# Themes

- Big Data
- Governance and Framework Development
- Measurement and Value
- A roadmap approach to Data Lifecycle Management

knoware
The Knowledge Warehouse Limited

# Data Lifecycle Management: What and why

## The What

- Data lifecycle management (DLM) is an approach to managing the flow of **data, information** and **associated metadata** through information systems and repositories, from creation and storage through to when it can be discarded.
- DLM encompasses **principles**, **policies**, **processes**, **practices**, **systems**, and **tools** used by an organization to manage information through every phase of its existence.
- DLM recognises that the importance and business value of data does not rely on its age, or how often it is used and reused.

## The Why

- Data and information has value for strategic and operational business needs, for managing risk, and for meeting legislative obligations.
- The value of data to the business decays over time. We don't want to, or need to, keep all of it.
- As business value decays, some information can be archived, and much of it can be discarded.
- Occasionally, sometimes unexpectedly, older data may need to be accessed again, quickly, completely and accurately

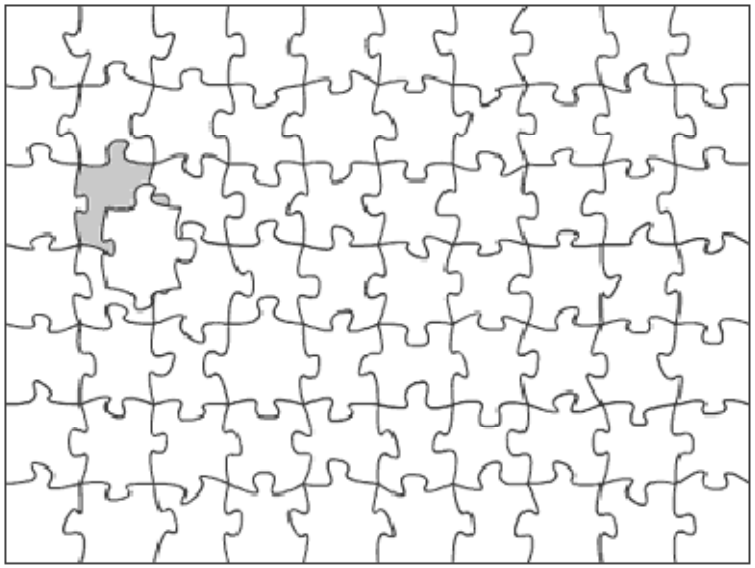
# Data Lifecycle Management: An Ill Organisation

- Siloed data – in line of business systems, databases, data warehouse, shared drives, email, filing cabinet, desk drawers…
  - No way to link **structured** and **unstructured** data together, so no single view
  - No way to determine **authoritative** source
  - Very little **described** or **defined** (limited metadata)
  - Everything is **kept** because there are no guidelines. Information is piling up, and yet…
  - …a lot is at **risk** of being hard or slow to access, **irretrievable**, **unreliable** or **lost**.
- Business requirements for lifecycle management not well understood (who needs what, when and for how long?)

# Data Lifecycle Management

**Scope:**

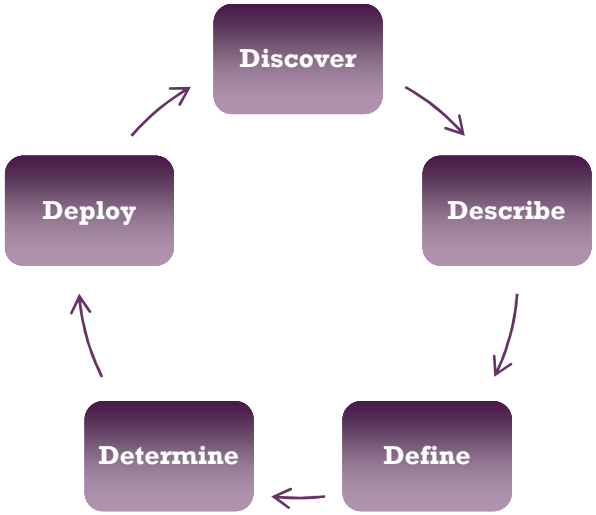
- All unstructured data or information held in personal and shared directories, email inboxes, databases that hold documents, web-based information including intranets, public websites and social media
- All structured data, including all repositories, transactional databases and the data warehouse
- Data or information in any format, including paper-based, electronic, and any other media or storage formats.



- Principles
- Policies
- Processes
- Practices
- Systems
- Tools
- Change management
- …and a Roadmap

knoware
The Knowledge Warehouse Limited

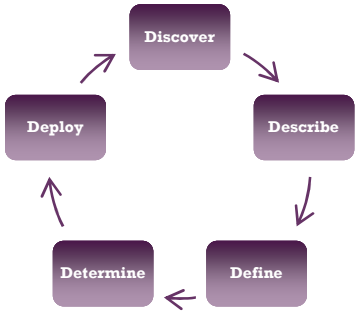
# Data Lifecycle Management: Roadmap Development




| Framework Component | Activity | 2011 | | 2012 | | | | 2013 | |
|---|---|---|---|---|---|---|---|---|---|
| | | Q3 | Q4 | Q1 | Q2 | Q3 | Q4 | Q1 | Q2 |
| **Principles** | Agree Principles | | | | | | | | |
| **Policies** | Record Policy | | | | | | | | |
| | Lifecycle Policy | | | | | | | | |
| | Metadata Policy | | | | | | | | |
| **Data value** | Business process analysis | | | | | | | | |
| | Information risk | | | | | | | | |
| | Legislative framework | | | | | | | | |
| **Data description** | Metadata Schema | | | | | | | | |
| | Classification schema | | | | | | | | |
| **Storage/Format Planning** | Existing systems | | | | | | | | |
| | Future approach | | | | | | | | |
| **Retention and Disposal** | Business requirements | | | | | | | | |
| | Disposal Authority | | | | | | | | |
| | Implementation Plans | | | | | | | | |
| | Disposal Schedule | | | | | | | | |

knoware
The Knowledge Warehouse Limited

# Discover



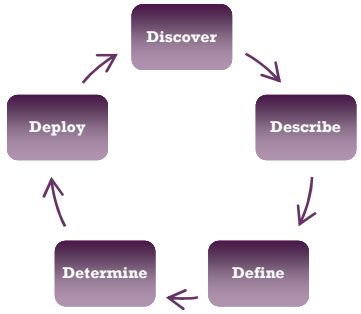



**Discover** what data assets exist and how they are used

- Map core business process data and artefacts to identify the point(s) in core business processes where active use of data diminishes.
- Determine what **datasets** are authoritative
- Determine what **data formats** are authoritative


knoware
The Knowledge Warehouse Limited

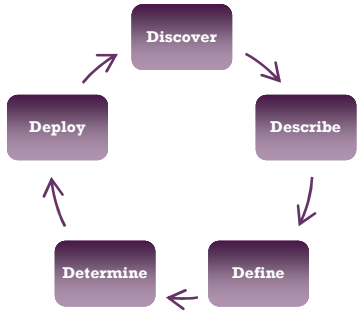
# Describe




**Describe**, characterise and label those data assets

- Develop or extend your metadata schema to include lifecycle management. Metadata includes retention period, trigger and disposal action
- Use metadata to enable data matching across systems. Data in one is metadata in another.
- You'll also need a function/activity business classification schema, so you can group things

# Define

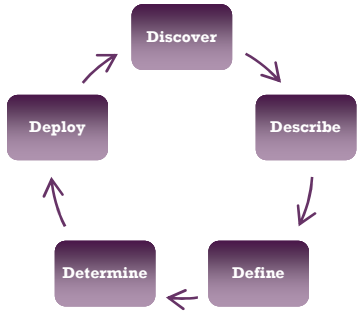





Identify and **define** business and legal requirements associated with the data assets

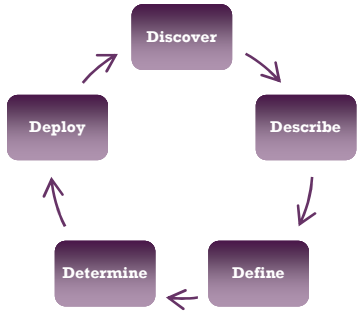
- Identify availability requirements for core datasets. Is it mission critical? How long will we need it, who needs it, and in what format?
- Assess the risk associated with data and data lifecycles. What happens if the data is not available? If it's wrong? If we got rid of it already?

# Define, cont.




- Establish a legislative framework for data lifecycle management
  - What records do we need to maintain? Which of them must be maintained beyond the lifespan of individual systems? How must we manage them?
- Plan for how to implement disposal
  - What formats and repositories? How to consolidate data into records? Do we use pre-disposal tiered storage?

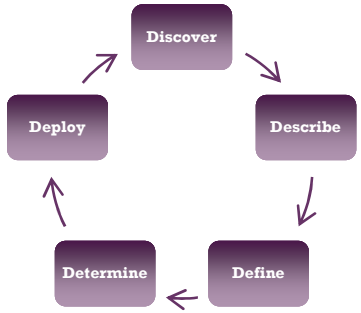
# Determine




**Determine** how you will manage the lifecycle, through principles, policies and rules for data assets

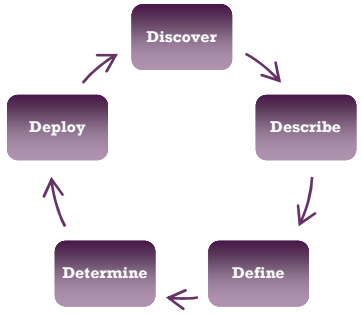
- Develop or extend principles. Examples might include:
  - Information will only be managed for only as long as it is required, and only as comprehensively as required.
  - Information that provides evidence of a transaction or business decision is deliberately managed throughout its lifecycle in such a way that it is complete, authentic, reliable, comprehensive, useable, tamperproof and has integrity (*we'll look at what this means for databases*)
  - Data and information assets will be managed throughout the lifecycle in accordance with security requirements (*note that this can change over the data lifecycle*)

# Determine, cont.




- Develop or extend other necessary policies, for example:
  - Lifecycle management policies: ie: automated systems management; storage; roles and responsibilities; decommissioning and migration
  - Records management
  - Metadata management (so that retention, triggers and disposal is captured)
- Develop business rules and guidance

# Deploy

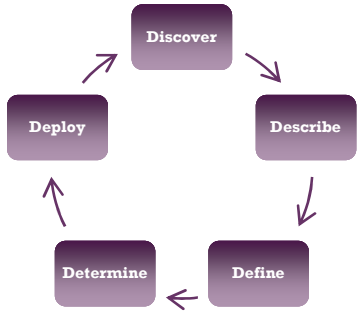





**Deploy** lifecycle management elements and artefacts to actively manage data assets over their lifecycle

- Map lifecycle artefacts (metadata, disposal classes, etc) to existing systems, and identify appropriate formats, so that…
- …data is stored in the format and repository suitable to is value at any given point in its lifecycle

# The ‘silent sixth D’




- **Do** it again, and keep doing it!
- Managing data lifecycles should become part of your overall data governance, data quality or information asset management programme

# EDW: Summary


- Next year: San Diego
- Is it worth it? Yes, if your context is a large(ish) organisation
- Who to send? Those involved in governance or architecture
- What to expect: Extremely diverse content, lots of case studies, and many 'data celebs"

We ♥ Data
DATAVERSITY

knoware
The Knowledge Warehouse Limited

t +64 4 499 9246
f +64 4 499 9245
www.knoware.co.nz

knoware

The Knowledge Warehouse Limited
Level 15 Equinox House
111 The Terrace
PO Box 10 541
The Terrace
Wellington 6143
New Zealand

# Enterprise Data World 2012 in Atlanta, Georgia

Selena Smeaton