DATA DYNAMICS

Use Case

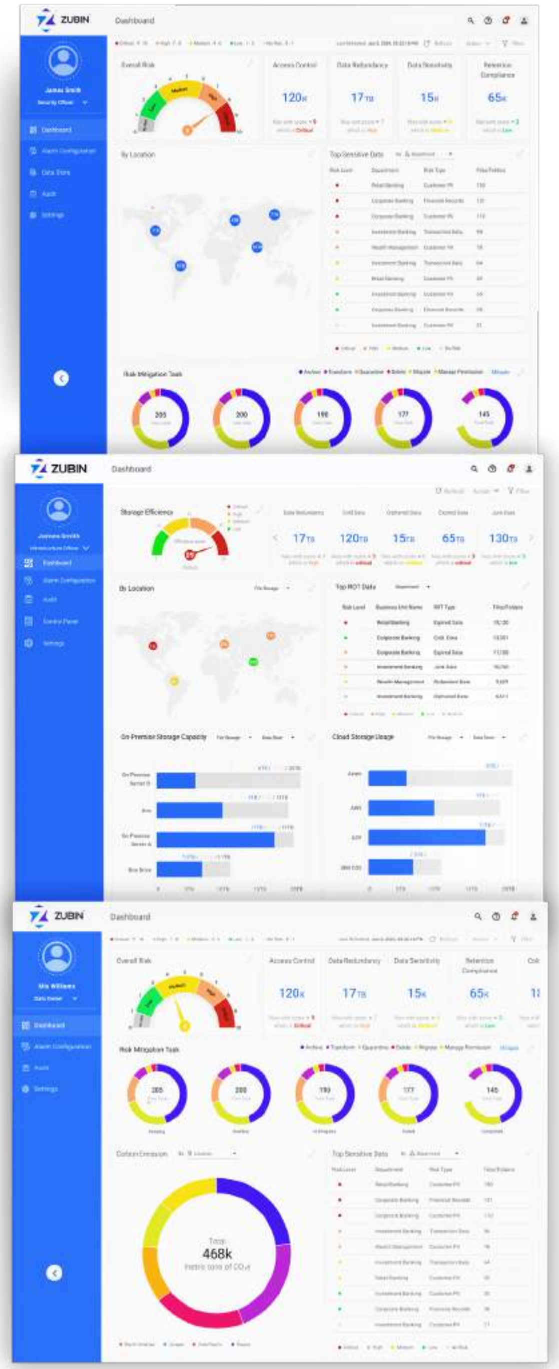
Privacy Management With Intelligent Data Classification for AI Models

In today's age of AI, organizations leverage vast amounts of data to fuel these powerful models. But a hidden challenge lurks: a significant portion of this data remains unstructured and unseen, creating a blind spot for privacy. This "dark data" can harbor sensitive information – a goldmine for those looking to exploit it.

This is where Zubin can help.

Zubin is Data Dynamics' AI-powered self-service data management software, bringing a fresh approach to privacy, security, compliance, governance and optimization in the world of AI-led workloads. It empowers enterprises by enabling users across all levels - from C-suite to data owners - to discover, define, act, transform, and audit data through a user-friendly interface. Zubin brings correlation, consistency, and standardization across your organization by delivering granular insights, deriving recommended workflows, and automating actions using personalized policies and RBAC-driven processes. This transformation fosters a culture of data ownership, where everyone becomes a data champion, and the organization fulfills its responsibility as a data custodian.

This solution brief explores how Zubin empowers you to navigate the tightrope walk of AI and data privacy. We'll delve into 6 crucial use cases that unlock the power of your data, from identifying and mitigating privacy risks to fostering trust with users and ensuring responsible AI development.



Comprehensive Data Discovery and Classification

Unidentified PII (Personally Identifiable Information) and PHI (Protected Health Information) within training data can lead to biased AI models perpetuating discrimination.

Minimizing AI Bias & Ensuring Data Fairness with Zubin

Zubin's Metadata Analytics goes beyond the surface, analyzing file formats, locations, and access control lists to pinpoint potential treasure troves of PII or PHI. It doesn't stop there. Content Analytics, powered by the muscle of NLP and advanced pattern recognition, dives right in, scanning the data for fingerprints of sensitive information like names, addresses, social security numbers, or medical records. This two-headed attack ensures you uncover every hidden piece of sensitive data, regardless of whether it's neatly structured or scattered across unstructured formats.

A large bank wants to develop an AI model to automate loan approval decisions. However, they can discover through Zubin's Metadata Analytics that customer loan applications are often stored in different formats (emails, PDFs, spreadsheets) making it difficult to identify and remove PII (Social Security numbers, income details) before feeding the data into the AI model. Content Analytics powered by NLP helps classify this sensitive data, allowing the bank to ensure fair and unbiased loan approval decisions without compromising customer privacy.

Multi-Layered Risk Assessment & Prioritization

Balancing robust data security with the need for AI models to access relevant data for training and analysis can be challenging.

Balancing Privacy with AI Innovation with Zubin

Building on the foundation of Zubin's Content Analytics, Risk Exposure Insights takes data classification a step further. It assigns risk scores based on the type of sensitive data identified, such as PII, PHI, or even business-sensitive information. This nuanced approach considers not just the data itself, but also its intended use. Furthermore, Statistical Sampling capabilities empower you to analyze representative subsets of data. This efficient approach provides valuable insights into your overall risk profile without overloading your processing resources.

An insurance company wants to develop an AI model to predict the risk of car accidents. This model requires access to a vast amount of data, including driving records and vehicle information. However, they need to ensure they are not accessing any personally identifiable information (PII) like driver names or addresses. Zubin can analyze the data and assigns a higher risk score to PII data compared to driving records or vehicle type. This allows the insurance company to prioritize data access requests, ensuring they leverage the most relevant data for AI development while minimizing privacy risks.

Click here for a demo

Data Lineage and Audit Trail for Regulatory Compliance

Regulations like GDPR and CCPA require organizations to demonstrate responsible data handling practices for AI models.

Demonstrating Responsible AI Practices with Zubin

Zubin's Data Usage & Traceability keeps a watchful eye on your data throughout the AI lifecycle. It meticulously records which data was accessed for AI training, by whom, and for what specific purpose. This comprehensive audit trail empowers organizations to confidently demonstrate compliance with data privacy regulations like GDPR and CCPA. These regulations mandate organizations to maintain a detailed record of all processing activities, and Zubin's solution provides the clear and concise documentation you need.

A pharmaceutical company develops an AI model to analyze patient data and identify potential drug candidates. Regulations require them to demonstrate they are not using patient data for AI development. Zubin can provide a clear data audit trail, showing regulators that all patient data was quarantined with appropriate access management and handled according to compliance standards.

Data-Centric Security with RBAC Down to the Data Owner Layer

Insufficient access controls and data sharing practices can lead to unauthorized access and potential misuse of data for AI training.

Minimizing Insider Threats & Data Sharing Control with Zubin

Zubin's RBAC (Role-Based Access Control) extends down to the data owner layer, working seamlessly with data classification to empower owners with granular control. This means you can define access not just for users, but also for specific AI models. This ensures that only authorized AI models can access specific data elements, minimizing the risk of unauthorized access or misuse of your data. Want to collaborate on AI projects? Zubin integrates seamlessly with Data Sharing capabilities. Define secure data sharing protocols for internal teams or external partners involved in your AI projects, ensuring sensitive data remains protected throughout the collaboration process.

A hospital wants to analyze medical images for early cancer detection. However, they need to ensure only authorized personnel (radiologists, data scientists) have access to sensitive patient medical images used for training the AI model. Zubin's RBAC Down to the Data Owner Layer allows radiologists (data owners) to define granular access controls, restricting access to only authorized personnel working on the AI project.

Click here for a demo

Data Wrangling and Curation for Responsible AI

Poor data quality and inherent biases within training data can lead to inaccurate and unfair AI models.

Ensuring Data Quality & Mitigating Algorithmic Bias with Zubin

Zubin empowers data scientists and data owners to work together on cleaning and transforming data for AI projects. This intuitive suite includes data discovery, classification, anomaly detection, and even helps identify potential biases within your data. Furthermore, Zubin's Content Analytics can pinpoint sensitive data points that might skew your AI model outputs. By ensuring high-quality, unbiased data, you can mitigate algorithmic bias and develop more trustworthy and reliable AI models.

An insurance company develops an AI model to assess the risk of homeowners' insurance claims. However, they discover through data analysis that the model is unfairly penalizing homeowners in certain zip codes. Zubin’s Data Wrangling and Curation tools allow data scientists to investigate the data and identify anomalies such as ROT in claims data recorded for different geographic areas. By correcting these inconsistencies and ensuring data quality, they can mitigate algorithmic bias and develop a more fair and accurate AI model.

Data Security Orchestration & AI Model Explainability

Maintaining ongoing data security for AI models and ensuring their explainability to stakeholders can be challenging.

Continuous Monitoring & Building Trust with Zubin

Zubin doesn't stop at finding your data – it empowers you to secure it too. Zubin's Data Security Orchestration, Automation, and Actionability provides continuous monitoring of data access patterns, keeping you vigilant against potential security threats related to AI models. Leveraging machine learning, Zubin's Data Observability and Root Cause Analysis detects anomalies in data usage within your AI pipelines. This proactive approach helps identify and address potential issues before they impact model performance.

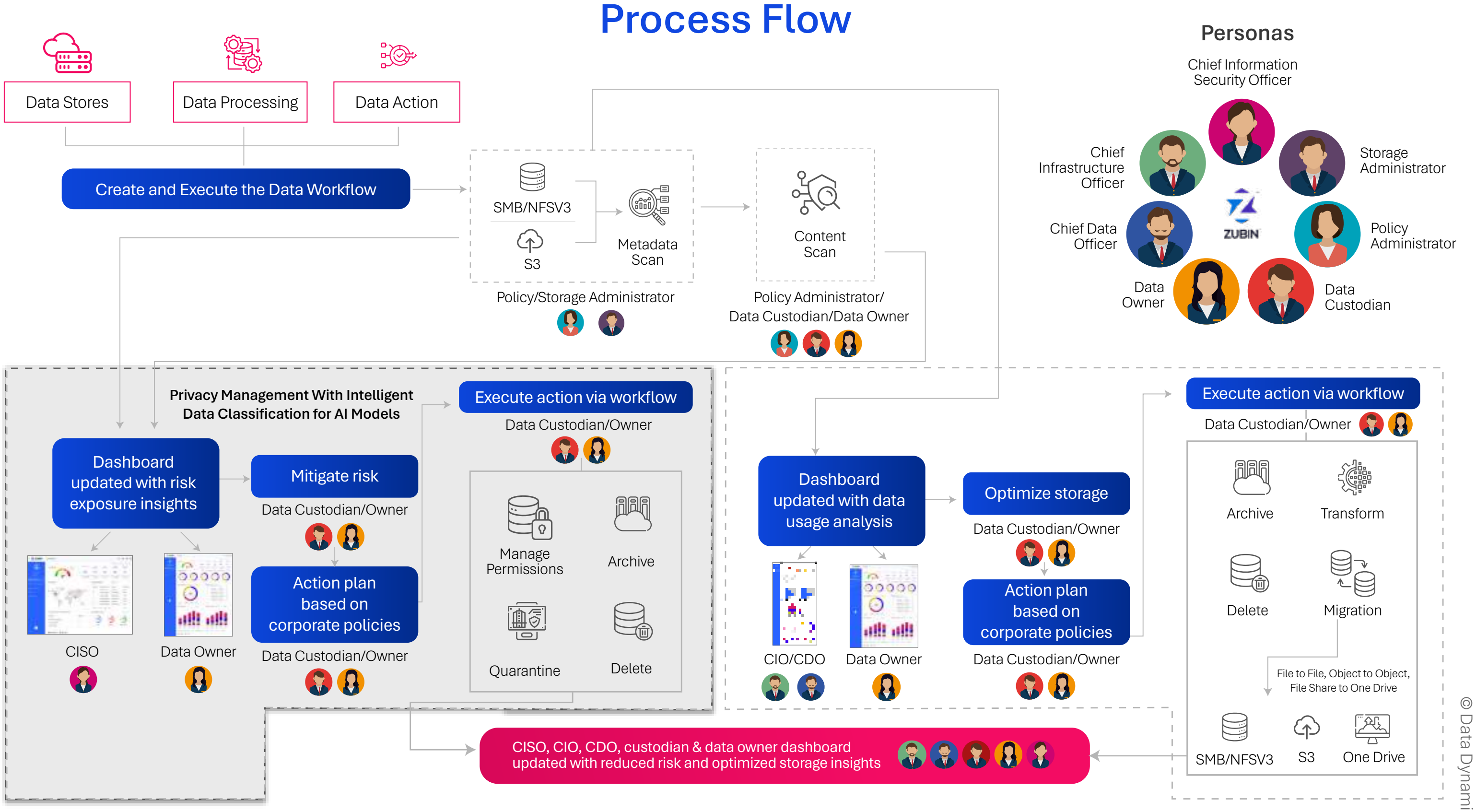
An energy company uses an AI model to predict energy consumption patterns and optimize grid operations. However, due to a lack of proper data monitoring, the model ingests faulty sensor data with inconsistencies. Zubin's Data Security Orchestration, Automation, and Actionability features can continuously monitor data access patterns and identify anomalies in sensor data. This allows data scientists to detect and rectify issues before they impact the AI model's performance. This fosters trust and transparency with stakeholders like regulators and the public, who rely on the accuracy of these AI-driven energy predictions.

Click here for a demo

A Closer Look at What Makes Zubin Stand Out

- Enterprise Risk Posture & Infra Usage BI & Reporting
- Data Risk & Usage by Team & Data Ownership
- Personalized Data Policy Creation & Workflow
- Metadata Analytics
- Content Analytics
- Data Classification
- Statistical Sampling
- Data Migration
- Data Minimization
- Data Archival
- Retention Compliance
- Data Transformation
- Data Tiering & Placement
- Cloud Integrated Risk Controls
- Access Control & File Re-permissioning
- Data Sharing
- Risk Exposure Insights
- Data Containment & Isolation
- Real-time Notifications and Reporting

Process Flow



Data Stores
Data Processing
Data Action
Create and Execute the Data Workflow
SMB/NFSV3
S3
Metadata Scan
Content Scan
Policy/Storage Administrator
Policy Administrator/ Data Custodian/Data Owner
Personas
Chief Information Security Officer
Chief Infrastructure Officer
Storage Administrator
Chief Data Officer
ZUBIN
Policy Administrator
Data Owner
Data Custodian
Privacy Management With Intelligent Data Classification for AI Models
Dashboard updated with risk exposure insights
Mitigate risk
Data Custodian/Owner
Action plan based on corporate policies
Data Custodian/Owner
CISO
Data Owner
Execute action via workflow
Data Custodian/Owner
Manage Permissions
Archive
Quarantine
Delete
Dashboard updated with data usage analysis
CIO/CDO
Data Owner
Optimize storage
Data Custodian/Owner
Action plan based on corporate policies
Data Custodian/Owner
Execute action via workflow
Data Custodian/Owner
Archive
Transform
Delete
Migration
File to File, Object to Object, File Share to One Drive
SMB/NFSV3
S3
One Drive
CISO, CIO, CDO, custodian & data owner dashboard updated with reduced risk and optimized storage insights

Your next chapter of success awaits; let's write it together with Zubin. Click here for a demo