating compliance and security exposure."

- *CDO Magazine, Unstructured Data: The Hidden Bottleneck in Enterprise AI Adoption*

The blind spot CIOs cannot afford to ignore

Archives were built to retain information, not actively govern it. Over time, they've accumulated decades of contracts, records, media and legacy formats, much of it never fully classified or reviewed.

As AI, search and analytics expand across the enterprise, dark data no longer stays hidden. Previously inaccessible content becomes discoverable without the context or controls needed to manage sensitive information safely. Rapid adoption of Microsoft 365 only accelerates this exposure, as growing content volumes collide with aging repositories and oversized SharePoint environments.

Why dark data is a growing enterprise risk

Cloud adoption has made it easy to store vast amounts of cold or low-value data. But placing unclassified content into cheap storage is a false economy. Unmanaged archives increase enterprise risk in three ways:

Security exposure

Sensitive or regulated data can surface unexpectedly through AI tools or discovery processes.

Compliance gaps

Records without clear provenance, classification or retention alignment struggle to withstand audit or legal scrutiny.

Cost creep

Redundant and unknown data drives long-term storage, retrieval and remediation costs.

For CIOs, this means the traditional definition of the attack surface is incomplete. Risk no longer resides only in live systems. It extends deep into the archive.

Turning dark data into a governed asset

- Metadata enrichment to understand what archived data actually contains
- Classification and PII detection to identify sensitive or regulated information
- Format migration to ensure records remain readable and defensible over time
- Fine-grained access controls and security policies, applied down to the individual file or asset level

Your 15-minute dark-data worksheet

Answer each question honestly. If you answer “no” or “not sure” to two or more, dark data is already an active risk in your environment.

Question	Yes	No	Not sure
Do you know how much unstructured data (files, emails, media) exists across your legacy archives and file shares?			
Can you identify which of that content contains sensitive or regulated information (PII, PHI, financial data)?			
Is that content classified, tagged or otherwise governed the same way your live systems are?			
Could your team produce a defensible answer if an auditor, regulator or AI tool surfaced that content tomorrow?			
Do you have a documented plan for reviewing and retiring dark data on a recurring basis?			



Cut your backlog up to 4x faster at 20% of the cost with Preservica's built in human-centric AI

preservica.com/roi-calculator-landing-page



info@preservica.com



preservica.com



@preservica