FORTINET

Beyond Compliance:
Building AI-Safe Data Defenses

Bill Carter, CISSP
National Director of Cybersecurity Strategy, Public Sector
bcarter@fortinet.com

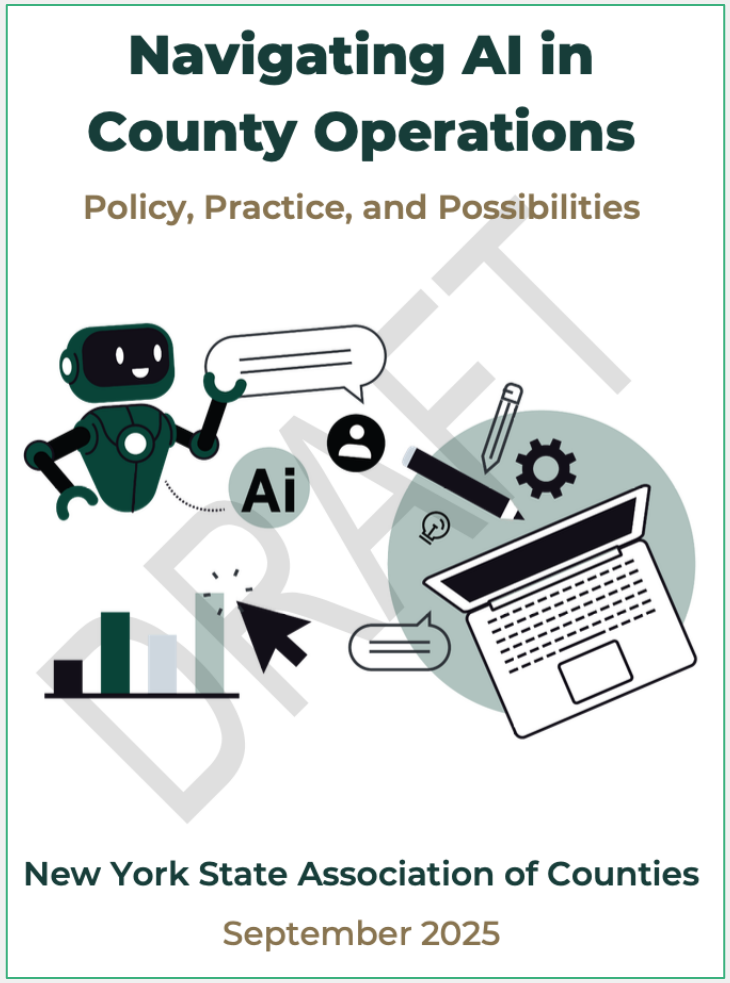

Mark LaVigne, PhD
Deputy Director
NYSAC

Key Ideas

- Realities for New York Local Government
- What does AI-Safe mean?
- Data Loss Trends
- MITRE ATT&CK and Data Loss
- Expectations of Modern DLP
- The New Perimeter
- Implementation Roadmap

- **Leadership Conversation, Not a Tech Demo**

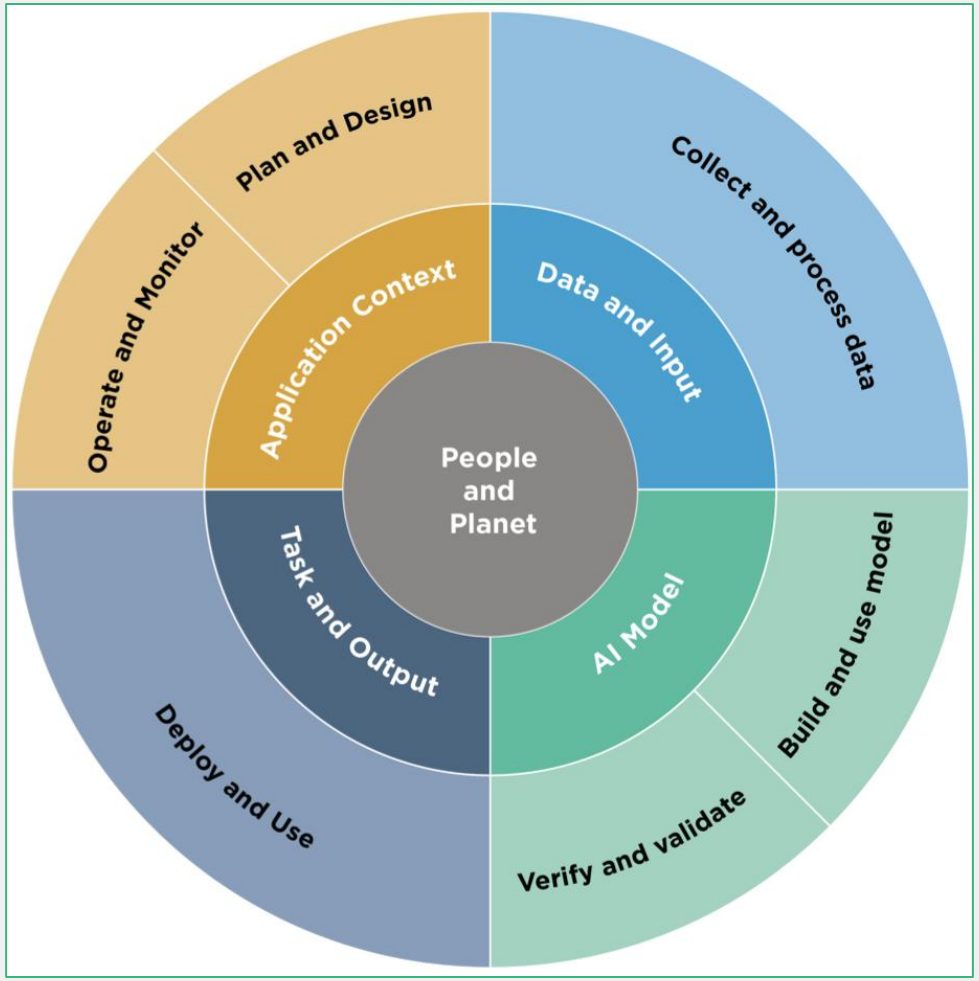
What Brings Us Together



- NYSAC Guide is Excellent in Showing How to Adopt
- Today We Cover What to Protect
- We Don’t Want Adoption to Create Irreversible Data Risk!

Compliance Is No Longer Enough

- Compliance is Static; Threats Evolve
- AI Accelerates Mistakes and Attacks
- Most Breaches Occur in Compliant Environments
- Risk Now Moves at Machine Speed



Realities for Local Government

Highly Regulated Environments

Posture is Mature, But Resources are Constrained

Shared Services and Vendors are Involved

AI Usage is Already Happening – With or Without Policy

What Does It Mean to be AI-Safe?

AI-Safe is how Governance becomes enforceable at Machine Speed

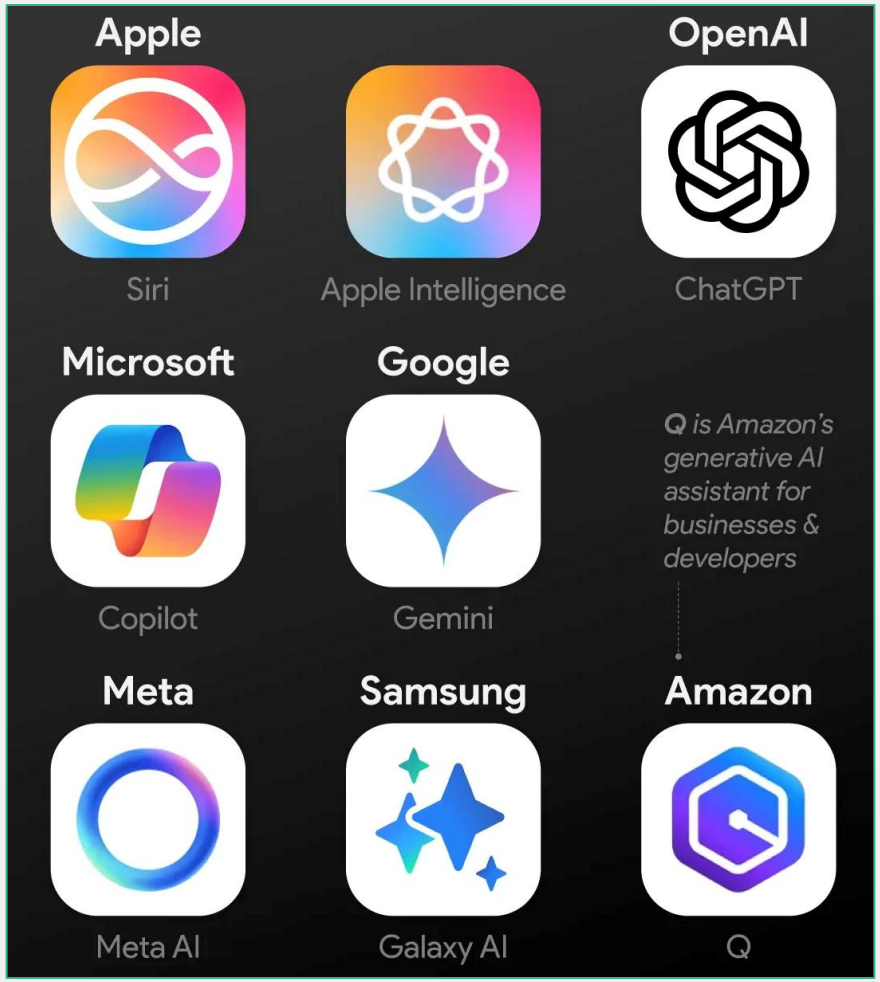
Secure AI Usage

Accidental Loss Prevention

Focused Governance

Enforced Controls

How AI Changes Data Risk



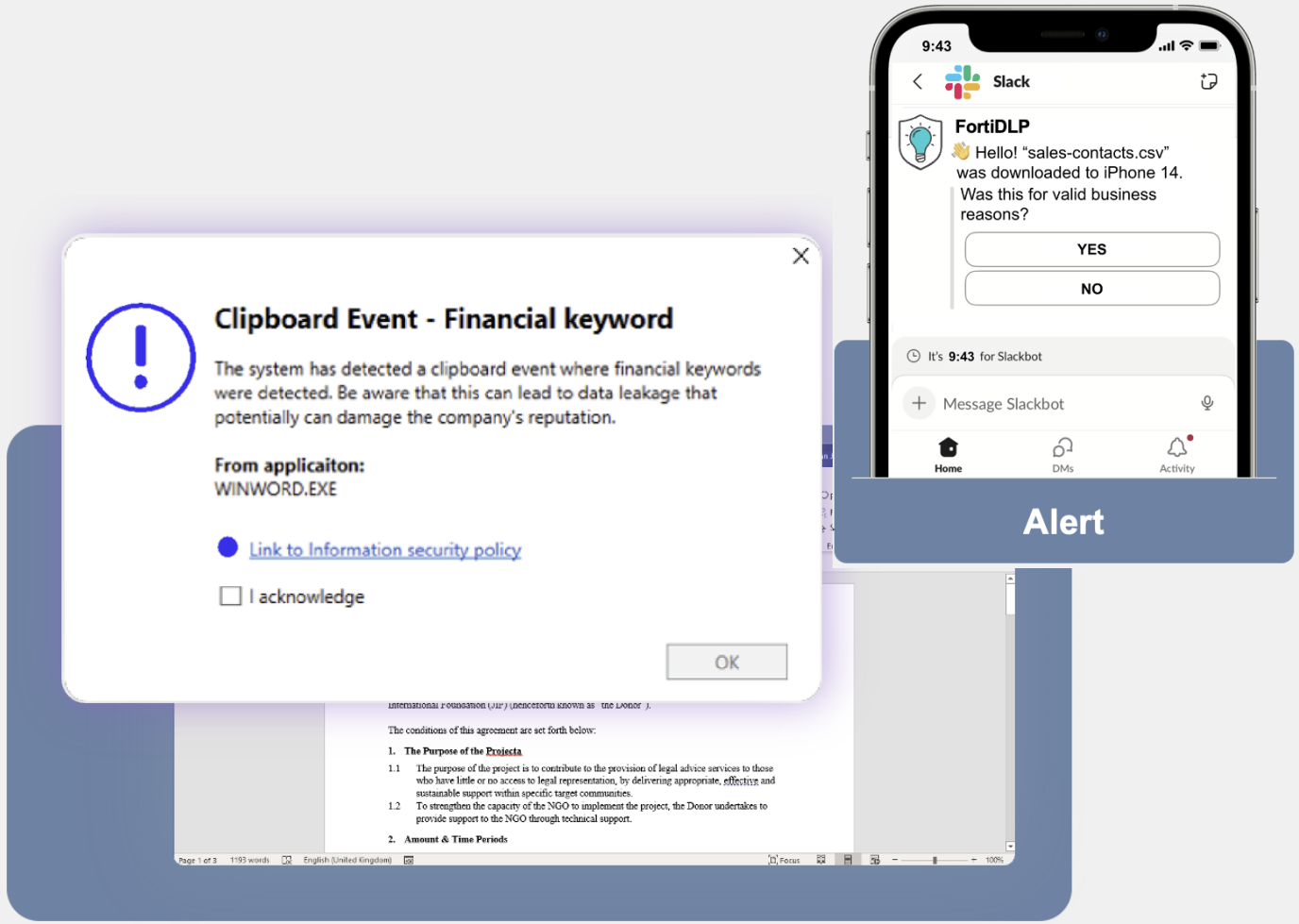
- AI Amplifies Human Error
- AI Tools Can Ingest Data Permanently
- Context-Aware Phishing
- Automated Harvesting

- And The Result Is This…

Accidental Data Loss Dominates

- Copy/Paste Into AI Tools
- Cloud Overshare
- Misclassification
- Well-Intentioned Users

- Most AI-related Data Loss is Unintentional



MITRE ATT&CK

- **MITRE ATT&CK Evolved For This Moment**
- No Longer Just Describes Attackers
- Latest Versions Describe Outcomes

- Today's Exfiltration Attempts Start With:
  - Legitimate Users
  - AI Tools

https://attack.mitre.org

MITRE ATT&CK: Accidental Exfiltration

Mitigations

ID	Mitigation	Description
M1057	Data Loss Prevention	Data loss prevention can prevent and block sensitive data from being shared with individuals outside an organization.[8] [9]
M1037	Filter Network Traffic	Implement network-based filtering restrictions to prohibit data transfers to untrusted VPCs.
M1054	Software Configuration	Configure appropriate data sharing restrictions in cloud services. For example, external sharing in Microsoft SharePoint and Google Drive can be turned off altogether, blocked for certain domains, or restricted to certain users.[10] [11]
M1018	User Account Management	Limit user account and IAM policies to the least privileges required.

- T1537 – Transfer to External Service
- T1213 – Data From Repositories
- User-Driven Cloud Sharing

Mitigations

ID	Mitigation	Description
M1047	Audit	Consider periodic review of accounts and privileges for critical and sensitive repositories. Ensure that repositories such as cloud-hosted databases are not unintentionally exposed to the public, and that security groups assigned to them permit only necessary and authorized hosts.[9]
M1041	Encrypt Sensitive Information	Encrypt data stored at rest in databases.
M1032	Multi-factor Authentication	Use two or more pieces of evidence to authenticate to a system; such as username and password in addition to a token from a physical smart card or token generator.
M1060	Out-of-Band Communications Channel	Create plans for leveraging a secure out-of-band communications channel, rather than existing in-network chat applications, in case of a security incident.[10]
M1054	Software Configuration	Consider implementing data retention policies to automate periodically archiving and/or deleting data that is no longer needed.
M1018	User Account Management	Enforce the principle of least-privilege. Consider implementing access control mechanisms that include both authentication and authorization.
M1017	User Training	Develop and publish policies that define acceptable information to be stored in repositories.

Malicious Exfiltration Evolves

- AI-Enhanced Social Engineering
  - Initial lure, credential grab
- Faster Reconnaissance
  - Find the golden egg
- Automated Exfiltration
  - Evasion of controls
- Cloud-Native Paths
  - Exfil over approved channels
- **Log In, Blend In, Walk Out**

MITRE ATT&CK: Malicious Exfiltration

- T1048 – Exfiltration Over Alt Protocol
- T1567 – Exfiltration to Cloud
- **T1020 – Automated Exfiltration**
  - **No mitigation, system-based!**
  - Compensate with SASE or ZTNA

Mitigations

ID	Mitigation	Description
M1057	Data Loss Prevention	Data loss prevention can detect and block sensitive data being uploaded via web browsers.
M1037	Filter Network Traffic	Enforce proxies and use dedicated servers for services such as DNS and only allow those systems to communicate over respective ports/protocols, instead of all systems within a network. Cloud service providers support IP-based restrictions when accessing cloud resources. Consider using IP allowlisting along with user account management to ensure that data access is restricted not only to valid users but only from expected IP ranges to mitigate the use of stolen credentials to access data.
M1031	Network Intrusion Prevention	Network intrusion detection and prevention systems that use network signatures to identify traffic for specific adversary command and control infrastructure and malware can be used to mitigate activity at the network level.
M1030	Network Segmentation	Follow best practices for network firewall configurations to allow only necessary ports and traffic to enter and exit the network.[12]
M1022	Restrict File and Directory Permissions	Use access control lists on cloud storage systems and objects.
M1018	User Account Management	Configure user permissions groups and roles for access to cloud storage.[13] Implement strict Identity and Access Management (IAM) controls to prevent access to storage solutions except for the applications, users, and services that require access.[14] Ensure that temporary access tokens are issued rather than permanent credentials, especially when access is being granted to entities outside of the internal security boundary.[15]

The Hard Facts

Human Error is the Predominate Cause of Data Loss to AI

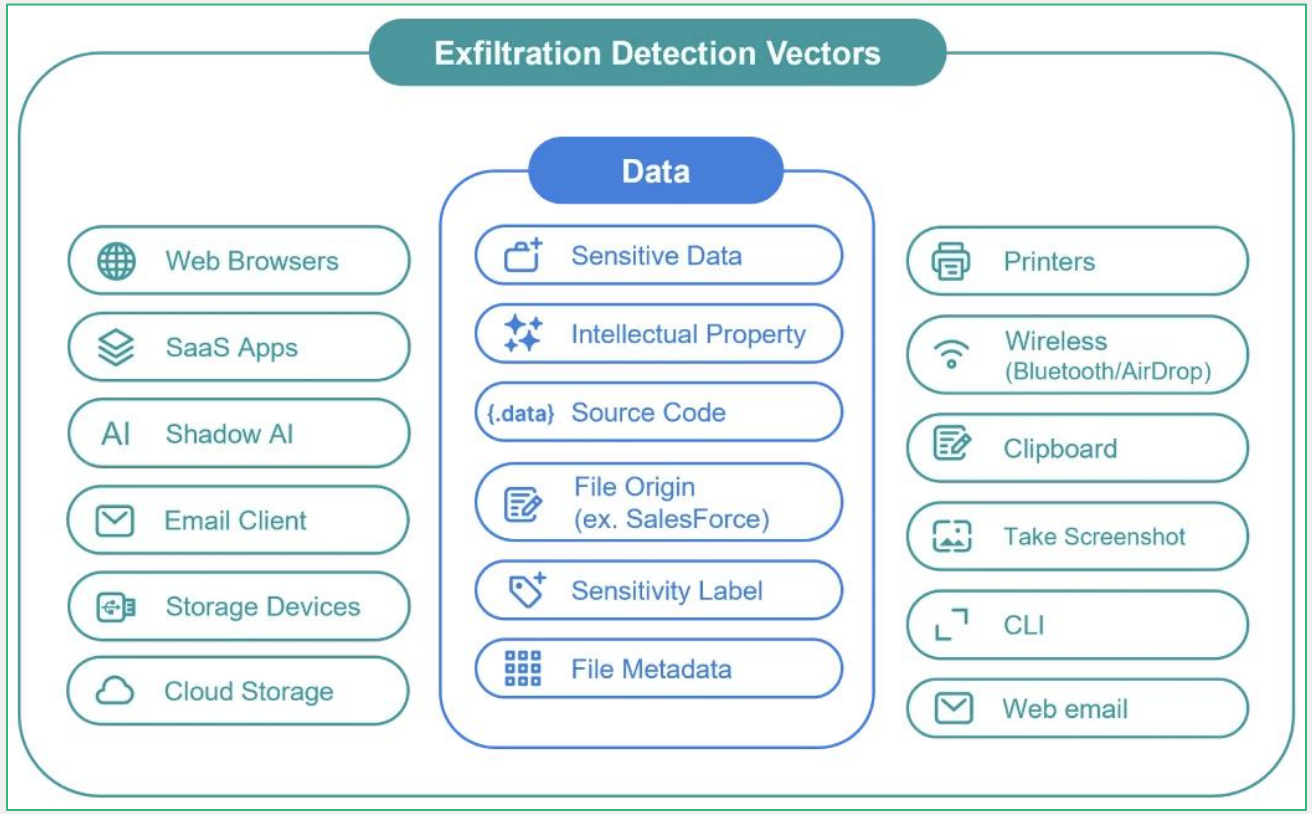
AI Accelerates Accidental Exposure More Than Malicious Theft

Malicious Exfiltration is Fewer in Number, but Higher in Impact

Modern DLP Must Prioritize Accidental Loss Prevention First

Modern DLP Requirements

Not About Blocking AI – Making AI Safer to Use



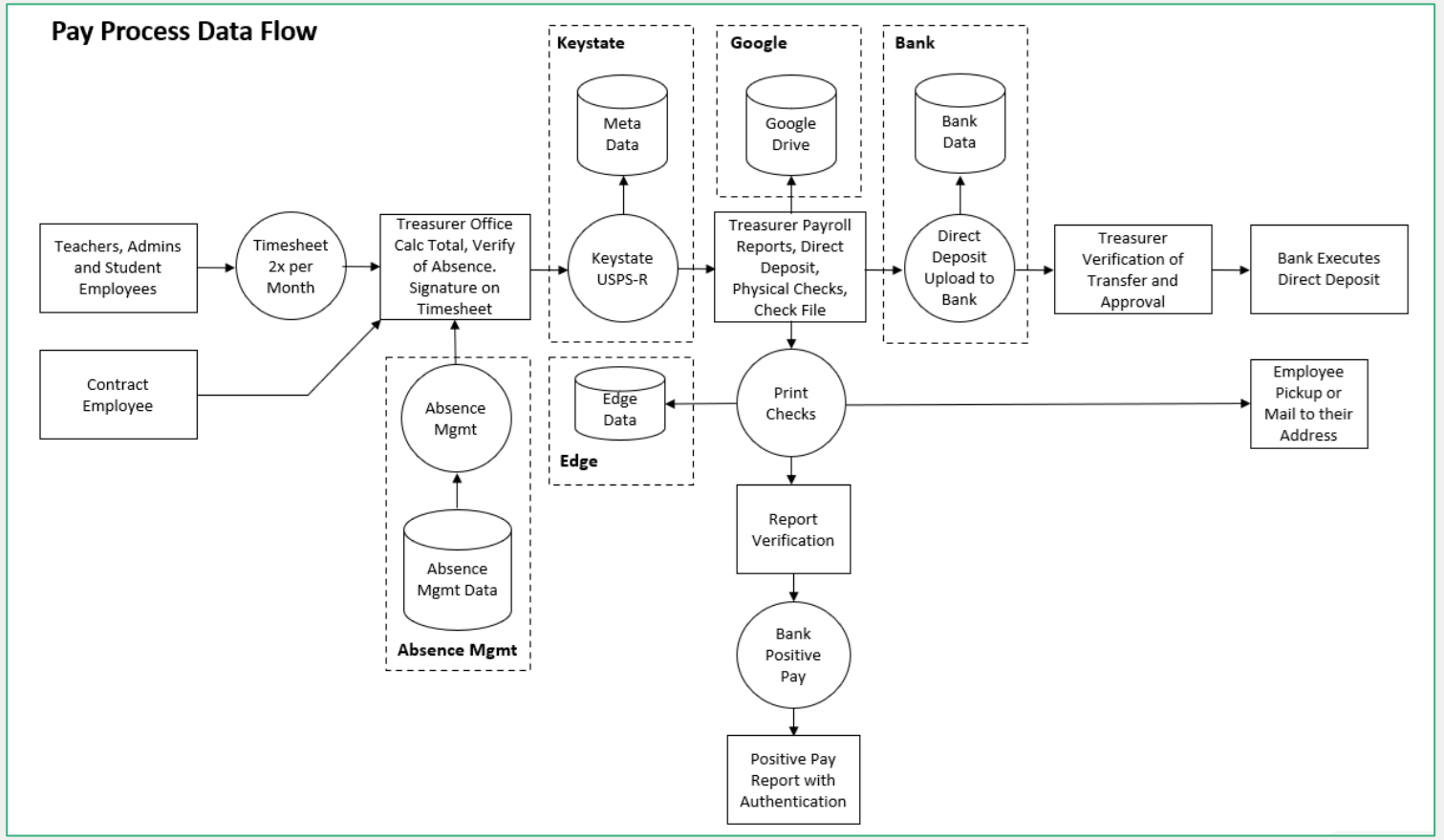
- Understand Intent
- Span Endpoint, Cloud, AI
- AI Enriched Content
- Act in Real Time
- Learn Continuously

The New Perimeter: Metadata

Policy Without Data Control Cannot Succeed

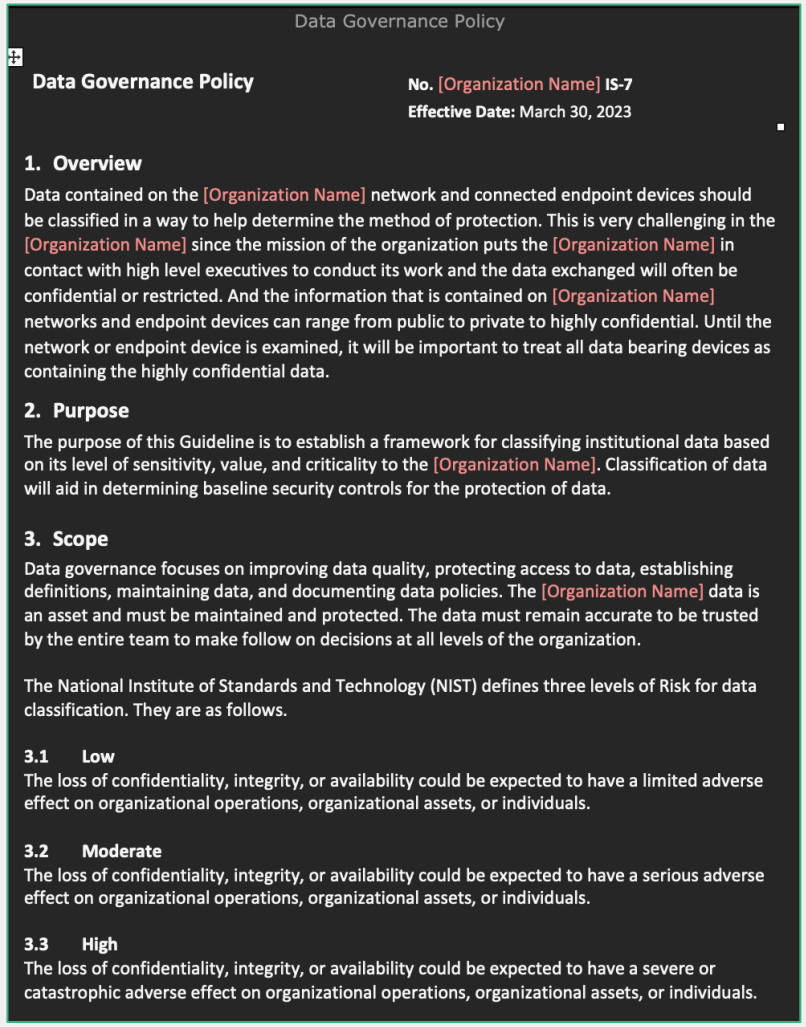
- Data Lineage
- User Behavior Context
- Sensitivity Follows Data
- AI Enriches Metadata
- Flow is Everything

- **Governance is Intent, Control is Protection**



AI-Safe Governance Principles

- Clear AI AUP
- Data Classification
- Least Privilege
- Continuous Tuning



Data Governance Policy

Data Governance Policy

No. [Organization Name] **IS-7**
Effective Date: March 30, 2023

1. Overview

Data contained on the [Organization Name] network and connected endpoint devices should be classified in a way to help determine the method of protection. This is very challenging in the [Organization Name] since the mission of the organization puts the [Organization Name] in contact with high level executives to conduct its work and the data exchanged will often be confidential or restricted. And the information that is contained on [Organization Name] networks and endpoint devices can range from public to private to highly confidential. Until the network or endpoint device is examined, it will be important to treat all data bearing devices as containing the highly confidential data.

2. Purpose

The purpose of this Guideline is to establish a framework for classifying institutional data based on its level of sensitivity, value, and criticality to the [Organization Name]. Classification of data will aid in determining baseline security controls for the protection of data.

3. Scope

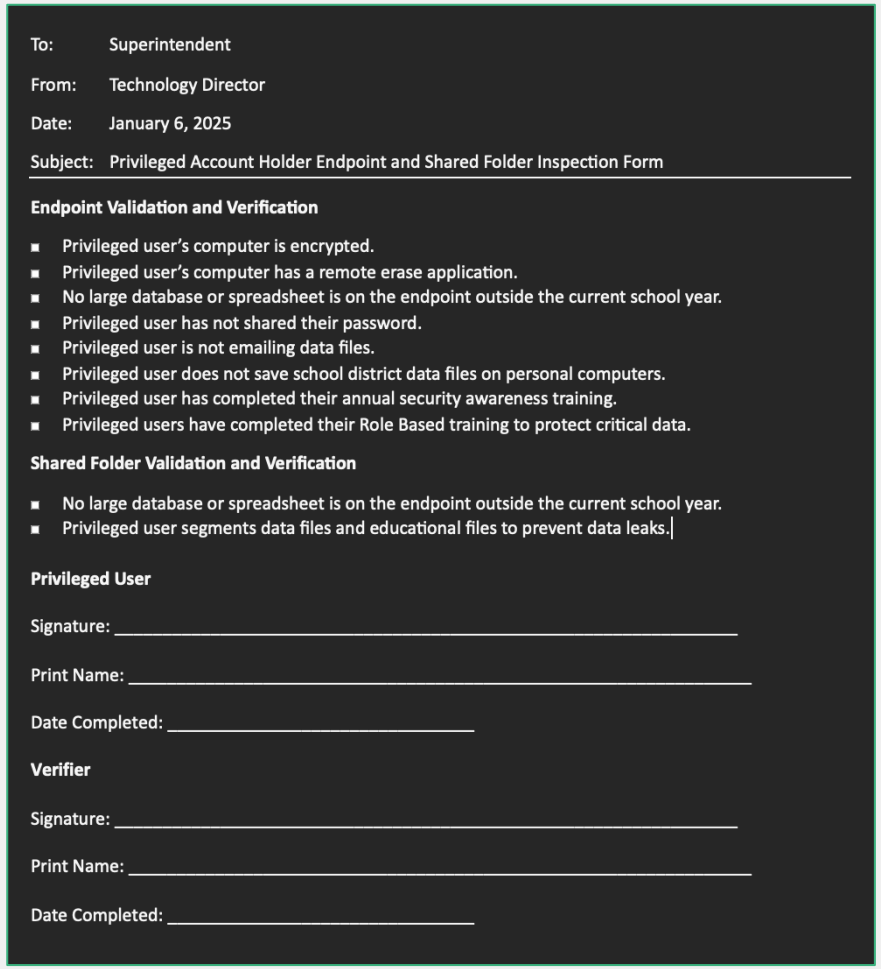
Data governance focuses on improving data quality, protecting access to data, establishing definitions, maintaining data, and documenting data policies. The [Organization Name] data is an asset and must be maintained and protected. The data must remain accurate to be trusted by the entire team to make follow on decisions at all levels of the organization.

The National Institute of Standards and Technology (NIST) defines three levels of Risk for data classification. They are as follows.

3.1 Low
The loss of confidentiality, integrity, or availability could be expected to have a limited adverse effect on organizational operations, organizational assets, or individuals.

3.2 Moderate
The loss of confidentiality, integrity, or availability could be expected to have a serious adverse effect on organizational operations, organizational assets, or individuals.

3.3 High
The loss of confidentiality, integrity, or availability could be expected to have a severe or catastrophic adverse effect on organizational operations, organizational assets, or individuals.



To: Superintendent

From: Technology Director

Date: January 6, 2025

Subject: Privileged Account Holder Endpoint and Shared Folder Inspection Form

Endpoint Validation and Verification

- Privileged user's computer is encrypted.
- Privileged user's computer has a remote erase application.
- No large database or spreadsheet is on the endpoint outside the current school year.
- Privileged user has not shared their password.
- Privileged user is not emailing data files.
- Privileged user does not save school district data files on personal computers.
- Privileged user has completed their annual security awareness training.
- Privileged users have completed their Role Based training to protect critical data.

Shared Folder Validation and Verification

- No large database or spreadsheet is on the endpoint outside the current school year.
- Privileged user segments data files and educational files to prevent data leaks.

Privileged User

Signature: _____

Print Name: _____

Date Completed: _____

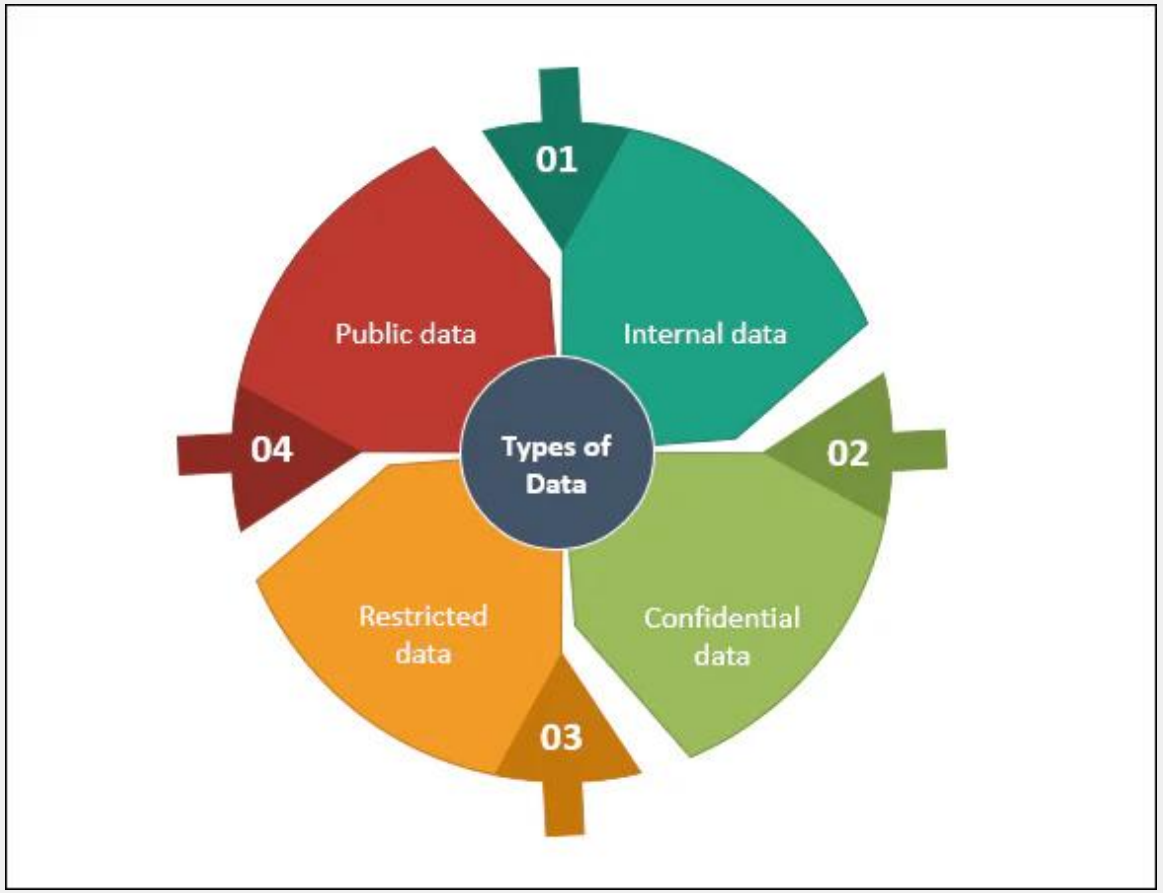
Verifier

Signature: _____

Print Name: _____

Date Completed: _____

Key Use Cases for Local Government



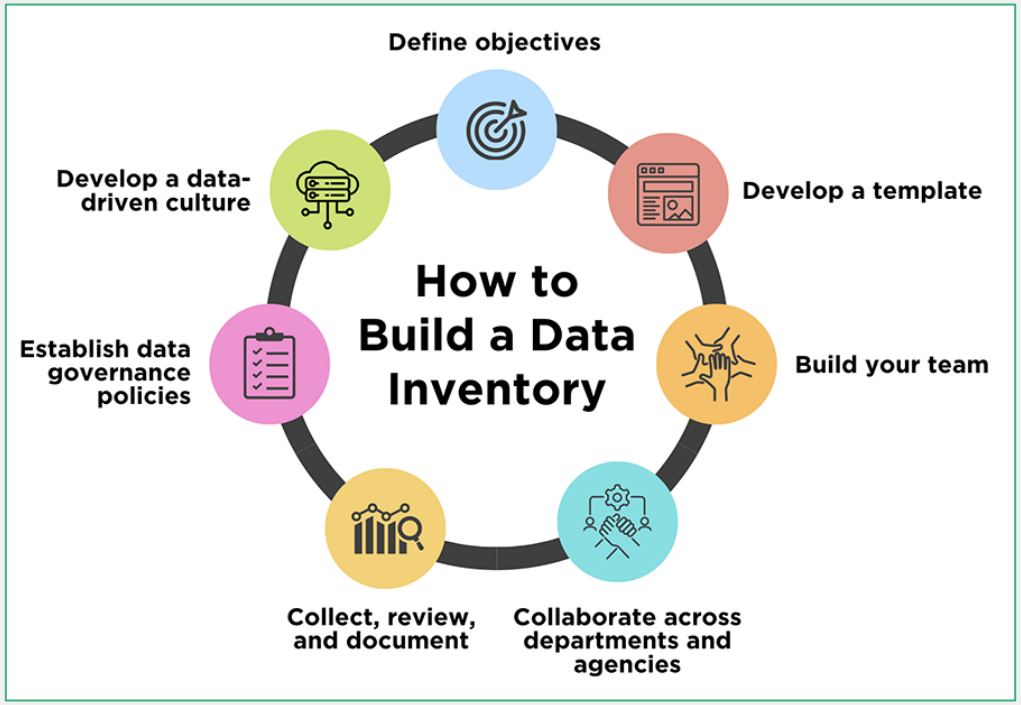
- Citizen Data
- Law Enforcement Records
- Educational Research
- Vendor Risk

Key Question: What data touches this AI, and what happens if it leaves the County?

Implementation Roadmap

- Inventory Data
- Identify AI Risks
- Deploy DLP Controls
- Measure and Adapt

- NOT Rip-and-Replace
- Many Counties have components in place!
- Don’t “Boil the Ocean” (p.13)



One of the biggest barriers counties face in adopting AI is knowing where to start. Counties don’t need to “boil the ocean.” Start small, learn, and build over time.

At the NYSAC AI Summit, county participants shared their challenges, and facilitators offered ways to identify AI opportunities that are realistic, impactful, and appropriate for local government. Below are several methods counties can use to generate AI ideas internally, along with examples drawn from the Summit.

Executive Takeaways

- Compliance <> Security
- AI Increases Risk Velocity
- Metadata is the Perimeter
- Governance is Intent, Control is Security

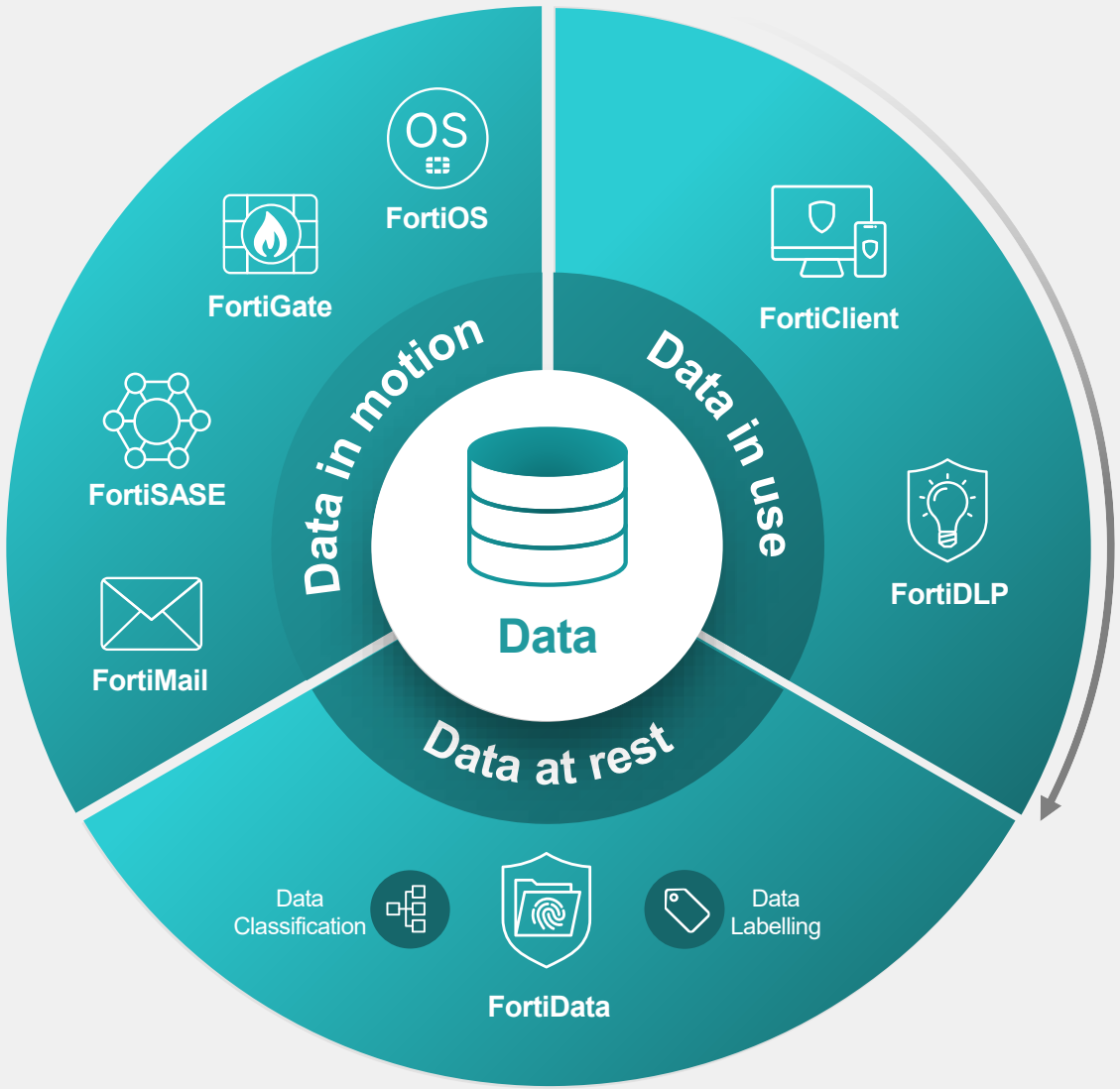
- **Being AI-Safe builds Trust**

"AI will define how counties serve the public,
But data protection will define whether the public trusts it."
- Uncle Bill

Fortinet Unified DLP

Policy Becomes Prevention Without Slowing Government

- FortiDLP - https://www.fortinet.com/products/fortidlp
- FortiAI - https://www.fortinet.com/solutions/enterprise-midsize-business/fortiai
- FortiSASE - https://www.fortinet.com/products/sase
- CASB - https://www.fortinet.com/content/dam/fortinet/assets/data-sheets/FortiCASB.pdf
- ZTNA - https://www.fortinet.com/solutions/enterprise-midsize-business/network-access



Bill Carter, CISSP
National Director of Cybersecurity Strategy, Public Sector
bcarter@fortinet.com