IBM Content 2014
Information. Insights. Results.

IBM

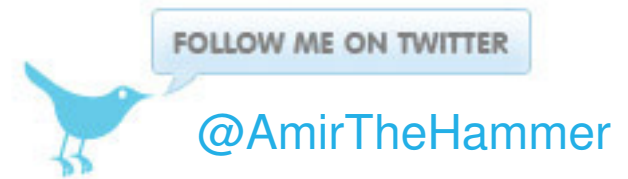Amir Jaibaji – ILG Product Management, Program Director
September 2014

Stop Data Hoarding
Cleaning up your legacy data

FOLLOW ME ON TWITTER

@AmirTheHammer

Enterprise big data dilemma

" Data capacity on average in enterprises is **growing at 40 percent to 60 percent year over year** due to a number of factors, including an **explosion in unstructured data**, such as email and documents that have to be stored due to 'regulatory requirements that continue to evolve and change. "

SOURCE: Computerworld, "Data growth remains IT's biggest challenge, Gartner says," Lucas Mearian, November 2, 2010.

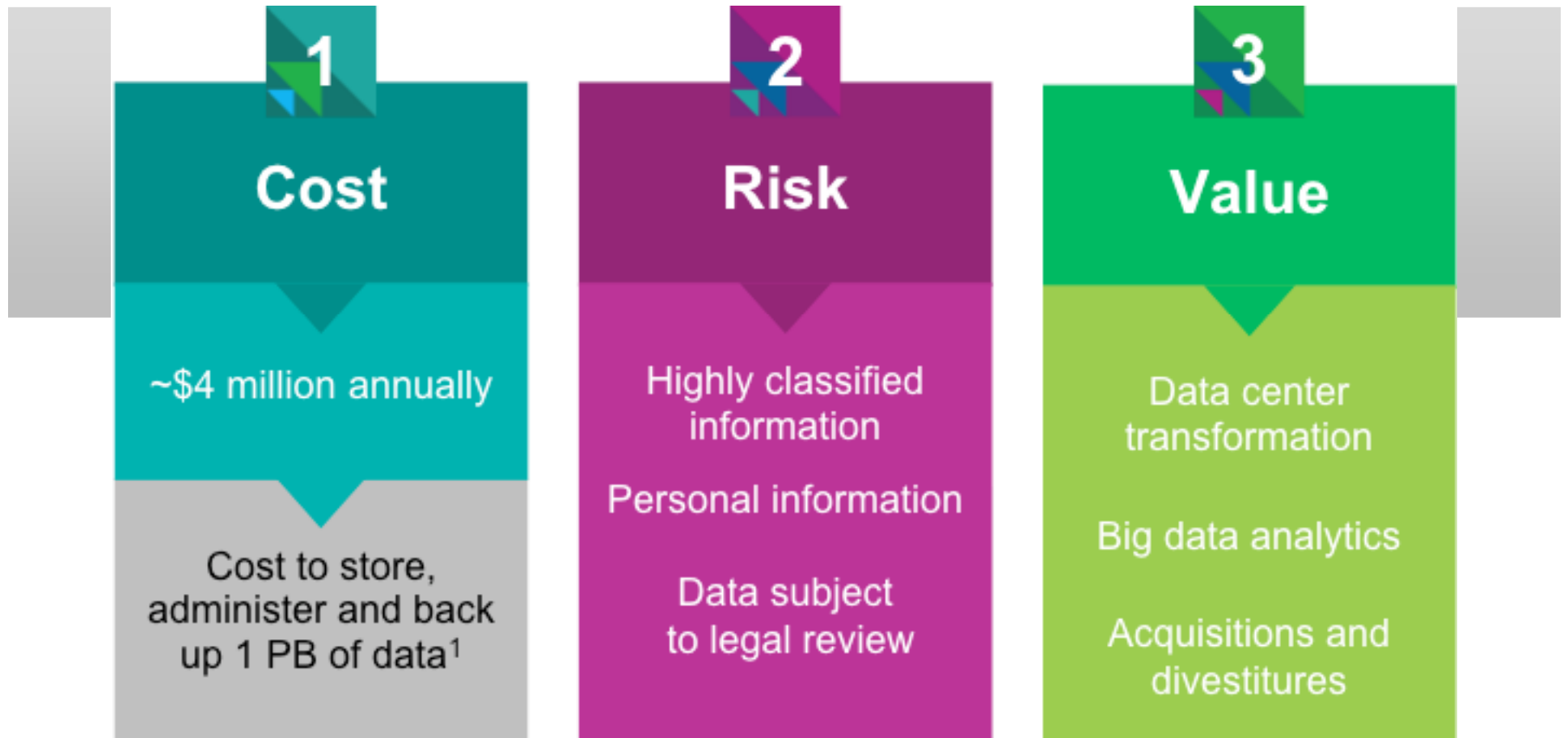# "Storage is cheap" is no longer the answer

- *Growth is out of control*
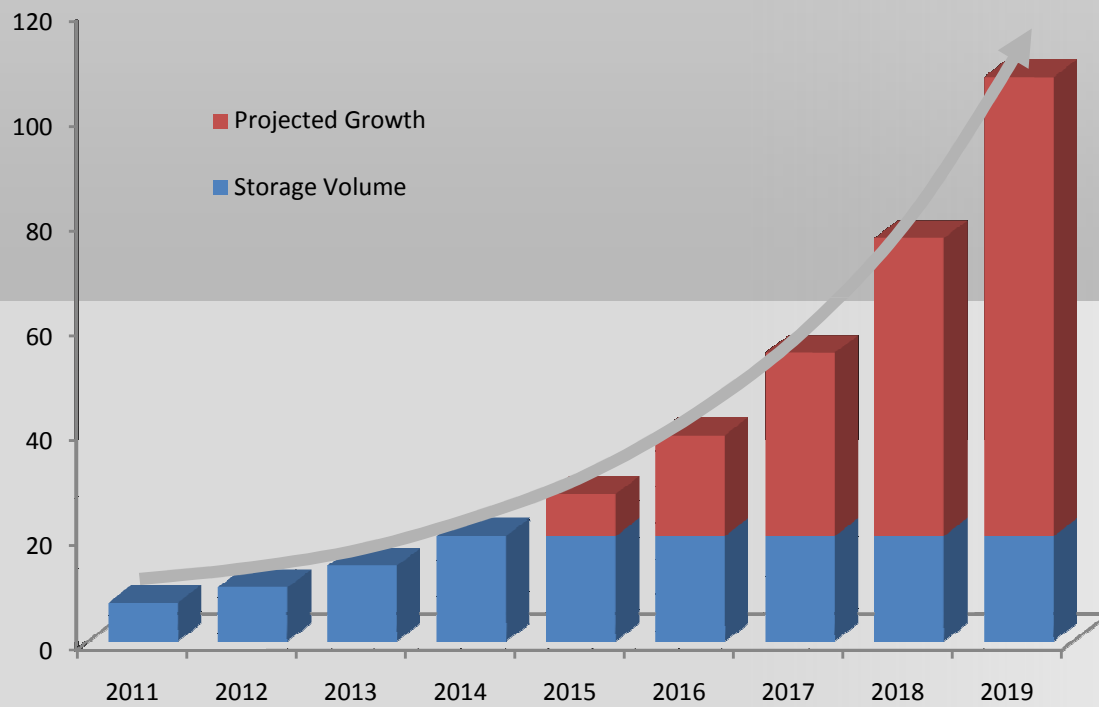
40%-60% Annual Growth

Storage Costs are Consuming IT Budget

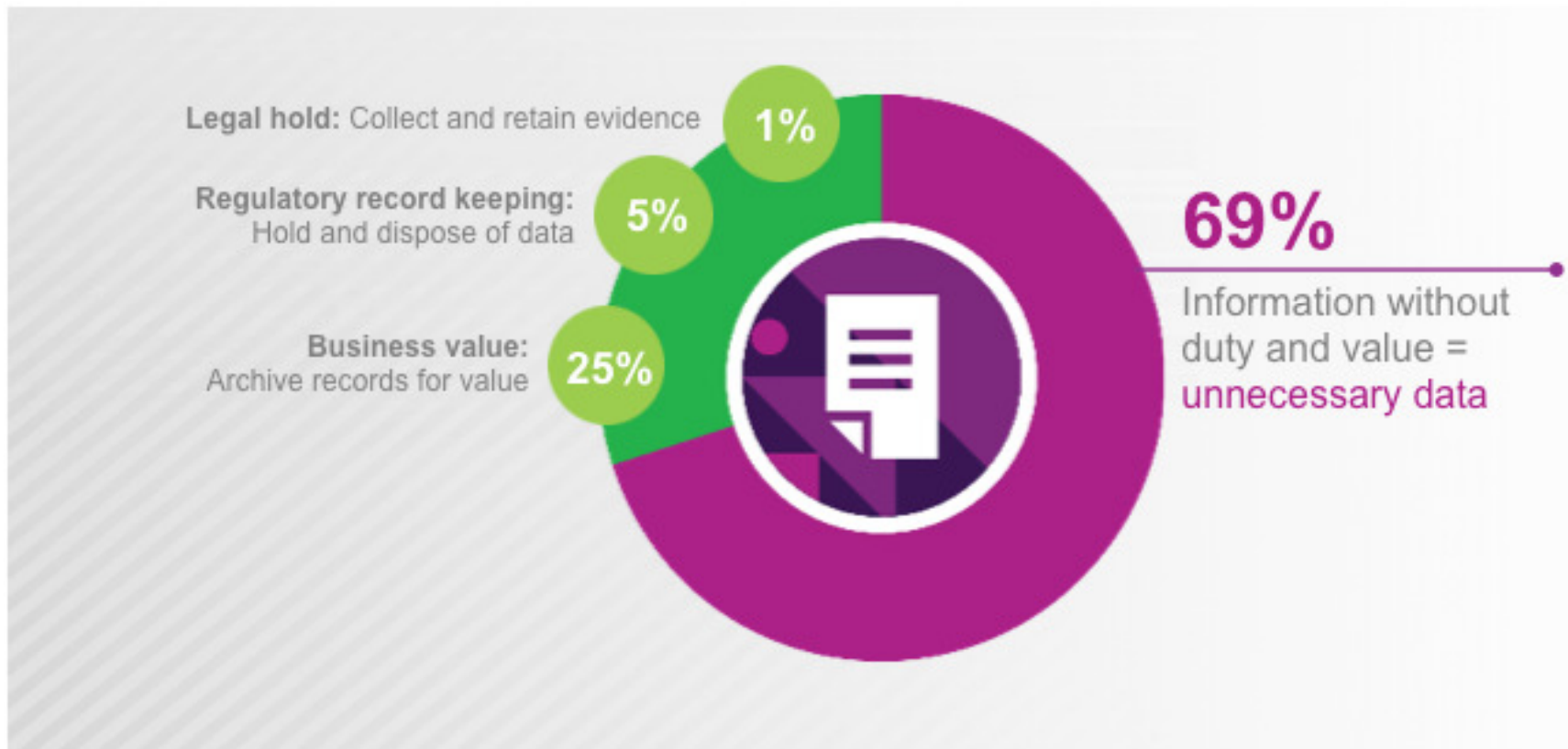**Supply-Side initiates are only a temporary stop-gap**
- **Virtualization**
- **Over-Allocation**

This growth represents an enormous challenge to IT organizations

1 Cost

~$4 million annually

Cost to store, administer and back up 1 PB of data[1]

2 Risk

Highly classified information

Personal information

Data subject to legal review

3 Value

Data center transformation

Big data analytics

Acquisitions and divestitures

Excess information = higher cost and greater risk

Dispose of unnecessary data = **reduce** cost and risk

Legal hold: Collect and retain evidence — **1%**

Regulatory record keeping:
Hold and dispose of data — **5%**

Business value:
Archive records for value — **25%**

69%
Information without
duty and value =
unnecessary data

Where do we start?

Litigation & Legacy Data Clean Up

Legacy data no longer has value but creates cost and risk

| Office Docs |
| Product Research |
| Sales/ Customer |
| HR |
| Financials |
| Messaging |
| IT Cost |
| Risk |

Trapped Value

Risk

Cost

Risk-to-Value Gap

Value

Cost-to-Value Gap

Legacy Data

6 Months   1 Year   2 Years   3 Years   4 Years   5 Years   > 7 Years   > 10 Years   > 15 Years   > 20 Years
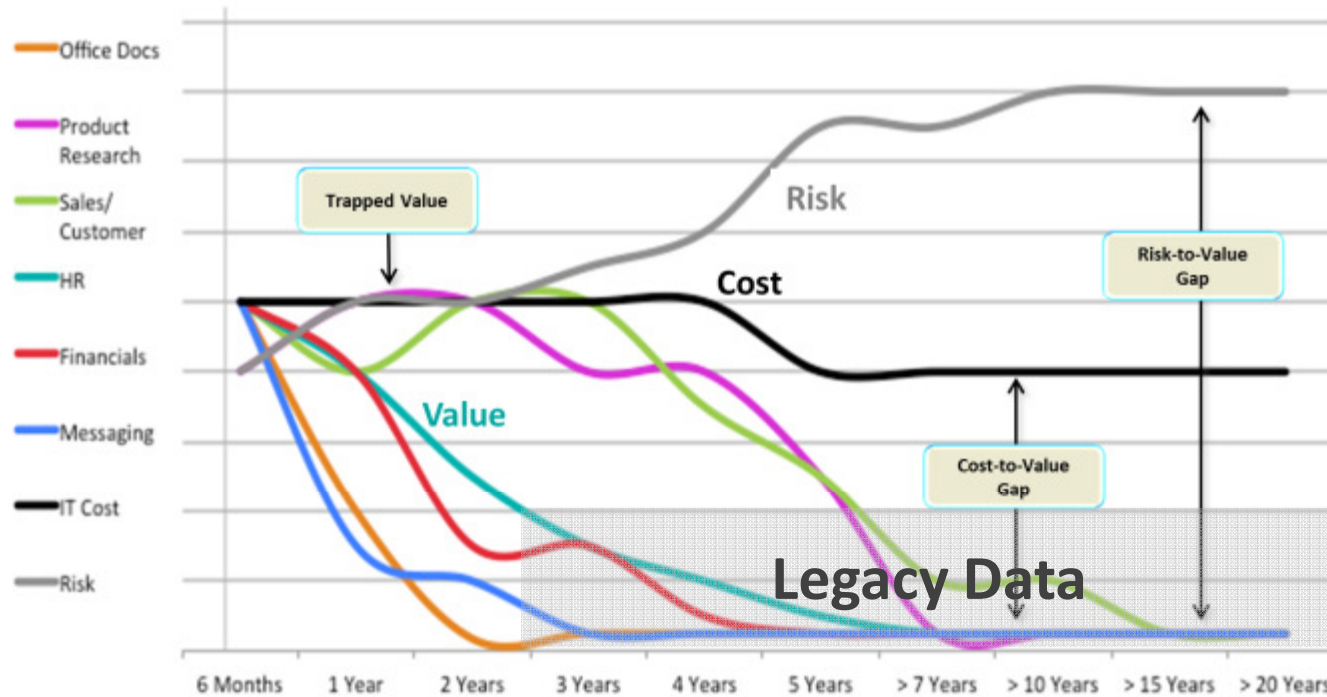
Legacy data is:

- Data that has

- Data that is **consuming COST without providing value** to the organization

- Data of **low value that still carries RISK** in legal actions

Use cases

Clean-up ROT Data

Redundant, Obsolete and Trivial data

Boolean
Search

Remediate Regulated Data

PII, PCI, HIPAA, HR, Financial, Records

Pattern Match
& Classify

Secure High Business Value Data

IP, Pricing, Sales and Market, Patent, Planning

Bayesian
Classification

eDiscovery Identification & Collection

Early Data Assessment & Targeted Collection

Boolean
Search

Three-phase reduction of ROT data

1

Phase 1: Remove trivial data
- Data that never had any value to the organization
- Violates acceptable use policy
- Multimedia files (audio, video and images)
- Temporary files

2

Phase 2: Move obsolete data for timed disposal
- Orphaned files (employee has left the company)
- Log files
- Files past their longest departmental retention or business value
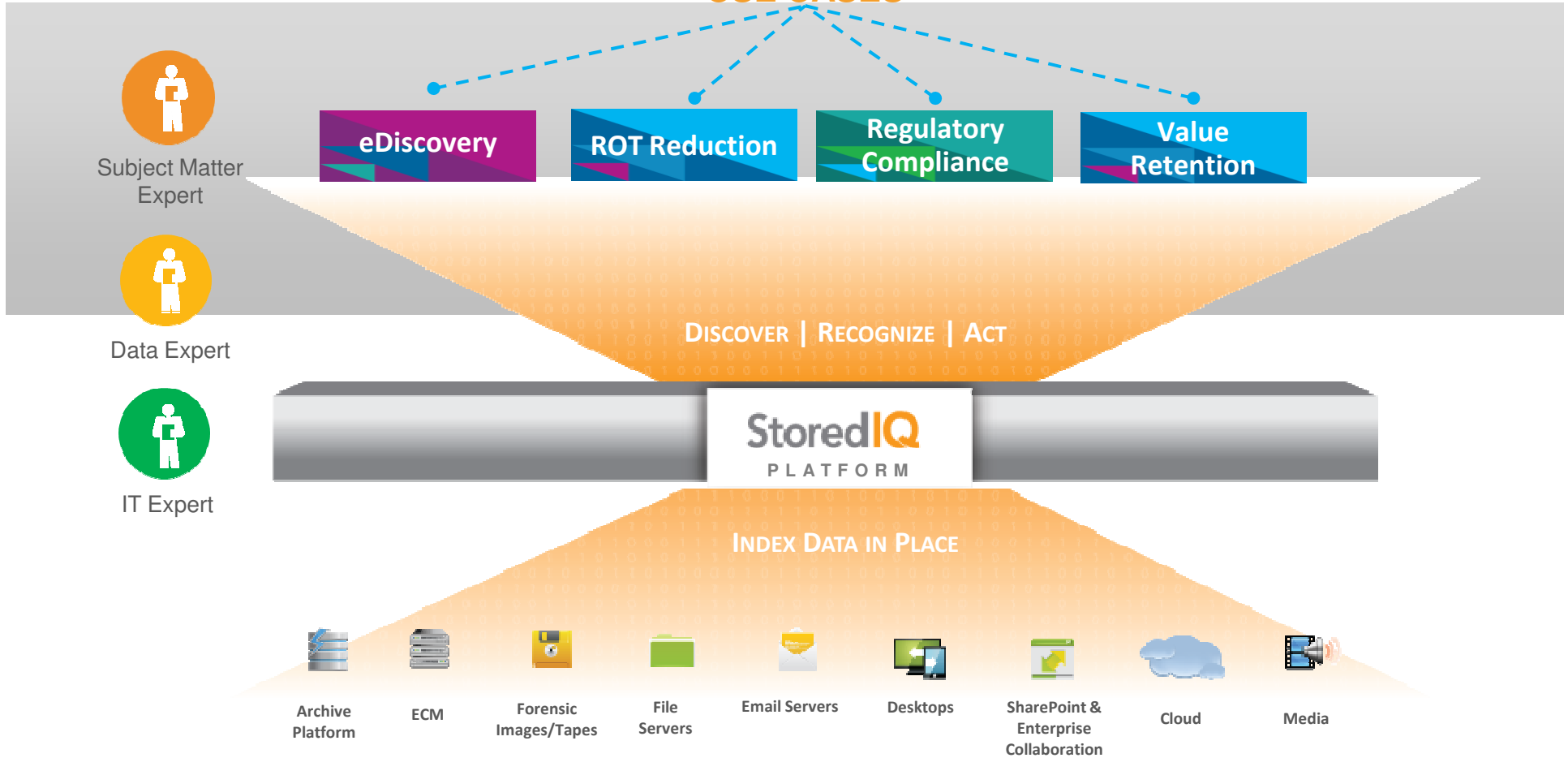
3

Phase 3: Remove departmental deduplication
- Create departmental master copy area (ensure that all department employees have access)
- Identify master copies and place in the correct area
- Deduplicate remaining share against the master area
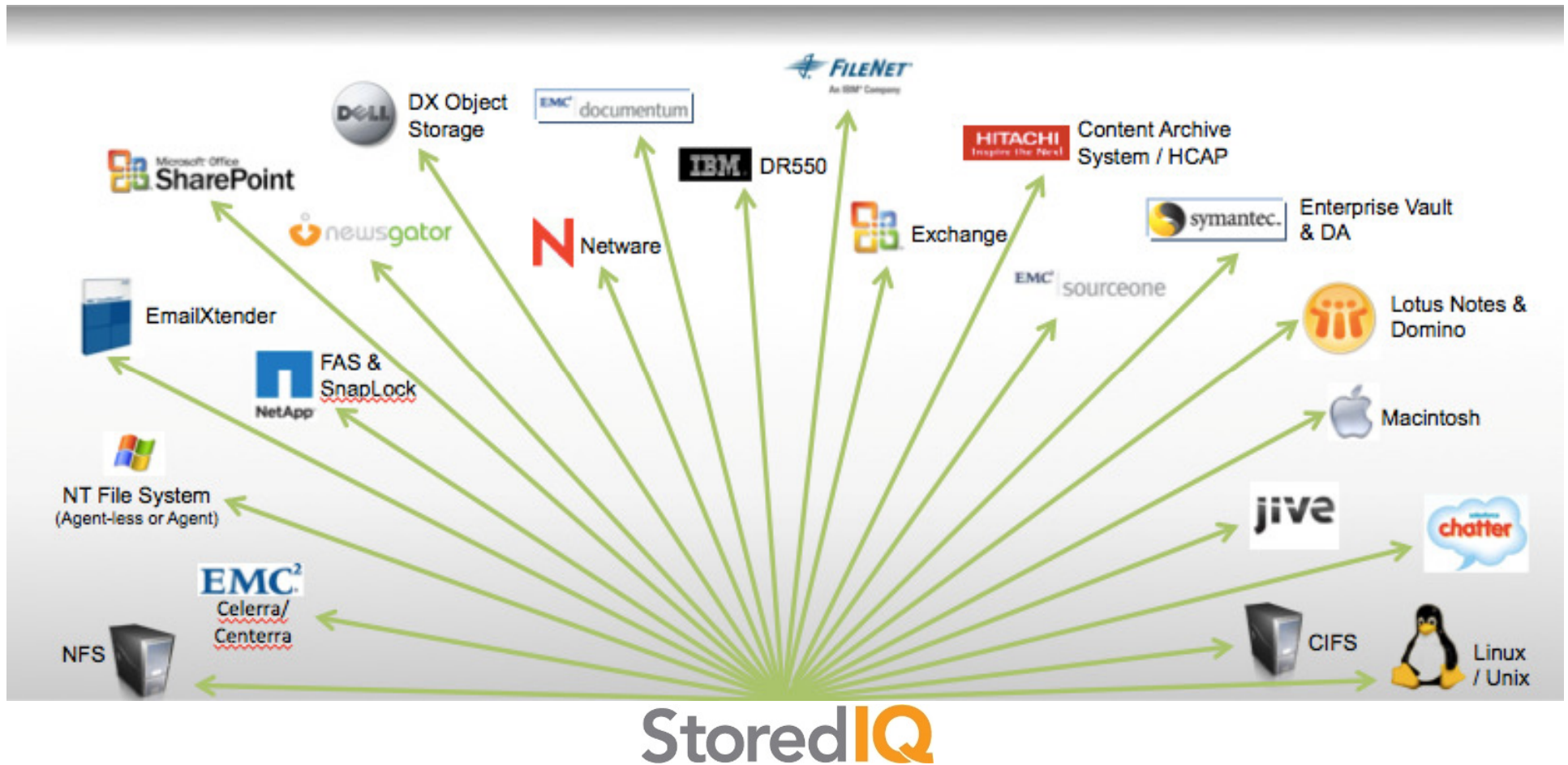
StoredIQ Solution

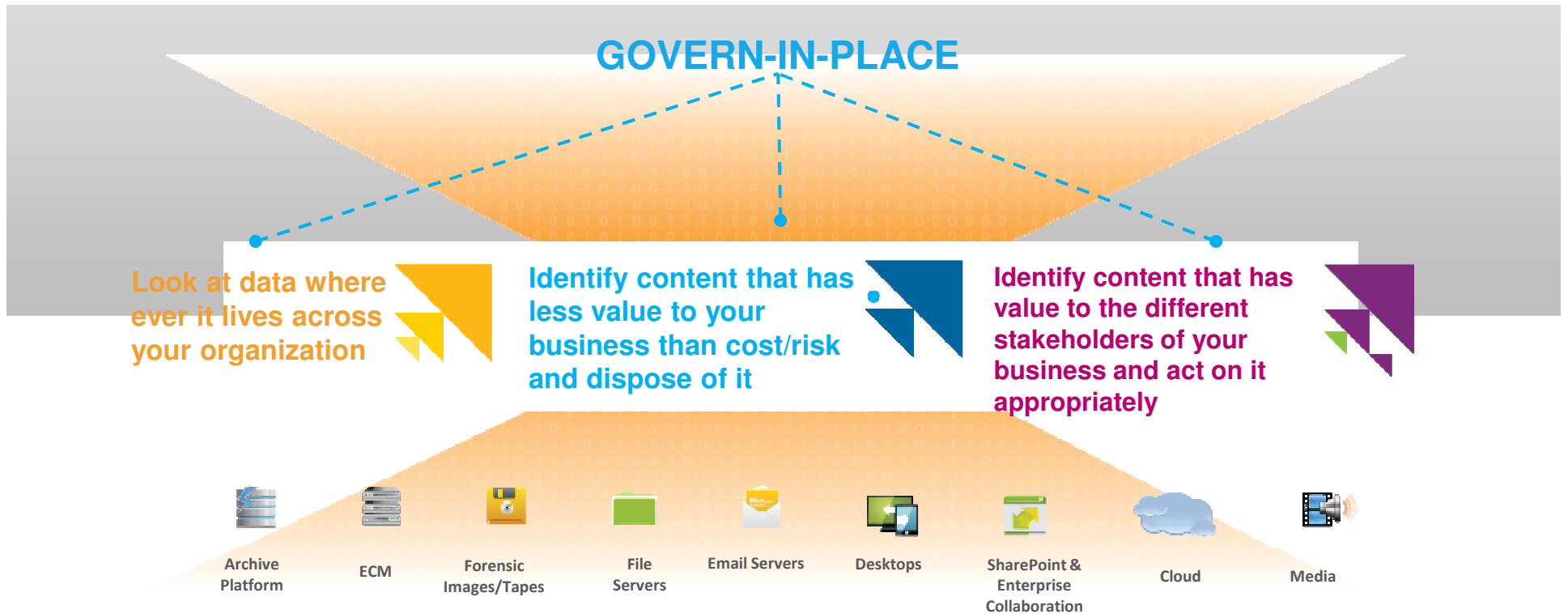## Clean Up Legacy Data

StoredIQ platform and solutions

USE CASES

eDiscovery

ROT Reduction

Regulatory Compliance

Value Retention

Subject Matter Expert

Data Expert

IT Expert

DISCOVER | RECOGNIZE | ACT

StoredIQ
PLATFORM

INDEX DATA IN PLACE

Archive Platform

ECM

Forensic Images/Tapes

File Servers

Email Servers

Desktops

SharePoint & Enterprise Collaboration

Cloud

Media

Connect to data in its native location

Support for 75+ data sources and 450+ file types

Govern-in-Place

GOVERN-IN-PLACE

Look at data where ever it lives across your organization

Identify content that has less value to your business than cost/risk and dispose of it

Identify content that has value to the different stakeholders of your business and act on it appropriately

Archive Platform   ECM   Forensic Images/Tapes   File Servers   Email Servers   Desktops   SharePoint & Enterprise Collaboration   Cloud   Media

StoredIQ approach: data about data

DATA INTELLIGENCE

DISCOVER

RECOGNIZE

ACT

1
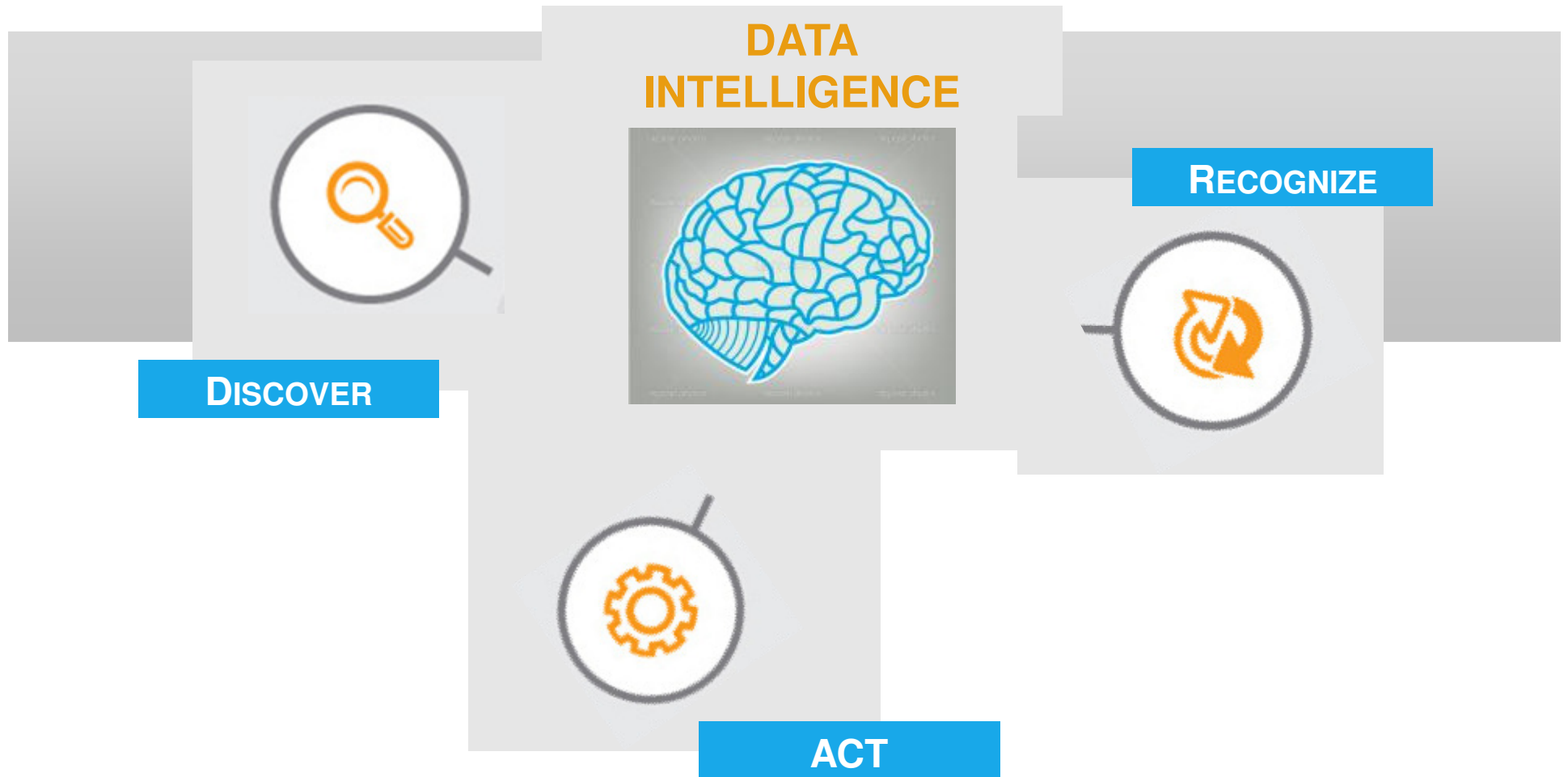
Advanced visualizations show what types of data are stored across your enterprise
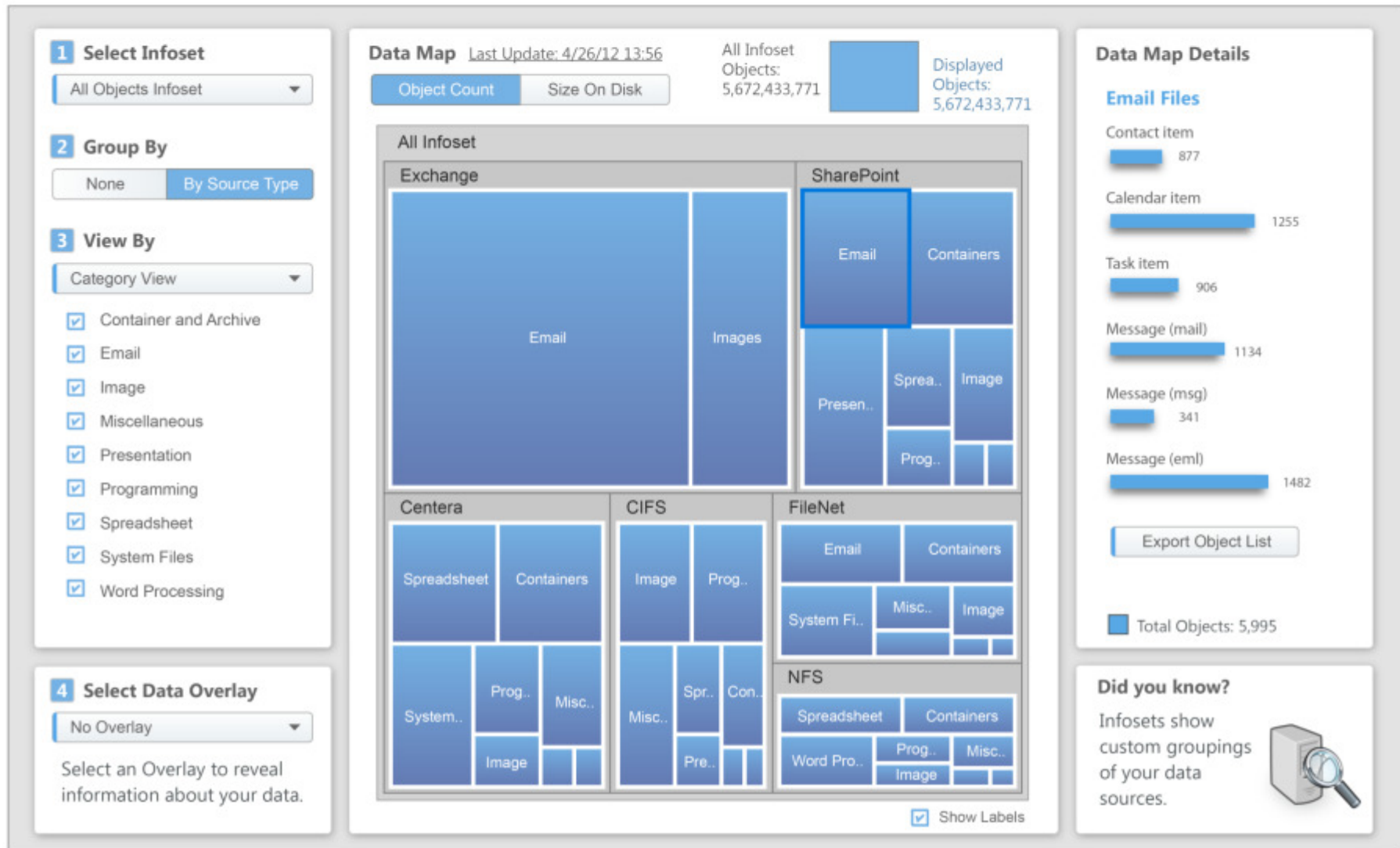
Discover where your oldest or least used data resides

Utilize intelligent overlays to spot potential compliance issues
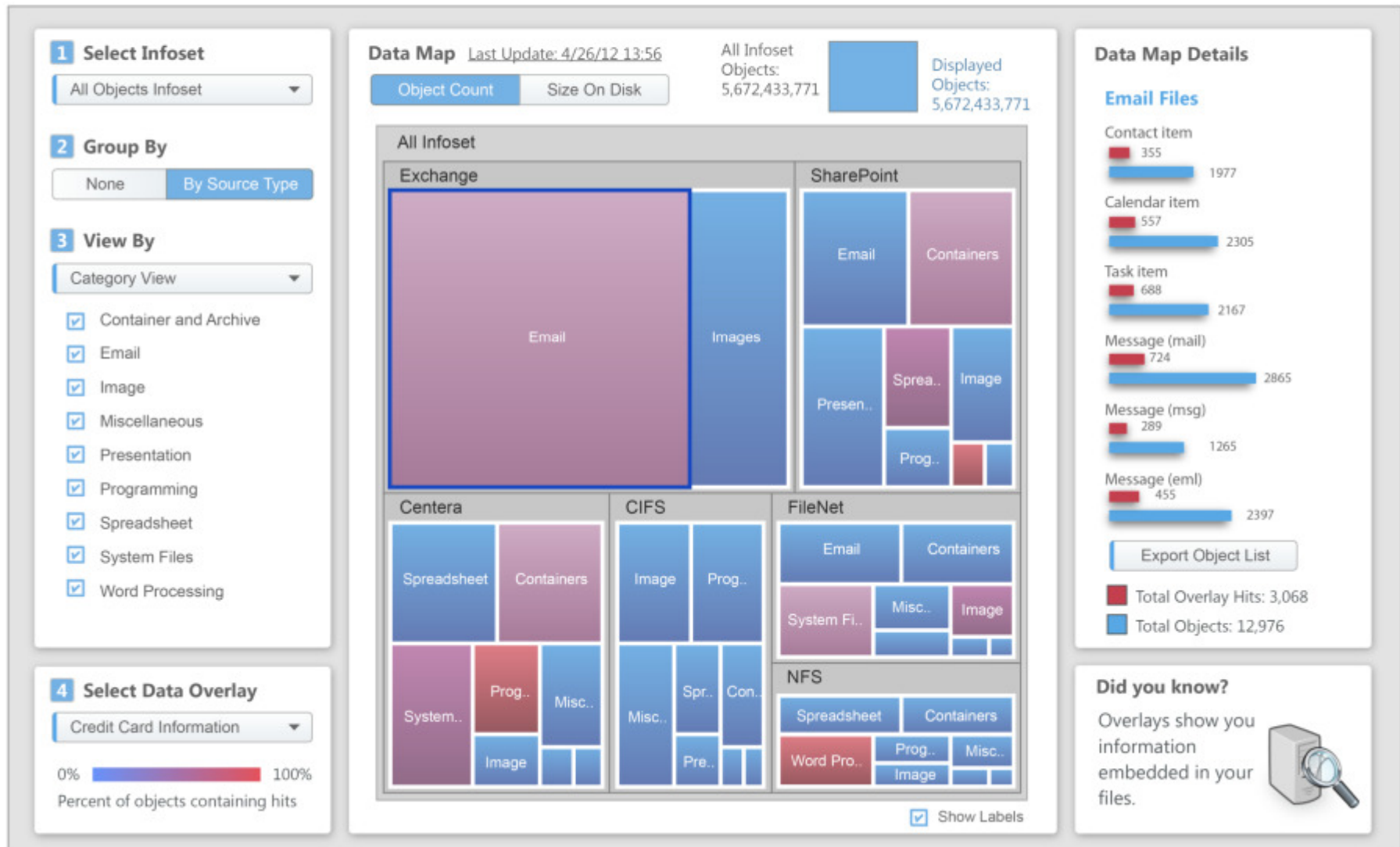
Key benefits

Find the data that matters — Properly discover, classify and manage information according to business value to reduce risk and cost

Get rid of old, obsolete data — Delete nonbusiness, aged and obsolete data to reduce data volume

Identify sensitive and toxic content — Find misplaced client data, PCI and privacy-regulated data

Stratify information to accelerate:

- Cloud migration
- Investigations
- Postacquisition and merger data integration

Business process readiness — Practice proactive audit, investigation or disclosure and discovery readiness

Thank You!
Follow me on Twitter - @AmirTheHammer

19