# **EDM** Webinar















A conversation with



**Peggy Tsai** CDO **BigID** 



Jennifer Mezzio **Managing Director** Silicon Valley Bank

















Moderator: Mike Meriton
Co-Founder & COO
EDM Council



- Joined EDM Council full-time 2015 to lead Industry Engagement
- EDM Council Co-Founder & First Chairman (2005-2007)
- EDM Council Finance Board Chair (2007-2015)
- Former CEO GoldenSource (2002-2015)
- Former Executive for D&B Software and Oracle
- FinTech Innovation Lab Executive Mentor (2011 – Present)





Peggy Tsai CDO BigID



- 18+ years of experience in data management, stewardship & governance at S&P Global, AIG, Morgan Stanley
- Adjunct Faculty at Carnegie Mellon's CDataO Program
- EDM Council Women Data
   Professional Forum's Global Co-Lead for Events
- CDMC Working Group member and committee member





Jennifer Mezzio

Managing Director, HR Divisional Data Officer

Silicon Valley Bank



- 25 years in IT strategy, business / systems / data analysis, data warehousing & data governance
- Current projects include building & executing and transforming HR into a data-driven organization
- Scope now includes Legal and Cybersecurity
- Masters in Organizational Leadership
   & Computer Science from Quinnipiac
   University
- Editorial Board Member for CDO Magazine

### Today's speakers



Moderator



Mike Meriton
Co-Founder & COO
EDM Council



Peggy Tsai CDO BigID



Jennifer Mezzio

Managing Director,

HR Divisional Data Officer

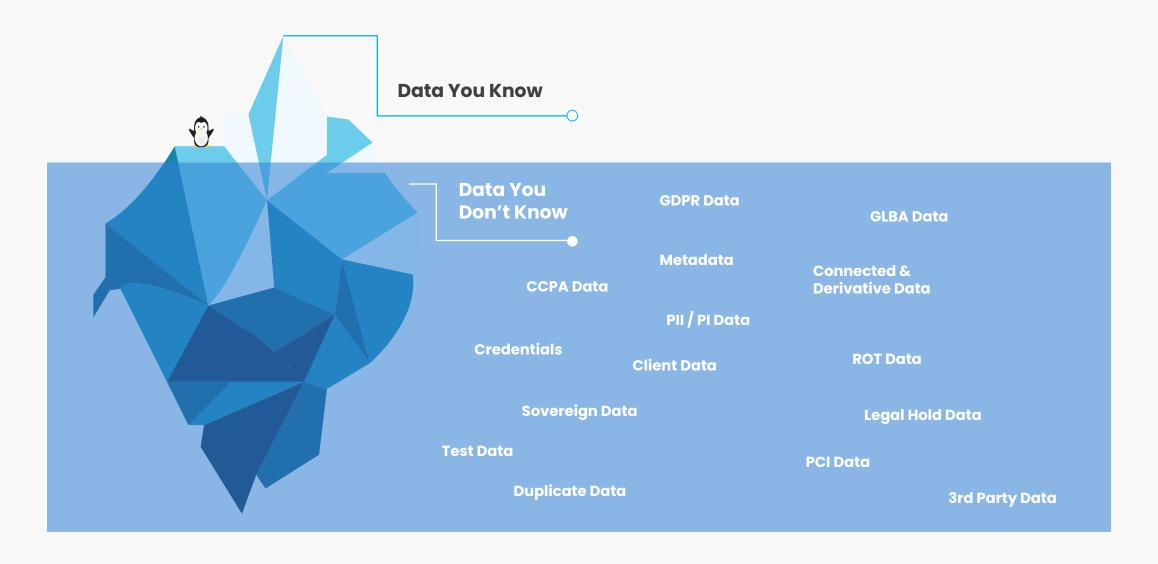
Silicon Valley Bank







## New Regulations Driving New Challenges for Managing Sensitive Data





### Sensitive Data Management Challenges in the Cloud and On-Prem

More Data Volume

### **More Places to Collect and Process Data:**

AWS S3, Redshift, DynamoDB, EMR, RDS, Aurora, Kinesis, Databricks, Snowflake, Azure Data Lake, 0365, ServiceNow, Salesforce, Confluent, Jira, GitHub, GitLab, Cohesity, Confluence, Athena

### More Speed of Data Change:

Data volumes exploding exponentially as it becomes easier to add, copy, move in cloud with ephemeral data sets, platforms like snowflake and databricks, email, files

More Data Velocity

More Kinds of Regulated, Sensitive, Critical Data To Track GLBA, HIPAA, PCI, SOX, GDPR, CCPA, PII, PI, PCI, Credentials, CID, Patents, Research, ML Models, Data Combos

More Data Variety

More Data
Vulnerability

How do you measure & remediate risk: open access, excessive privileges, duplicate data, sovereign data, misconfigurations, ROT data, credential leaks, embedded code





## The innovation economy's financial partner

### **Facts at a Glance**

Since 1983, we have fueled the innovation economy by enabling our clients to grow and thrive through the power of our integrated financial services.

## Nearly half

US venturebacked technology and life science companies bank with SVB<sup>1</sup> 44%

of U.S. venture capital-backed technology and healthcare IPOs YTD bank with SVB<sup>2</sup> 2022

Forbes' Best Banks in America<sup>3</sup>

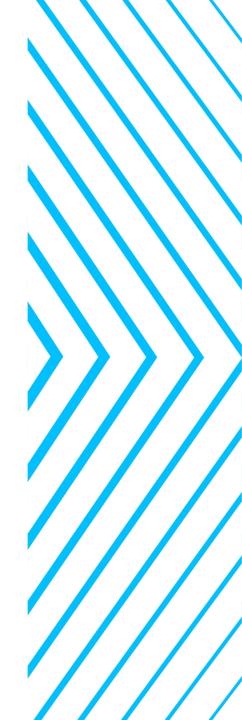












### **Problem Statements**



- Due to the inexpensive cost of storage, companies have gone from "data collectors" to "data hoarders"
- Data is in various systems both in and outside of HR
- Difficult to easily identify data outside of the retention policy i.e. – out of compliance
- Retention policy has multiple retention limits for one record making purge extremely difficult
- Purging is a highly manual and tedious process
- Purging for 1 record has to happen across all levels physical, source system, structured and unstructured repositories, digital etc.





# What is your organization's biggest challenge with sensitive data management? Select one.

- 1) Difficult to identify data outside of retention policy
- 2) Inexpensive storage costs encourage 'hoarding' of data
- 3) Manual and tedious process to identify data to be in compliance
- 4) Lack of prioritization across teams to manage sensitive data
- 5) Other (specify in chat)

## **HR Initiatives and Challenges**



- Purging of data is highly manual due to the complexity of scenarios, multiple targets, structured and unstructured repositories, and lack of automation
- Employee data is more than just the worker record, it's spouse/dependent data, investigations, training, surveys, workforce planning, etc.
- Risk and exposure of sensitive data is reduced by having the IT and data verticals reside within the HR Organization - what about outside HR?
- Need a secure way for Human Resources data outside of HR to be consumed by teams across the Enterprise
- Need to document data lineage all the way to final rest including 3<sup>rd</sup> party vendor applications



### **Intended Outcomes**

- Robust inventory of data elements which includes classification and source to target lineage - currently manually updated
- Tag and catalog the data elements across various repositories
- Easily identify the data elements and records that need to be retained or purged in accordance to the retention policy
- Create an automated retention and purge process that is accurate and operationalized across all of HR and SVB
- Align purge and retention efforts across all unstructured and structured repositories throughout the Enterprise









# What is your organization's technology-readiness level to help manage sensitive data? Select one.

- 1) Innovators We successfully implemented solutions and processes
- 2) Early Adopters We implemented solution(s) and developing the policies and processes for our teams
- 3) Early Maturity We started to explore solutions and will do so within the next 12 months
- 4) Late Maturity We do not plan to adopt until 2024 or later
- 5) Laggards It is not in our roadmap

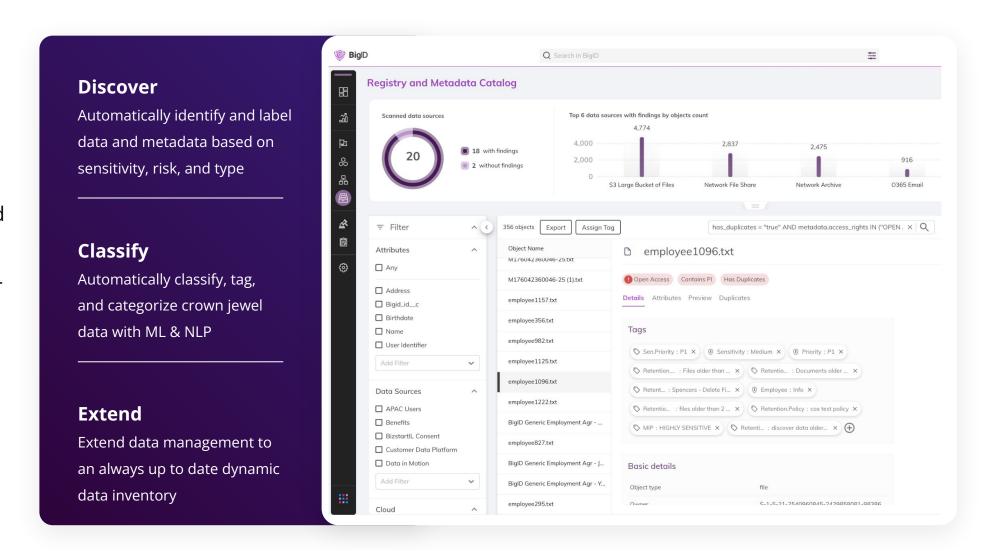


# Additional Industry Use Cases for Discovery, Classification & Data Mapping

### Industry Use Case: Al-Discovery and Know Your Data



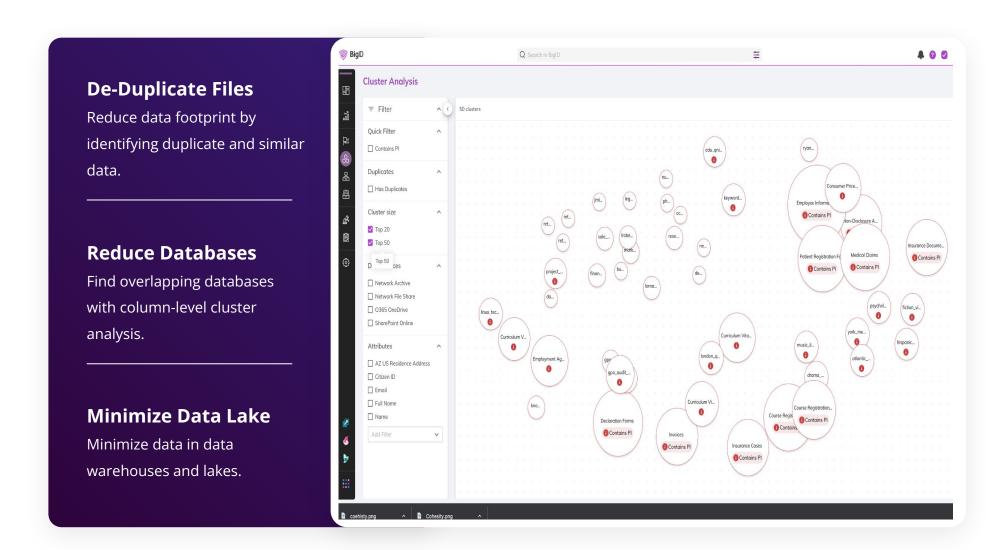
Automatically classify and tag data for data management across your entire data landscape.



## Industry Use Case: Data Minimization



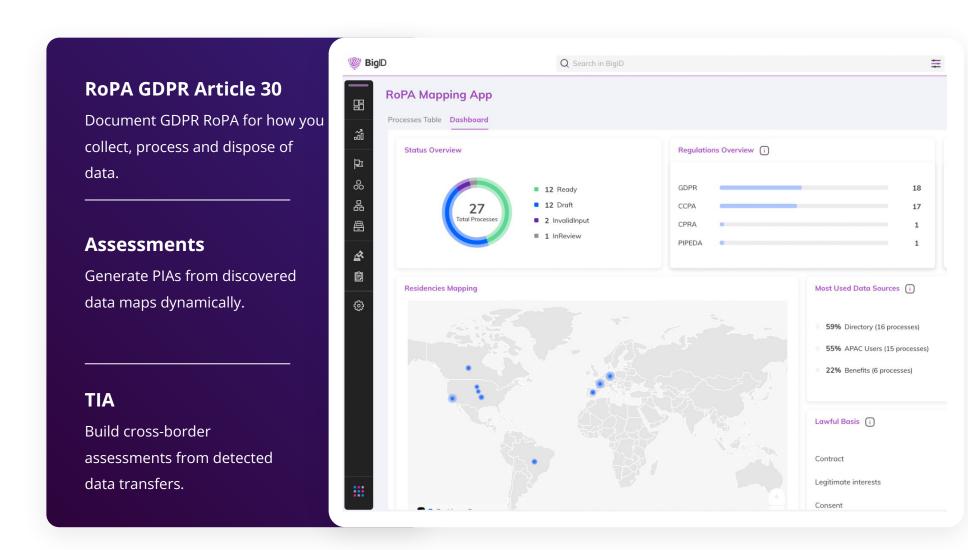
Enable data and database rationalization and hygiene with deduplication and similarity analysis.



## **Industry Use Case: Automate Privacy Reporting**



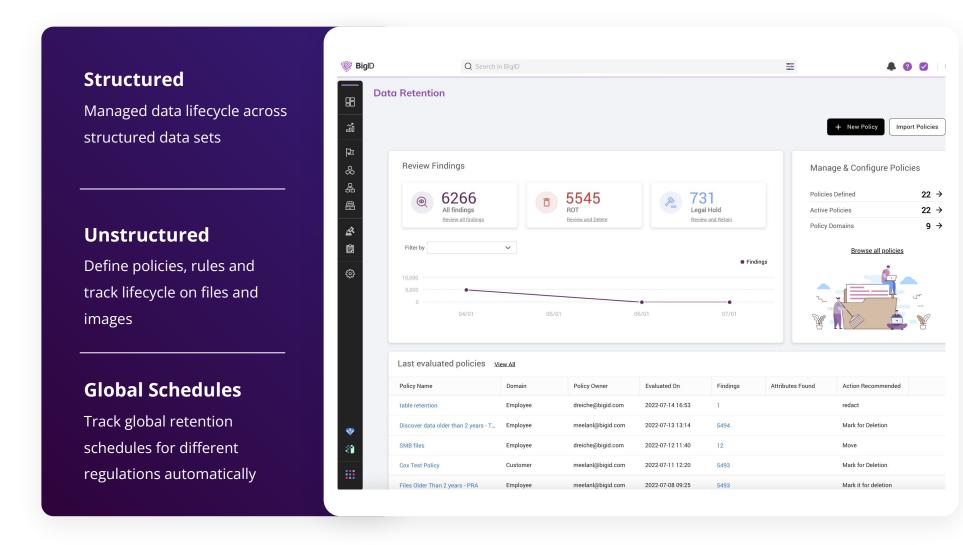
Automate preferences, data rights, assessments, RoPA and more from dynamic data map.



### **Industry Use Case: Streamline Retention**



Consistently manage data aging, reduce redundant data, improve system hygiene, and manage data retention based on age, policy, and more.





## **Questions?**



Break down the data silos with

### One Platform.

Improve visibility and control with a single interface - and modular apps within the platform. Drive better decisions and outcomes with greater context, insight, and accuracy of your data.



Discover dark data and find all sensitive data, improve security posture, reduce risk



Accelerate cloud migrations, data validation for M&A, and avoid audits



Fill gaps in the security stack with a data-centric approach to get more value from your existing tech stack



Achieve compliance, comply with data privacy and protection regulations, and avoid audits



Enable security frameworks: NIST, HITRUST, CIS CSC, Zero Trust



Make better decisions and accelerate security initiatives with more accuracy and confidence in understanding enterprise data



## Open Extensible Data Visibility & Control Platform

One Platform Deployable as SaaS, Edge-to-Cloud, or Self-managed

**BigID** Œ Low Quality Data

**Control Your Data** 

**Know Your Data** 

**Connect Your Data** 

BigID App Marketplace for privacy, security, and governance controls

4-in-1 ML Augmented Data Discovery for deeper data visibility, insight, and inventory

Agentless scalable scanning
AWS• Azure• GCP•
SaaS• laaS• Code Repos•
Pipeline• NoSQL• Email• Apps•
Messaging• Mainframe



### **Know Your Data**

Automatically scan, find, and classify the data that matters most - wherever it lives (and whatever it is)





### **Enrich your Investments**

Extend and enrich your existing investments with an open ecosystem designed for actionable context and integrations across the board



PaaS. SaaS. Hybrid. Everywhere in-between.

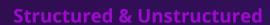






### Scale

Modern architecture designed for deployment where you want it, at petabyte scale.



Not all data looks the same - your data management needs to be customized to your challenges & environment.

Make technology work for you modern ML methodology to reduce time, improve accuracy, and accelerate value



## Thank you!



Additional Questions? info@bigid.com

www.edmcouncil.org







