

AI-Powered Information Governance



June 2019

IDENTIFY



90%

of data in existence today was produced in the last two years¹

70-80%

of the data created in the last two years is unstructured, unidentified, uncontained, dark data¹

CLASSIFY



83%

of organizations store sensitive data in the public cloud¹

1 in 4

Organizations have experienced data theft from the public cloud (found for both SaaS and IaaS)¹

OPTIMIZE



97%

of organizations use cloud services (public, private, or hybrid)¹

66%

of IT professionals say security is their greatest concern when adopting a cloud computing strategy²

Solution Overview

MinerEye interpretive AI™ is a fusion of Byte level pattern recognition and unsupervised clustering of patterns. MinerEye interpretive AI™ uses computer vision to read bytes of an exemplar file and create a signal – a mathematical vector.

The exemplar & its signal is the rule

The image collage illustrates the MinerEye AI system's capabilities. It features several examples of 'Learned' and 'Matched' patterns:

- Learned:** A 3D model of a mechanical part.
- Matched:** A 3D model of a car body.
- Learned:** A blue 3D model of a car body.
- Matched:** A scanned document titled 'Request for Proposal' from 'RUE SCATIONS'.
- Learned:** A scanned document titled 'Request for Proposal' from 'symtrex inc.'.
- Matched:** A scanned check from 'Grange National Bank'.
- Learned:** Another scanned check from 'Grange National Bank'.

See Beyond Data



Security

- Identify
- Analyze
- Protect

Compliance

- Classify
- Define
- Enforce

Tactical Value & Impact

- Execute Mission
- Enhance Oversight

Strategic Value & Impact

- Governance
- Cybersecurity

Identification

Locate data to determine infrastructure and enhance data governance

Data Minimization

Remove redundant data in multiple locations to reduce storage and attack surface

Automated Classification

Set sensitivity labels, automatically cluster and track dispersed data

Forensic Analysis

Data spillages, FOIA requests, accidental or malicious release of data

Data Optimization

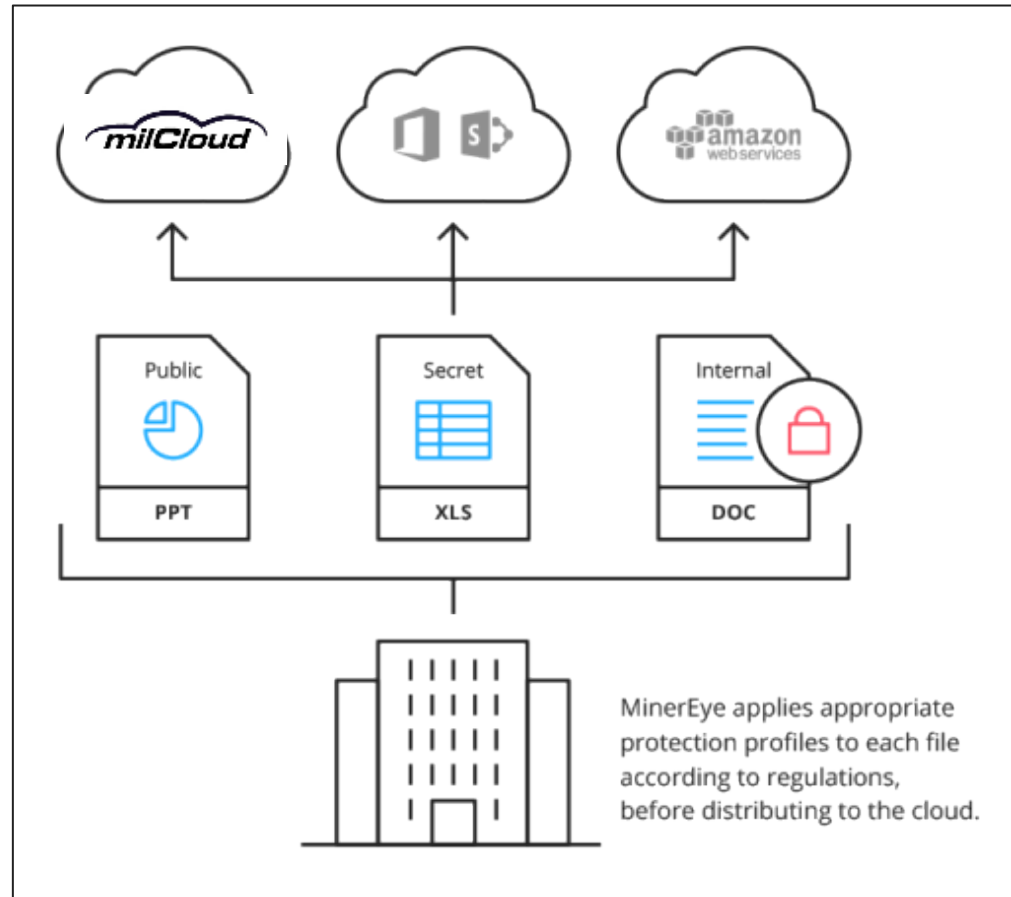
Minimize data, define criticality levels, provide visibility on-prem and cloud

Preparing for a Secure Cloud Migration

MinerEye stands as a mediator, ensuring the privacy and protection of data that travels between the premise and the cloud, as enabled by hybrid cloud models

Features enable customers to clearly define data impact and criticality categories, segregate one data group from another, establish and apply appropriate protection profiles

This allows for an in depth data analysis that provides deep visibility into infrastructure



Potential Savings



The Association of Information and Image Management (AIIM) reports that on average, half of an organization’s retained information has no business value.

The Compliance, Governance, and Oversight Council (CGOC) estimates that a company with 10 petabytes of data could be spending as much as \$34.5 million on data that could be deleted.

Site	Storage	Estimated Potential Cost Savings	Site	Storage	Estimated Potential Cost Savings
Army Customer	46.3 TB storage	Estimated \$160k in redundant data	Pentagon	22TB/day 8.03PB/yr	\$28M in redundant data costs
DoD Customer	288 TB storage	\$994k	Potential Total		~\$28M redundant data
USAF Customer	11 TB storage	\$345k			
Potential Total		~\$1.5M redundant data	<p>“The Pentagon collects 22 terabytes of data every day...You cannot exploit 22 terabytes worth of data the way we are doing things today.” -Lt. Gen Jack Shanahan, Director, Joint Artificial Intelligence Center, Office of the Department of Defense Chief Information Officer, the Pentagon.</p>		

United States Financial Services Group

- Problem – Unknown company proprietary data usage and unidentified ROT data
- MinerEye Solution - located 5 million obsolete documents that customer reduced
- MinerEye Solution - Customer is now utilizing data forensics to analyze dark data worldwide

Worldwide Integrated Circuit Manufacturer

- Problem - Internal Intellectual Property leakage and GDPR & PII compliance
- MinerEye Solution – Identify IP existing outside approved facilities
- MinerEye Solution – GDPR and PII identified and subsequently reduced or protected

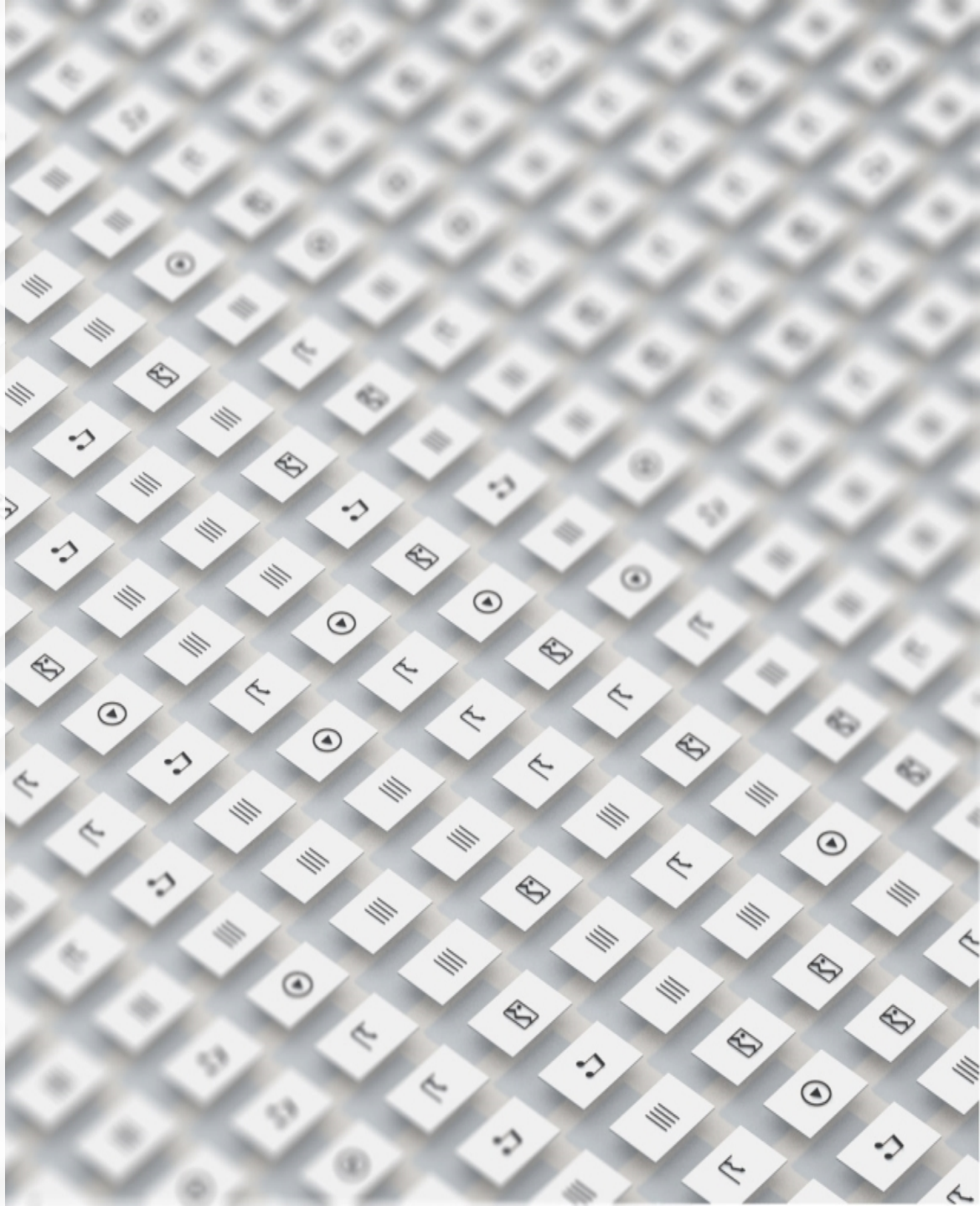
United Kingdom Bank

- Problem – Unknown extent of sensitive customer data in business data sets
- MinerEye Solution - found thousands of rogue files that contain sensitive customer data in the form of Excel and Access DB (.mdb) that were subsequently classified and cleaned

- Enhanced Cybersecurity
- Operational Efficiency
- Regulatory Compliance
- Risk Management
- <https://agile-defense.com/services/minereye>



Supporting Documents



MinerEye Data Tracker™ GUI



Dashboard Vgadmin Vgadmin

15M (10TB)
Total Files

10
Scanned Data Sources

7
Geo-Locations

7000
Data Exceptions

Recommendations

Data Minimization	100 Recommendations	zoom
Data Collaboration	1000 Recommendations	zoom
Data Segregation	100 Recommendations	zoom
RTBF	9000000 Recommendations	zoom
HR-CVs	0 Recommendations	zoom

Classification Progress

● Classification Progress

Data Minimization Status

Full Duplications Near Duplications

● Duplications ● Near Duplications

Top Categories by file size

● Data Minimization ● Data Segregation
● Data Collaboration ● RTBF ● HR-CVs

Files Info

Types Sizes Entities Last accessed

● Types

Geo-Distribution

Source
Category
Entities

MinerEye Data Tracker™ GUI



☰
🔔

Query Builder

+ Rule
+ Query

AND

👤 Credit Card Starts with 4580

OR

📄 Extensions Equals PDF

OR

🔄 In Group Red Label Green Label +

OR

📄 Similar to HR_agrment.xls 🔍 ⬆️ Duplication

Recommendations

Data Colaboration ▾

- Data Colaboration Set Label Red 10K files zoom
- Data Colaboration Set Label Red 10K files zoom
- Data Colaboration Set Label Red 10K files zoom
- Data Colaboration Set Label Red 10K files zoom

1 - 4 of 100

Visualization
Data

Locations
Entities
Sources
Categories
Types

Distribution by usage

State
Size
Age

- 76% more than 6 years old
- 60% Less than year
- 78% 3-5 years old

Classification Distribution

Risk Level
Categories

- Data Minimization
- Data Segmentation
- Data Segregation

Duplicates Distribution

Size
Source

- Duplicates
- Near Duplicates

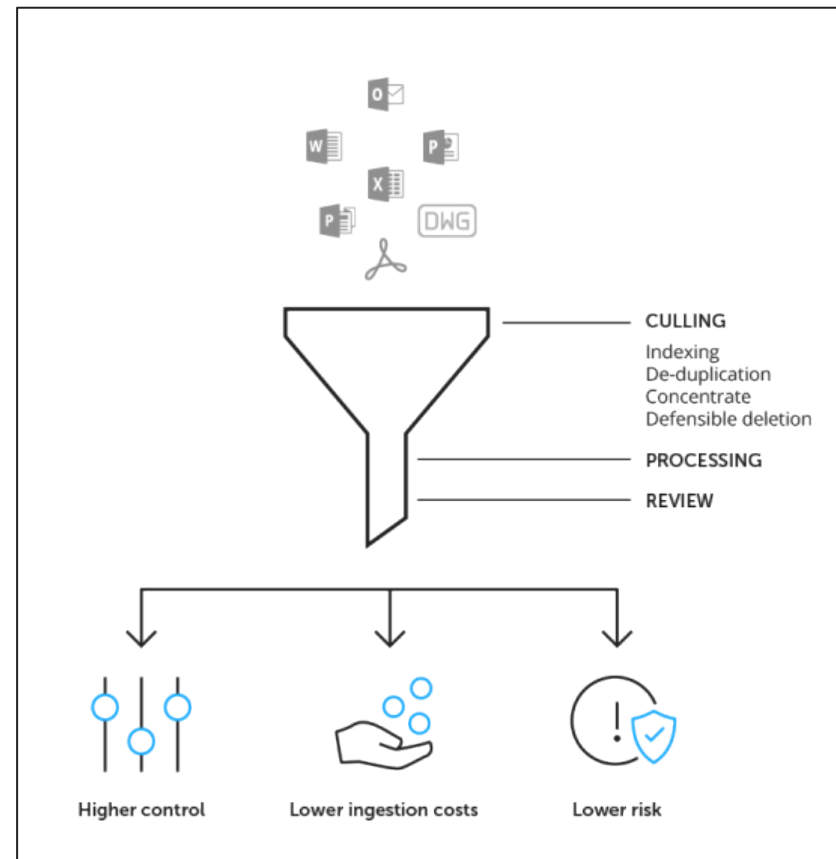
Zoom Results

SMB2 > Share > Scan_Files > Wikileaks ←

625	139
408	127
	51
	3

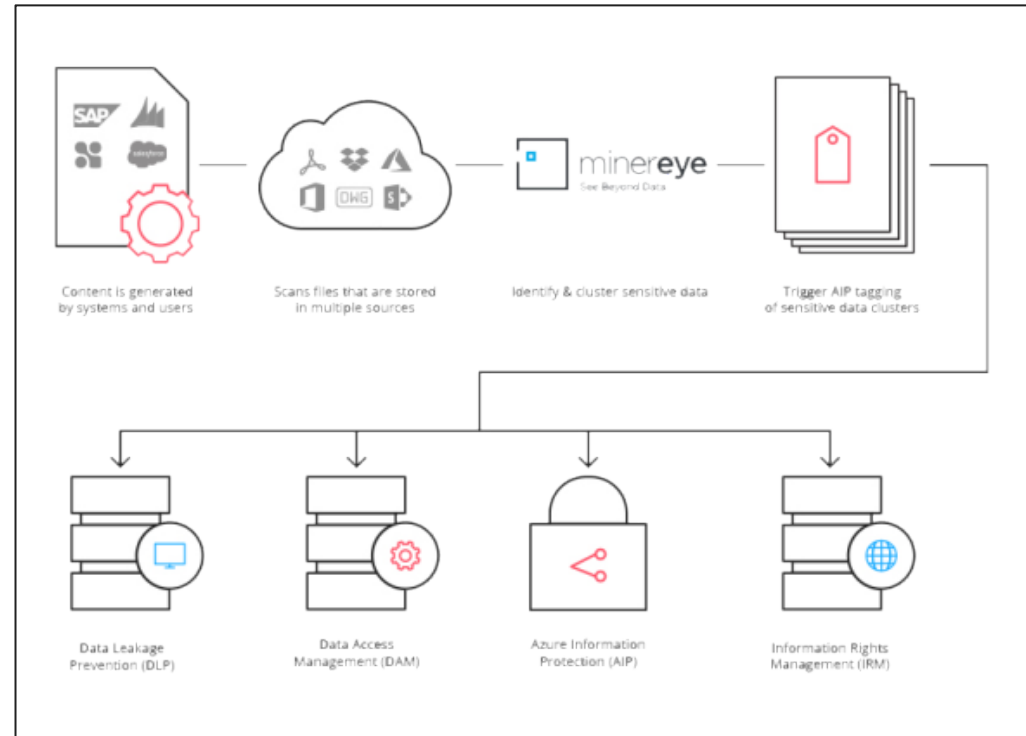
Use Case: Data Minimization

- MinerEye helps companies minimize obsolete and redundant data that has accumulated over time and takes up precious data storage.
- According to the Governance and Oversight Council (CGOC), a company with 10 petabytes of data may reduce storage costs by approximately \$34 million by properly disposing of its redundant data.



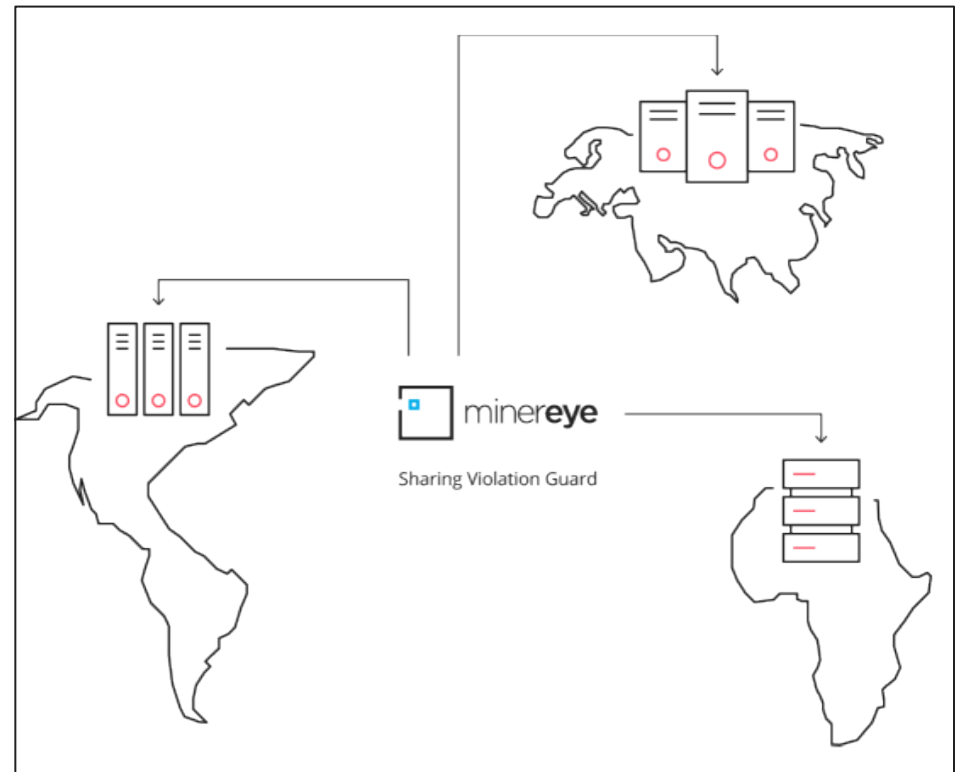
Use Case: Automated Data Classification

- MinerEye tailors automated classification capabilities to your current data protection systems, ensuring consistent and comprehensive classification of all essential data.
- MinerEye's integration with data protection software tools, such as Azure Information Protection (AIP), propagates AIP labels to unlabeled data, and assures comprehensive and efficient data classification and protection.



Use Case: Data Segregation

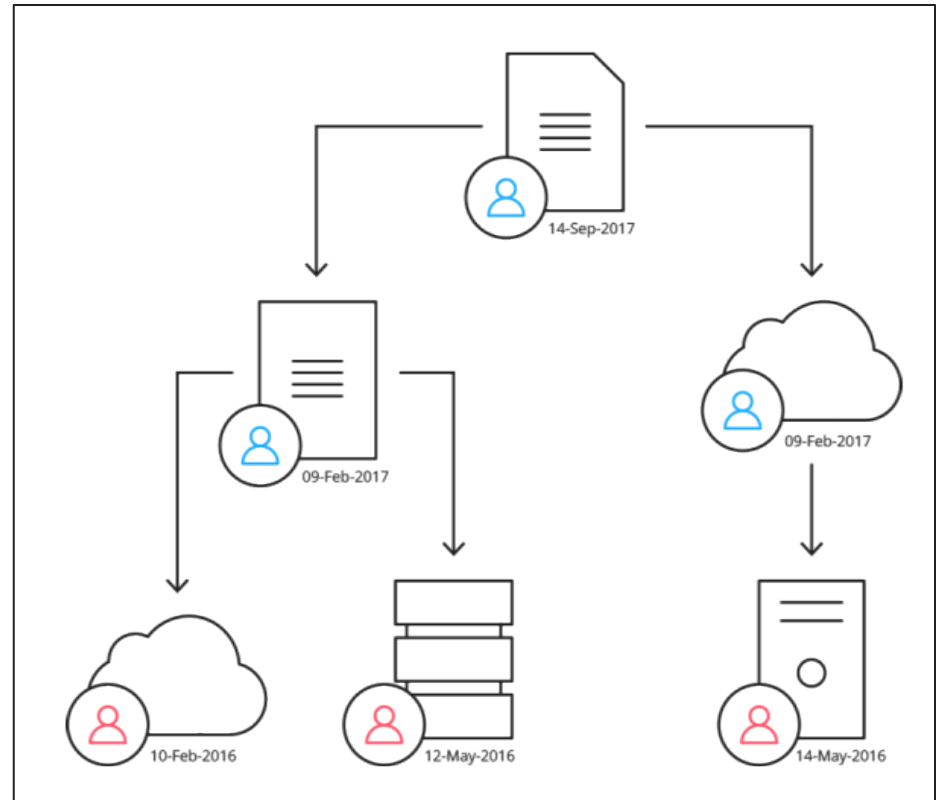
- MinerEye's Data Tracker™ continuously maps and alerts on data segregation issues and data sharing violations that are prevalent among large dataset carriers.
- MinerEye distributed architecture enables companies deploy a net of virtual machines (VM) across company sites, visualizing Data Residency and achieving sovereignty regulation compliance

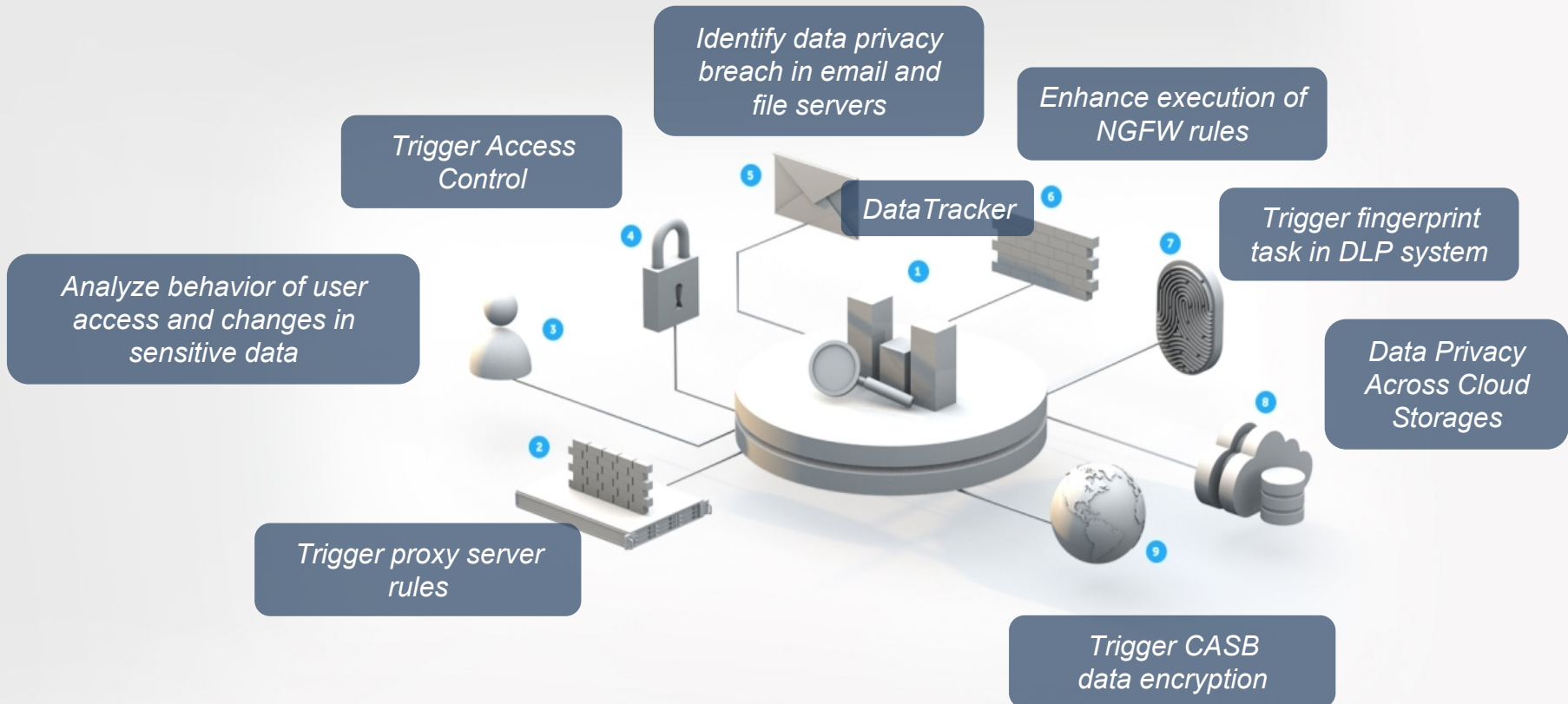


Use Case: Forensic Data Analysis

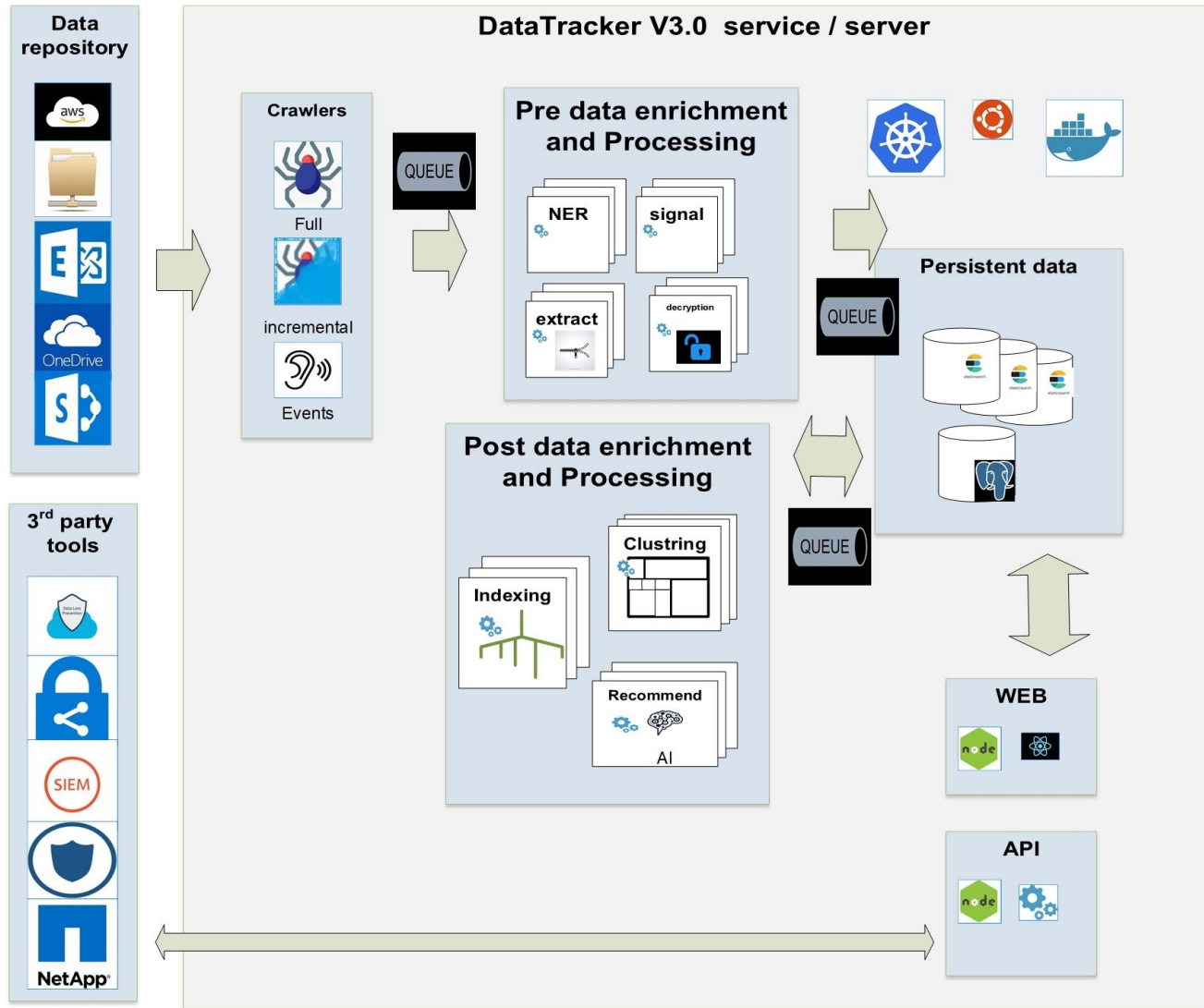


- MinerEye logs data behavior by scheduled scanning of your sensitive information clusters.
- MinerEye can instantly identify and present where your data is or was located at different times, whether it has been modified, who handled it and more.
- MinerEye automates alternative months of manual work to investigate the movement and tampering of data enabling companies to cut time to know and report on data breach according to regulations





Reference Architecture



Deployment Open Virtual Machine

One VM supports up to 120TB source data



Optimal OVA Compute Resources	
CPU	16GHz / 8 Core
RAM	64GB
Disk	2TB

