



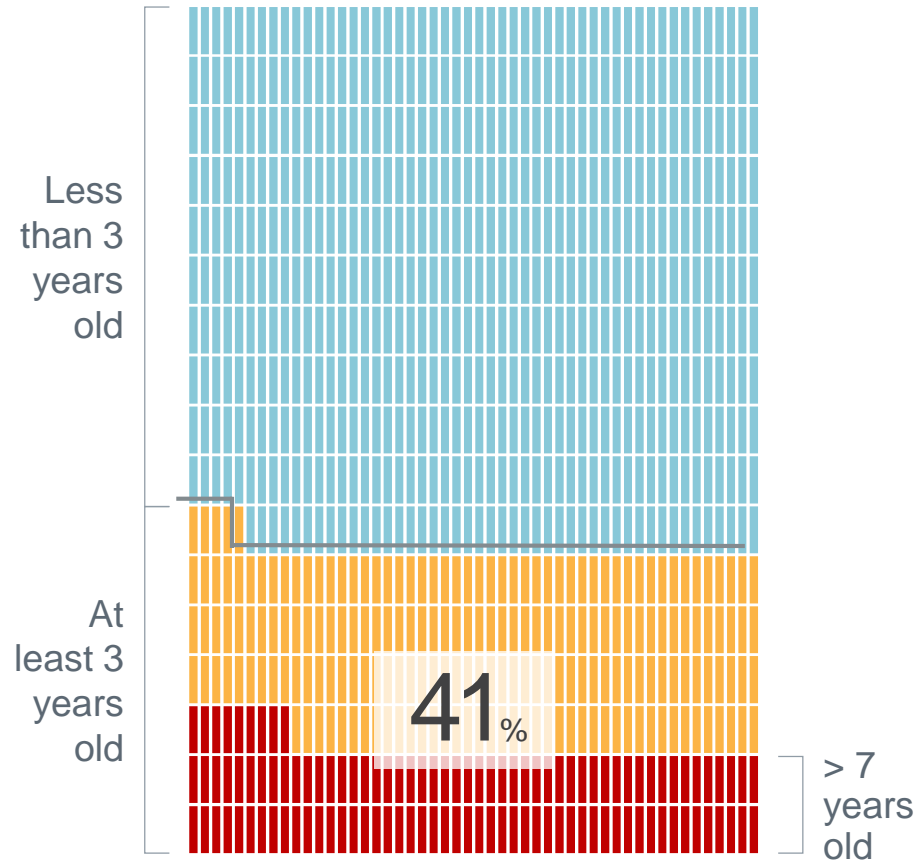
Preventing Information Chaos: The Importance of Gaining Insight into Your Unstructured Data

Marcus Loh

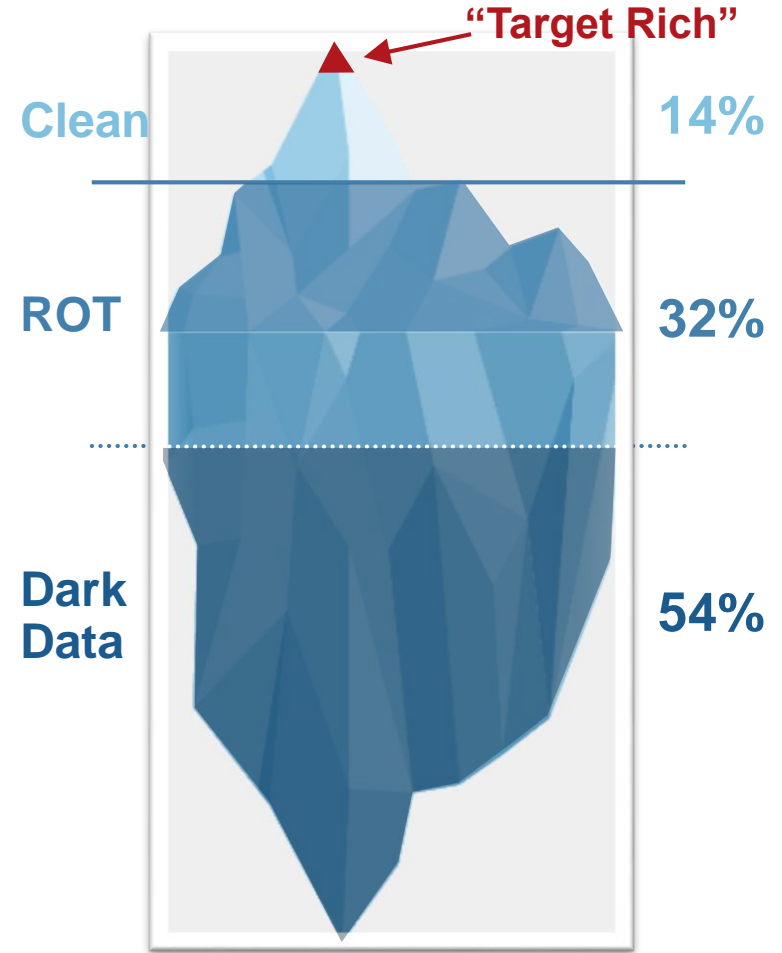
Director – Technology Practice

Too Much Data, Not Enough Information

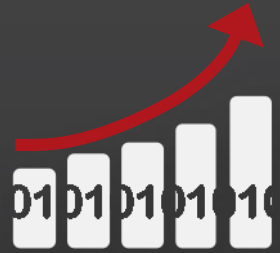
Data Genomics Index



Survey Results



INFORMATION CHALLENGES



EXPLOSION

More devices generating/storing more data.

- More volume
- Rising storage spend



FRAGMENTATION

Multiple clouds, hypervisors, and storage platforms.

- Increased complexity
- Diminished agility

99.999

AVAILABILITY

Always-on, always-accessible requirements

- Less visibility
- Less control
- Less assurance
- No leverage

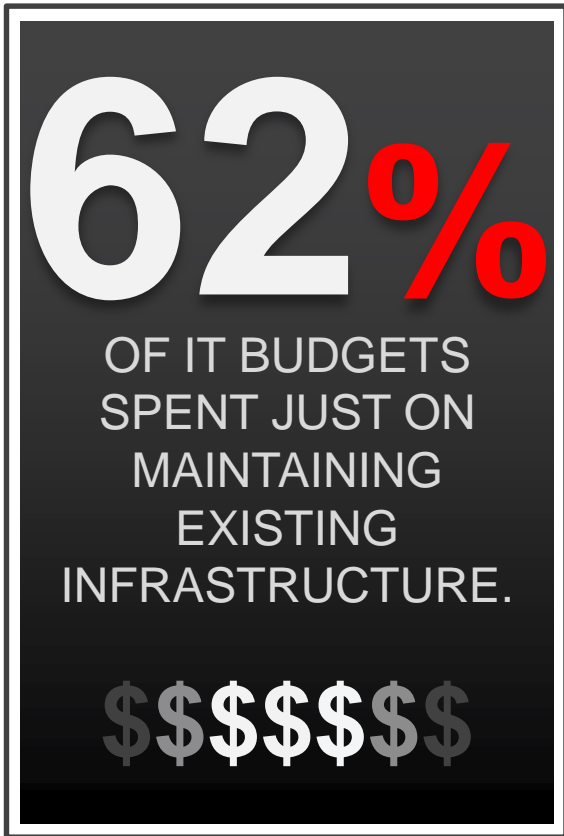


SECURITY

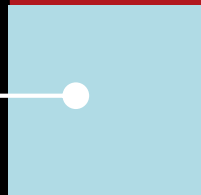
Protection against malicious threats & data breaches

- Complex requirements
- Increasing exposure

BUSINESS CHALLENGES



Structured
Data



Unstructured
Data







Problem 1

Most organizations have no idea what information they have, what it is worth, or why they keep it. But they will keep it. Forever.



Problem 2

Most organizations have no idea where their sensitive, regulated, or most valuable information is, nor who has access to it.



Problem 3

When organizations need to actually find information it takes too long, costs too much, and usually is missing results or has too many results. Or both.



Keep applying the same solution...

As Data grows – The risk exposure increases

Financial Risk

- Storage costs become exorbitant
- Data discovery costs are extreme

Legal/Compliance Risk

- Ability to meet deadlines is compromised
- Data overload limits effective formulation of case strategy

Security Risk

- Lack of intelligence regarding information under management opens door to malicious activity

Productivity Risk

- Employees waste time managing their inbox or searching for important documents

Reputational Risk

- Compounding effects of risk can lead to negative press coverage or potential customer distrust

INFRASTRUCTURE
IS IMPORTANT, BUT INFORMATION
IS WHAT MATTERS

OVERCOMING INFORMATION CHALLENGES, REQUIRES A **FOCUS ON INFORMATION.**

Insight – A Better Way Forward



KNOW
WHAT YOU HAVE

LEVERAGE
WHAT YOU KNOW

MAKE BETTER
DECISIONS

- **Gain Visibility:** Understand your gap and devise strategy
- **Take Action:** Implement policy that aligns behavior with goals
- **Assume Control:** Improve data management and risk practices

Make better informed decisions with consistent insight into your environment

Challenge → variables → gather & view data → make decision

What can I get rid of?

What should I get rid of?

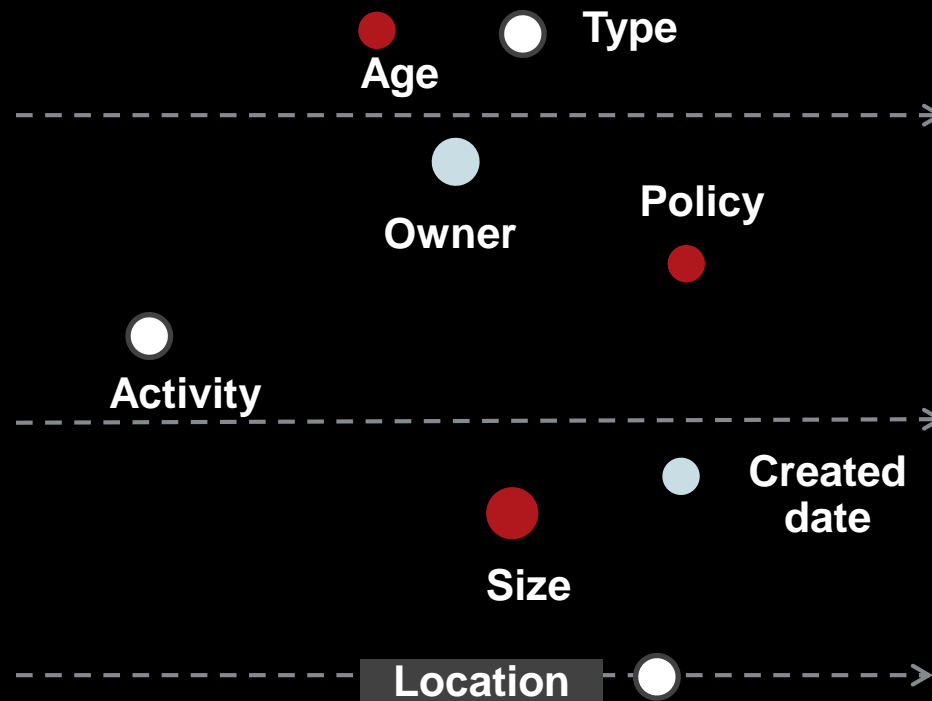
What can I archive?

What policy should it be under?

Where should I look?

What is being used?

What is valuable?



- Decommission servers and shares
- Address stale data
- Departmental

- Focus eDiscovery collections on particular custodian datasets
- Prioritize high activity shares for extra protection

- Identify PSTs for migration
- Identify non-business data for deletion

Dashboard



<p>Total</p> <p>21.46TB</p> <p>45,347,205 items</p>	<p>11%</p> <p>Orphaned</p> <p>2.54TB</p> <p>12,557,920 items</p>	<p>35%</p> <p>Stale</p> <p>7.60TB</p> <p>27,667,672 items</p>	<p>1.1%</p> <p>Non-Business</p> <p>238.18GB</p> <p>102,290 items</p>
--	--	---	--

Locations (3)

Total Size

Reading	15TB
San Francisco	5TB
Tokyo	2TB

List view



Content Sources (7 of 21)

Total Size

plead	5TB
face	3TB
admit	2TB
learn	2TB
win	1TB
soak	1TB
load	1TB

List view

Owners (7 of 3,588)

Total Size

aa	7TB
NA	3TB
aftermath	623GB
additerebbero	530GB
aaahs	429GB
aardwolf	356GB
agaceries	353GB

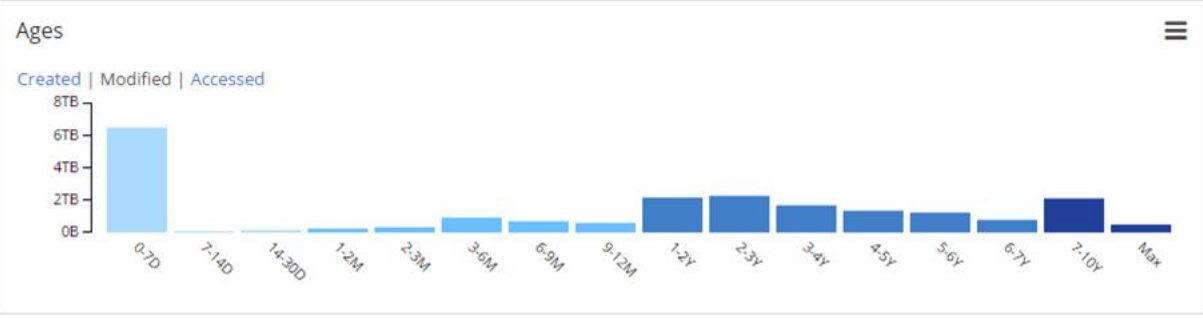
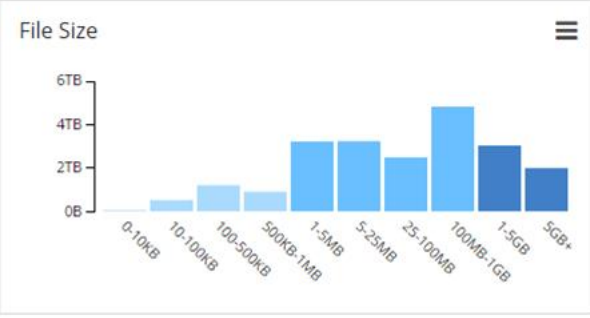
List view

Shares (7 of 38)

Total Size

tourism	3TB
computer	2TB
diet	2TB
announcement	2TB
neglect	1TB
means	1TB
wording	1TB

List view



Item Types (7 of 24)

Total Size

Unknown	6TB
Spreadsheet	3TB
Document	3TB
Compressed/archive	3TB
Text	1TB
Email	1TB
Database	1TB

List view

Preliminary Information Map Findings

Select Customer Deployments

Customer A

45TB

34% Stale

Customer B

80TB

45% Stale
1,004 Servers

Customer C

151TB

47% Stale

Average 41.2% of
data hasn't been touched in
3 years !!!

- 1 share with **73% stale info**; another **90% older than 1 year**
- Millions of small 'audio' files located
- **31 TB of unknown** non-business data

Early Case Study: “I don’t know what I don’t know”

No Visibility, No Action, No Control

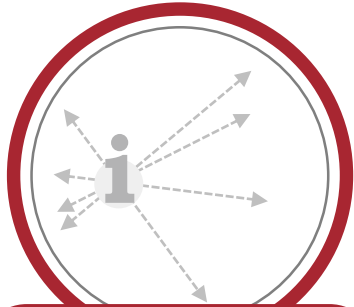
Upon ingestion of 64.81 TB noticed some crazy things:

- *45% of the environment was stale* – untouched in 3 yrs.
ACTION #1 – Use this to kick off an Archive project.
- Explored some more and realized 2 outlier datacenters with <50 GB of data
ACTION #2: Build a plan to decommission the locations and bring them into the centralized datacenter
- They’re coming up on an email refresh.
ACTION #3: Identify and audit PSTs throughout the environment so they only migrate what is needed.
- While exploring the data noticed 3 million items with audio extensions but <10kb in size.
ACTION #4: Investigate the anomalous data set further.

The Next 90 Days: Key Next Steps

- 1. Understand What You Have** – Develop a plan for getting visibility into your unstructured data including where it's kept and ownership
- 2. Know the Influencers and Start the Policy Discussion** – Getting agreement amongst CIO, Business Units, Legal, can take time; CDO can use a mandate break down governance silos
- 3. Prioritize & Tackle One Problem at a Time** – Determine where the greatest liability or gain can be had (e.g., Sensitive Information Access, PST remediation, stale file cleanup)

VERITAS | NEXT GENERATION SOLUTIONS



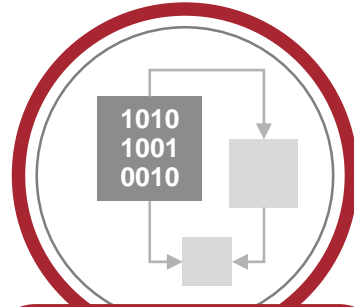
**Veritas
Velocity**



**Veritas
Resiliency
Platform**



**Information
Map**



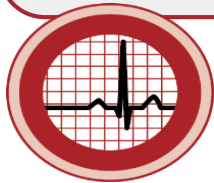
**Retention
Management**



**BACKUP &
RECOVERY**



**STORAGE
MANAGEMENT**



**BUSINESS
CONTINUITY**



**DISASTER
RECOVERY**



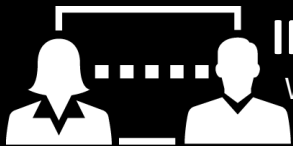
ARCHIVING



eDISCOVERY



**FILE
CLASSIFICATION**



INFORMATION AVAILABILITY

WHERE YOU NEED IT, WHEN YOU NEED IT,
WHEREVER IT RESIDES



INFORMATION INSIGHT

KNOW WHAT YOU HAVE, LEVERAGE WHAT
YOU KNOW, MAKE BETTER DECISIONS