Beyond ROT:
Mastering Data Minimization & Least Privilege

Today's webinar presented by:

Valora Technologies

Sandy Serkes &
Jennifer Nelson

Sapient

MeR

Helpful Hints

Questions

Handouts

Feedback

Recording

Valora Technologies

Beyond ROT: Mastering Data Minimization & Least Privilege

What every RIM/IG person needs to know

Sandra Serkes
President & CEO

Jennifer Nelson
VP Strategic Solutions

Agenda

- What is Data Minimization
- Where do Data Minimization & RIM/IG intersect?
- How to stand up a data minimization program
- Using AutoClassification to enable DM

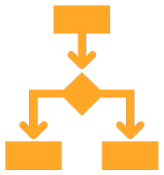
Polls

Q&A

What is Data Minimization?

Data minimization, in the domain of Information Governance, is a principle and practice that involves **collecting, processing, and retaining only the minimum amount of data necessary to achieve a specific purpose or business objective**. It is a fundamental aspect of data protection and privacy, especially in contexts where sensitive or personal information is involved.

Compliance regulations require organizations to have a lawful basis for processing personal data and **to ensure that they do not collect or retain more data than is necessary for the stated purpose**.



Data minimization is a **key component of responsible data management** and plays a crucial role in protecting individuals' privacy and **reducing the potential for data breaches and misuse**.

Key aspects of Data Minimization should include

Collecting only what's necessary

Organizations should limit the data they collect to what is strictly required for a specific, legitimate purpose.

Processing Limitations

Data should only be processed for the purposes for which it was originally collected. Additional processing beyond this scope should be avoided unless there is a valid reason for doing so.

Retention Periods

Data should only be retained for as long as it is necessary to fulfill the purpose for which it was collected. Once the data is no longer needed, it should be securely deleted or anonymized.

Minimize Data Sharing

Share data only to the extent that they need to know something to perform their role or task. Only the necessary data should be shared, and in a way that protects the privacy and security of the individuals involved.

Data Anonymization

Where possible, organizations should attempt to use techniques like data anonymization to further minimize risk associated with sensitive data. Anonymized data is less likely to be linked back to specific individuals.

What does “collect data” mean in this context?

- In a data privacy and information governance context, "collecting data" refers to the process of creating, gathering, obtaining, or acquiring information or records in any format, whether physical or digital, about individuals, entities, or events. This includes:

Direct Interaction: Voluntary submission through forms, webpages, purchases, & communications

Automated Systems: Automatic data collection from systems, software, & devices

Third-Party Sources: Data brokers, publicly databases, social media platforms, & partners.

Surveys & Research: Surveys, questionnaires, market research, & customer feedback mechanisms

Physical Documents: Paper forms, contracts, invoices, reports and forecasts, etc.

Employee Records: Hiring, termination and payroll details, performance evaluations, etc.

What is a “lawful basis” under which you can collect data (or keep, use, improve, manage, report it, etc…)

- **“Lawful basis” means a legitimate reason or legal justification** that allows an organization to collect, process, and use personal data, as established by regulations to ensure that privacy rights are respected and that personal data is handled in a transparent and appropriate manner
- Organizations are *required* to determine their appropriate lawful basis for collecting/having/processing personal data and to document this justification
- Below are GDPR’s defined lawful bases for processing personal data:

Consent

Contractual Necessity

Legal Obligation

Vital Interests

Legitimate Interests

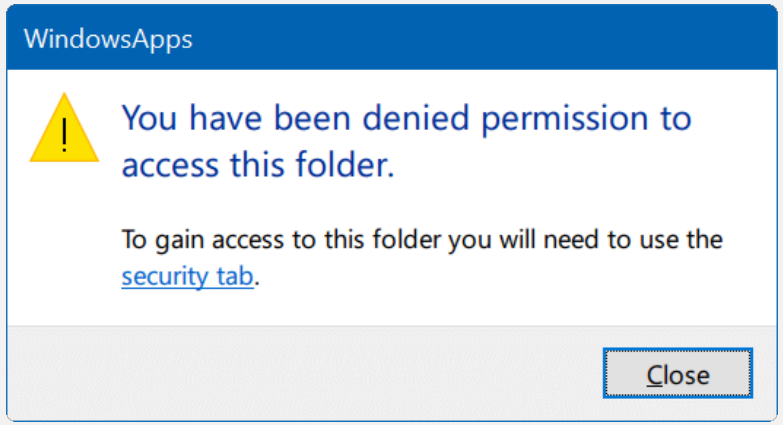
Public Task

Special Categories of Data

What is Least Privilege?

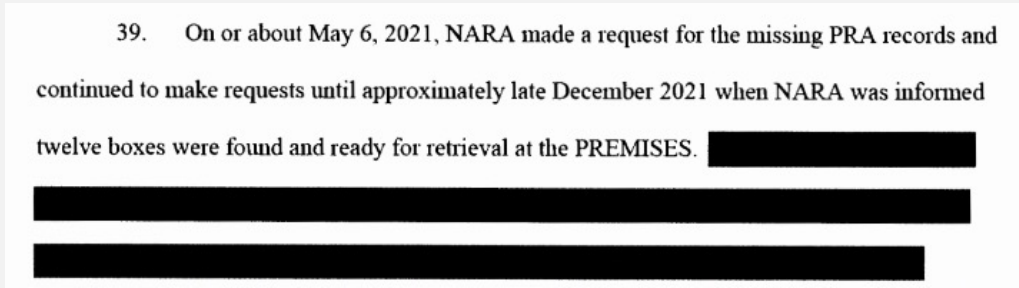
Current use of this term is typically in a *data security* context:

Least Privilege is a principle and access control strategy that restricts individuals (or systems) to only the minimum levels of access or permissions *to systems or applications* required to perform their job functions or tasks.



Emerging use of this term is in a *data governance/records* context:

Least Privilege is a principle and access control strategy that restricts individuals (or systems) to only the minimum levels of access or permissions *to data, records or content* required to perform their job functions or tasks.

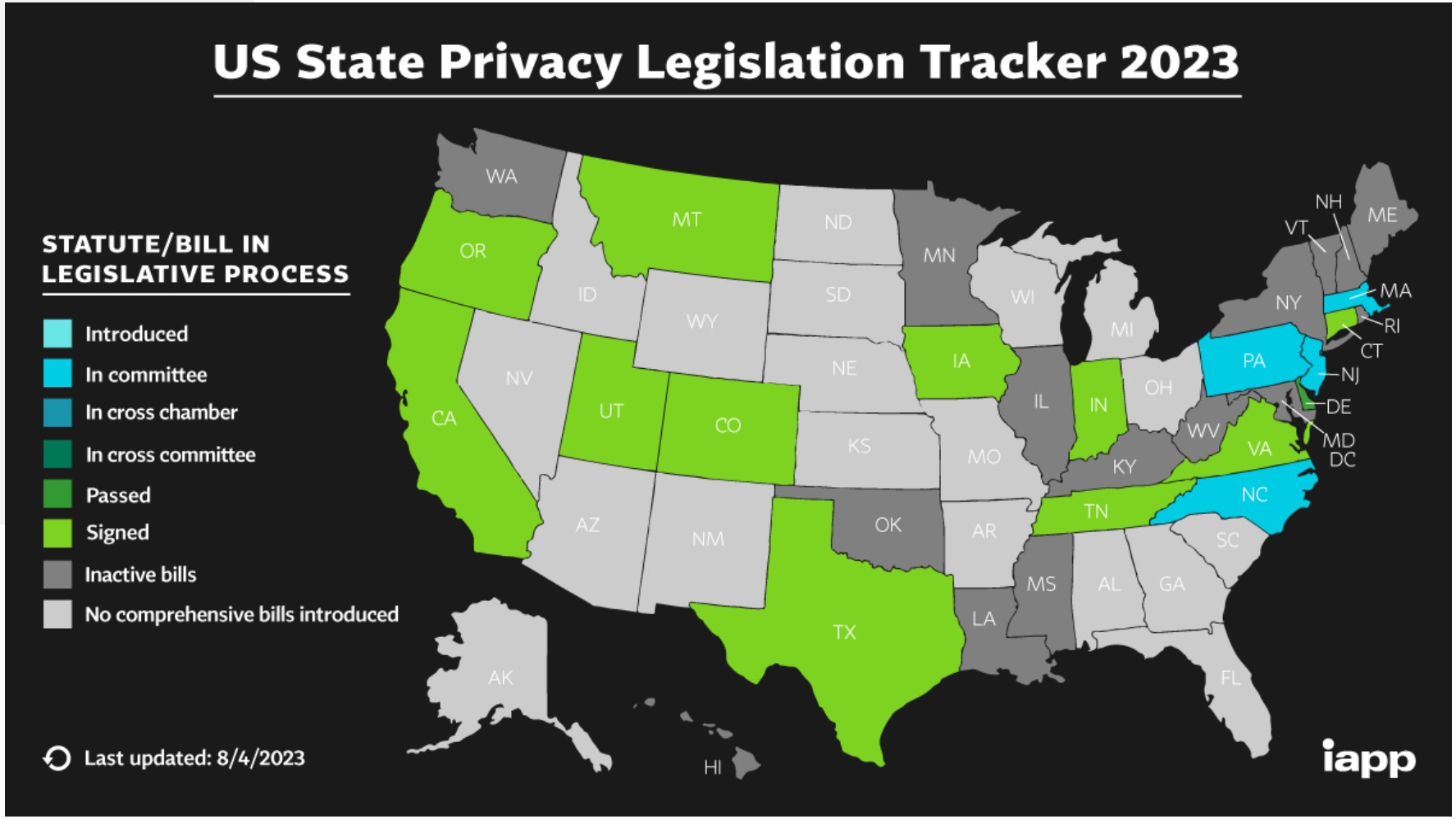

39. On or about May 6, 2021, NARA made a request for the missing PRA records and continued to make requests until approximately late December 2021 when NARA was informed twelve boxes were found and ready for retrieval at the PREMISES. [redacted]

Data Minimization requirements are standard fare in data privacy regulations

- Current Laws
  - GDPR: any personal data collected (held) must be "*adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.*"
  - CPRA: business and organizations *"shall not retain a consumer's personal information or sensitive personal information . . . for longer than is reasonably necessary"*
- Still pending
  - Data Care Act: "duty of confidentiality" and "duty of loyalty" that restrict disclosure or sale of personal data, particularly those that benefit the seller "*to the detriment of the end user.*"
  - Online Privacy Act: "*requirements for covered entities, including data minimization*"

Why is this important now?

- Get to know IAPP
- Current status of data privacy regulations in the US (Aug 2023)



Why is this important now?

- Stop thinking data privacy is “someone else’s problem”
- Just like eDiscovery, this will become a RIM/IG “problem”
- Good thing you are getting out ahead of it!

US State Privacy Legislation Tracker 2023

Comprehensive Consumer Privacy Bills

				CONSUMER RIGHTS									BUSINESS OBLIGATIONS				
STATE	LEGISLATIVE PROCESS	STATUTE/BILL (HYPERLINKS)	COMMON NAME	Right to access	Right to correct	Right to delete	Right to opt out of certain processing	Right to portability	Right to opt out of sales	Right to opt in for sensitive data processing	Right against automated decision making	Private right of action	Opt-in default (requirement age)	Notice/transparency requirement	Risk assessments	Prohibition on discrimination (exercising rights)	Purpose/processing limitation
LAWS SIGNED (TO DATE)																	
California		CCPA	California Consumer Privacy Act (2018; effective Jan. 1, 2020)	X		X		X	X			L	16	X			X
		Proposition 24	California Privacy Rights Act (2020; fully operative Jan. 1, 2023)	X	X	X	S	X	X		X	L	16	X	X	X	X
Colorado		SB 190	Colorado Privacy Act (2021; effective July 1, 2023)	X	X	X	P	X	X	X	X~		S/13	X	X	X	X
Connecticut		SB 6	Connecticut Data Privacy Act (2022; effective July 1, 2023)	X	X	X	P	X	X	X	X~		S/13	X	X	X	X
Indiana		SB 5	Indiana Consumer Data Protection Act (2023; effective Jan. 1, 2026)	X	X	X	P	X	X	X	X~		S/13	X	X	X	X
Iowa		SF 262	Iowa Consumer Data Protection Act (2023; effective Jan. 1, 2025)	X		X		X	X		X~		S/13	X		X	X
Montana		SB 384	Montana Consumer Data Privacy Act (2023, effective Oct. 1, 2024)	X	X	X	P	X	X	X	X~		S/13	X	X	X	X
Oregon		SB 619	Oregon Consumer Privacy Act (2023; effective July 1, 2024)	X	X	X	P	X	X	X	X~		S/13	X	X	X	X
Tennessee		HB 1181	Tennessee Information Protection Act (2023; effective July 1, 2025)	X	X	X	P	X	X	X	X~		S/13	X	X	X	X

Who is leading the way?

Healthcare

Restricts access to patient data, only authorized personnel can view & use

Financial Services

Limits the collection and retention of financial information

Retail & eCommerce

Streamlines customer data management and protect payment card information

Technology & Software

Implements data minimization practices in their software and platforms

Govt & Public Sector

Implements data minimization practices in their software and platforms

Energy & Utilities

Minimizes the data collected from smart meters and IoT devices

OK, ok, data privacy, I get it. Is there any other reason or driver to get on board with data minimization?

Efficiency &
Cost Reduction

Performance
Optimization

Simplified Data
Governance

Reduced Legal &
Regulatory Risk

Environmental
Sustainability

Improved
Data Quality

Enhanced
Data Security

Improved
Decision-Making

Streamlined
Data Migration

Content Lifecycle
Management

... just a few!

Organization's Data ~~Privacy~~ Responsibilities

DATA PRIVACY	RECORDS MANAGEMENT
What sensitive data do we have?	**What** information do we have?
Where is sensitive data located?	**Where** is our data located?
Who has access to sensitive data?	**Who** has access to what data?
Are we **handling** sensitive data correctly?	Are we **handling** our corporate data correctly?
What sensitive data is **accessible** that shouldn't be?	What corporate data is **accessible** that shouldn't be?
Can I **produce** the relevant information when a DSAR request is made?	Can I **produce** the relevant information when a request is made?

Which group or department is typically responsible for data minimization practices?

Like information governance, data minimization is a multidisciplinary effort, often involving:

1. Chief Information/Data Officer (CIO/CDO)
2. Data Governance Team
3. Data Privacy or Compliance Office
4. Information Security Team
5. Legal and Compliance Department
6. IT Department
7. Business Units and Data Owners
8. Data Stewards
9. Audit and Internal Controls Teams
10. Training and Awareness Teams

Describes you?
Sounds like RIM/IG?

How to stand up a data minimization program in 10 “easy” steps

You will need:

1. An understanding of the regulations and legal basis for your requirements
2. Policies documenting how you will operate and comply
3. Executive support to implement, maintain and enhance these programs
4. Appropriate tools to execute and ensure compliance
5. Ongoing diligence and management

You will do:

1. Data inventory (aka discovery, mapping)
2. Establish tagging/retention criteria
3. Baseline tagging & disposition
4. Remove low hanging fruit (ROT, dupes, past retention, etc.)
5. Implement automated retention that is data privacy-aware (ensures data minimization)

What does a typical DM program/policy look like?

- It looks like a records retention schedule!
- With some key additions:
  - ✓ Least privilege – who can see/do what with this file/record
  - ✓ Intended lifecycle vs. record lifecycle
  - ✓ What time/event triggers force the disposition forward?
  - ✓ Data privacy citations
  - ✓ Sensitivity
- This means single-tab record-keeping is not enough
  - Record Class is insufficient (and non-compliant!)
- New retention/disposition labels required to retain only as long "as needed" (per the stated intention at collection/creation time)

Things can get complicated quickly:

We keep client transaction data for 3 years past the transaction date

- unless the client requests their "right to be forgotten," in which case it is 45 days from DSAR date
- unless it is under legal hold.

Watch out!
Suddenly “past retention” takes on new, onerous meaning

- It’s no longer ok to keep data “past retention” because it is in direct violation of data minimization that is expressed as “no **longer** than….”
- Vs records retention, which is implicitly expressed as “no **shorter**/less than…”
- New lifecycle management tags, expression, and management policy as a retention “window:”
  - Minimum retention period (RIM) – no less than
  - Maximum retention period (DM) – no more than
  - Subject to current Legal Hold

Remember:

If you retain it,

it is discoverable!

Historically, why were you unable to delete ROT?

- “No tools.” (Delete, delete, delete…)
- “Can’t find it, ROT is mixed in with the good stuff.”
- No one will let us delete things.

No real penalty to “just keep forever” (and violate our RRS, but ok)

Now, there are **real** consequences to not deleting data that is supposed to be deleted (past retention), and that we said we would delete/retain for only X purposes or time period.

We are now **out of compliance** (not just sloppy).

Meh, so we're out of compliance what's the big deal?

- GDPR fines for data privacy compliance violations
  - Meta - €1.2 billion ($1.3 billion)
  - Amazon - €746 million ($781 million)
  - Instagram - €405 million ($427 million)
  - Facebook - €265 million ($275 million)
  - WhatsApp - €225 million ($247 million)
  - Google LLC - €90 million ($99 million)
  - H&M - €35 million ($41 million)
  - British Airways - €22 million ($26 million)
  - Marriott International - €20.4 million ($23.8 million)

How can tools like AutoClassification help establish, manage and maintain the principles of data minimization?

AutoClassification tools can play a significant role in establishing, managing, and maintaining the principles of data minimization within an organization. These tools use automated processes, machine learning, and predefined rules to categorize and classify data based on its content, context, and relevance.

Identification of Sensitive Data

Data Classification

Data Retention Policies

Access Control & Permissions

Data Minimization in Backup & Archives

Data Discovery & Mapping

Compliance with Privacy Regulations

Data Incident Response

Policy Enforcement

Regular Auditing & Monitoring

How long does it take to set up or configure an AutoClassification tool to perform this work?

Complexity

Of your data environment: cloud, on-prem, structured, unstructured data environments

Goals & Requirements

Input from stakeholders for enterprise-wide approach to data governance.

Amount of Data

Smaller organizations with less data will be faster to process than large enterprise.

Weeks

Months

Phased Implementation

Planning

Prioritize data environments, align stakeholders, set KPIs

Proof of Concept

Limited data set to benchmark and forecast scope/scale

Implementation

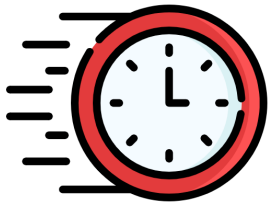
Implement scale to process all enterprise content

Delta Monitoring

Continuously look for and process new & edited content

Weeks

Months



coming soon!

New eBook:

Data Privacy, Data Minimization & Least Privilege

What every IM/IG professional needs to know

Important Info about Data Minimization

Important Terms

- Need to know
- Lawful basis
- Collecting data
- Least privilege
- Data subject

Where to Get More Info

- www.valoratech.com/resources
- Iapp.org
- European Data Protection Supervisor
- Electronic Privacy Information Center

Valora Technologies

www.valoratech.com

Book a Demo

Webinar Series

/valora-technologies

Q&A

Thank you

Sandy Serkes
President & CEO
sserkes@valoratech.com

Jennifer Nelson
VP Strategic Solutions
jnelson@valoratech.com

Sapient Upcoming Webinars

DATE	TITLE	PRESENTED BY
September 26 In Partnership with OpenText	**Strengthen Content Protection with Zero Trust Information Governance**	Mike Safar, OpenText Tracy Caughell, OpenText Greg Clark, OpenText
October 12 In Partnership with Epiq	**How Microsoft Purview Brings Information Governance, Security, and Legal Together**	Anne Costello, Epiq Brandon Hollinder, Epiq
October 24 In Partnership with EncompaaS	More information to come	
November 4 In Partnership with Valora	**Managing Records Retention with AutoClassification**	Sandy Serkes, Valora Jennifer Nelson, Valora