DATUM IN INTEGRITATE, SECURITATEM, PRIVATUS
Avoid “Headline Risk” by Understanding and Managing Private and Sensitive Data Vulnerabilities

DONALD SOULSBY
SANDHILL CONSULTANTS
Session D102 - May 16, 2017
Who we are:

Sandhill Consultants
Is a group of federated people, partners, products and processes that help our clients:
• Build comprehensive Data Architecture
• Resulting from a persistent Data Management Process
• Founded on a robust Data Governance Practice

Donald Soulsby
Vice President Architecture Strategies, Sandhill Consultants
• CMMI – Enterprise Data Management Expert (EDME)
• Enterprise Architecture, Business Intelligence, Metadata Consultant
• Data/Process modeller
• Educator & Industry Speaker

Meta
Wright
person who creates, builds, or repairs something specified
No Good News

No headline about the airship that didn’t crash…

World reels from massive cyberattack that hit nearly 100 countries
It is not just about Locks and Keys!

Companies are no longer judged by whether they have a breach, but rather on how they respond when a breach occurs.

Michael Bruemmer, VP, Experian Data Breach Resolution

Two-thirds of Canadian companies say they are losing the war on cyber security

CEO.CA 2017-02-09
…By any other name

<table>
<thead>
<tr>
<th>Entity</th>
<th>Year</th>
<th>Recs (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yahoo</td>
<td>2013</td>
<td>1,000</td>
</tr>
<tr>
<td>Yahoo</td>
<td>2014</td>
<td>500</td>
</tr>
<tr>
<td>Adobe Systems</td>
<td>2014</td>
<td>152</td>
</tr>
<tr>
<td>eBay</td>
<td>2014</td>
<td>145</td>
</tr>
<tr>
<td>Anthem Inc.</td>
<td>2015</td>
<td>80</td>
</tr>
<tr>
<td>Sony PS Network</td>
<td>2011</td>
<td>77</td>
</tr>
<tr>
<td>JP Morgan Chase</td>
<td>2014</td>
<td>76</td>
</tr>
<tr>
<td>Target Corporation</td>
<td>2014</td>
<td>70</td>
</tr>
<tr>
<td>Home Depot</td>
<td>2014</td>
<td>56</td>
</tr>
</tbody>
</table>

*advanced persistent threat
When things don’t go...

According to plan!

Get the blueprints.
Data Management Blueprints
Data Vulnerability Risk Assessment

- Classification
- Valuation
- Detection
- Mitigation
Risk Assessment: Classification

“Data is a Managed Corporate Asset”

Sales Revenue Amount  Tower of Tuna Count

Not all assets need to be managed the same way!
It’s 10pm, do you know where your DATA ASSETS are?

Mel Epstein, New York's WNEW-TV, 1967
Classification

Aristotle

Carolus Linnaeus
Data Asset Inventory

Data Breach Target
Data Governance Metamodell

Business Concept or Category
- Business Term
  - Qualified Term (Element)

Subject Area
- Entity
  - Attribute

Association Entity

Business Information Model

Technical Data Model

Data Element

Relationship

IDS

KEYS
Data Asset Prioritization

Critical data elements (CDEs) are defined as “the data that is critical to success” in a specific business area (line of business, shared service, or group function), or “the data required to get the job done.”

Competing with High Quality Data: Concepts, Tools, and Techniques for Building a Successful Approach to Data Quality
by Rajesh Jugulum July 2014
How much is your Data Worth?

- Passwords
- Health information and medical records
- Social Security numbers
- Payment details
- Purchase history
- Physical location (GPS)
- Home address
- Photos and videos
- Marital Status
- Name and gender

Critical Data Element – ISO 11179

Metadata Rules:

• State in the singular
• State what the concept is, not only what it is not
• State as a descriptive phrase or sentence(s)
• Contain only commonly understood abbreviations
• Express without embedding definitions of other data or underlying concepts
Critical Data Element - Risk

- **Accessibility**
  - Secure - Intranet
  - Monitored - Extranet
  - Unsecure - Internet

- **Reliability**
  - Trusted
  - Verifiable
  - Untrusted

- **Confidentiality**
  - Sensitive
  - Confidential
  - Internal Use
  - Public
Finding Critical Data Elements
Data Reference Model
Critical Master Data

- Process
- Transaction
- Molecule
- Customer
- Calendar
- Product
- Location

$42
Identifying Critical Data Elements

Regulatory Compliance
Big Data Model
Model Driven Impact Analysis

- Business Glossary
- Business Domain
- Conceptual Data Model
- Logical Data Model
- Physical Data Model
- DATABASE
- Technology Domain
- Critical Data Element
- Concurrent Engineering

- TOP DOWN
- MIDDLE OUT
- BOTTOM UP

PREVENTIVE

FORENSIC
Investigate the Data Management Process

Data Management Body of Knowledge
The DMM model, first announced in May 2014, was developed using the principles and structure of CMMI Institute’s Capability Maturity Model Integration (CMMI®).

The DMM℠ model outlines data process improvement across business lines — allowing executives to make better and faster decisions using a strategic view of their data.
Data Management Maturity Model

LEVEL 1: PERFORMED
LEVEL 2: MANAGED
LEVEL 3: DEFINED
LEVEL 4: MEASURED
LEVEL 5: OPTIMIZED

- Data Management Strategy
- Data Governance
- Platform & Architecture
- Data Operations
- Data Quality

Supporting Processes
Form & Substance

**FORM**
- Data Structure & Context
- The definition and meaning of the data

**CONTENT**
- Data Content
- Data Quality dimensions and standards

---

Data Governance

Data Management Strategy

Data Quality

Platform & Architecture

Data Operations
Column Stewardship

Governance Steward

- Data Metadata - Party who is responsible for the decision about WHAT data the enterprise will collect and maintain the data represented by the Column

<table>
<thead>
<tr>
<th>COLUMN</th>
<th>COLUMN</th>
<th>COLUMN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FORM
Cell StewardSHIP

Data Qualification Steward
• Compliance Metadata - Party who is responsible for qualifying entries per the constraints on the data.

<table>
<thead>
<tr>
<th></th>
<th>COLUMN</th>
<th>COLUMN</th>
<th>COLUMN</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROW</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROW</td>
<td>CONTENT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ROW</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Row Stewardship

Process Stewardship

• Process Metadata - Party who is responsible for the decision about HOW the enterprise will collect and maintain the data represented by the Row
Data Governance Capabilities

Best practices for ensuring broad participation in the practice and senior oversight of the effectiveness of data management.
# Glossary - Critical Data Element

- **Business Term**
  - Name
  - Description
  - Workflow Status
  - Abbreviation
  - Acronym
  - Synonym

- **Term Type**
  - Topic
    - Term
      - Qualified Term (element)

- **Origin**
- System of Record
- Business Rule
- Authority
- Scope
- Organization
- Domain

[Diagram showing relationships between Business Term and other terms, with multiple connections indicated]
Best practices for defining and implementing a collaborative approach for detecting, assessing, and cleansing data defects to ensure fitness for intended uses in business operations, decision making, and planning.
Data Quality Issues

ACCOUNTABILITY
# Measurement

## Metadata Repository Report

![Metadata Repository Report](image)

*Data Quality indicators*
Data Quality Standards

"You can't manage, what you can't measure"

Measure = Observation Standard

You can’t measure, what you can’t define
Data Quality Dimensions

- Completeness
- Accuracy
- Timeliness
- Integrity
- Coverage
- Conformity
- Consistency
- Duplication
Conformity, Duplication, Masking
Data Quality going Forward

OBSERVATIONAL Analytics

DQ%

ALGORYTMIC Analytics

DQ%

2015
Risk Assessment: Mitigation

Data Traceability Ecosystem

Business Domain

IT Domain

BREACH
Forensic Metadata

Exposure: 1/2000 sec at f/8.0
Focal Length: 400 mm
Exposure Bias: 0 EV
ISO Speed Rating: ISO 400
Flash: Did not fire
Exposure Program: Manual
Metering Mode: Pattern
Make: Canon
Model: Canon EOS DIGITAL REBEL XTi
Lens: EF100-400mm f/4.5-5.6L IS USM
ISO 20022 Financial Industry Data Dictionary
The integration of business and technical metadata has become the holy grail of Data Governance.

“But just like the holy grail, it seems to be a never-ending quest fueled by the belief that technology is going to solve all our problems.”

Rick Sherman
DM Review Magazine
Risk Governed Data

Business Information Model

SEMANTIC LINEAGE

Technical Data Model

DATA MOVEMENT LINEAGE
In Summary

• Classification
  • Definitions & Standards

• Valuation
  • Quality criteria

• Detection
  • Protection & Risk

• Mitigation
  • Life-Cycle
General Data Protection Regulation

- Implementation – May 2016
- Enforcement – May 2018
- Involved Parties:
  - Data Subject (DS)
    - Natural individual of EU member nations
  - Data Controller (DC)
    - Says how and why personal data is processed
  - Data Processor (DP)
    - Acts on the controller’s behalf
- Guiding Principles:
  - Rights of individuals
  - Data protection by design and default
  - Grounds for lawful processing
  - Mandatory personal data breach reporting
Foundation of GDPR

Data Ownership

Personal Data in Systems

The rights of the Individual
How to Get Started?

CMMI – Data Maturity Model (DMM)

- Need a Map
- Need to know:
  - Where you are
  - Where you want to get to
  - How to get there

Sandhill Consultants
DMM Assessment
Data Management Maturity Assessment

POWER OF 3

EDME
Sandhill Consultants Solutions

Expertise ◇ End-to-End Solutions ◇ Reputation

- Strategy
- Governance
- Quality
- Architecture
- Operations

SANDHILL CONSULTANTS Capability Delivery

CMMI® Institute Partner
Tibetan Proverb

“If upstream is dirty, downstream will be muddy”

Thank You

Don.Soulsby@SandhillConsultants.com

Contact:
Robert.Lutton@SandhillConsultants.com