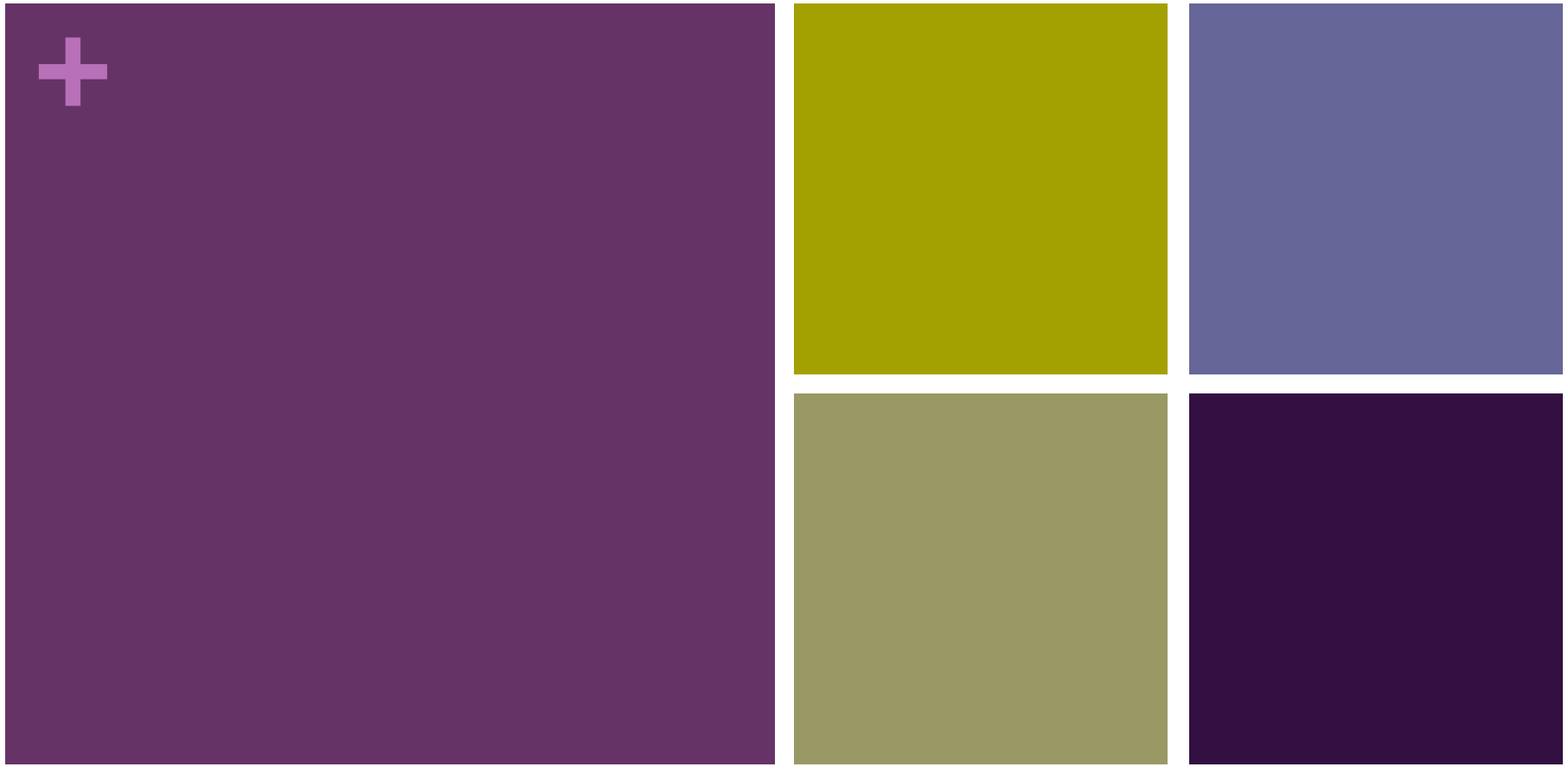
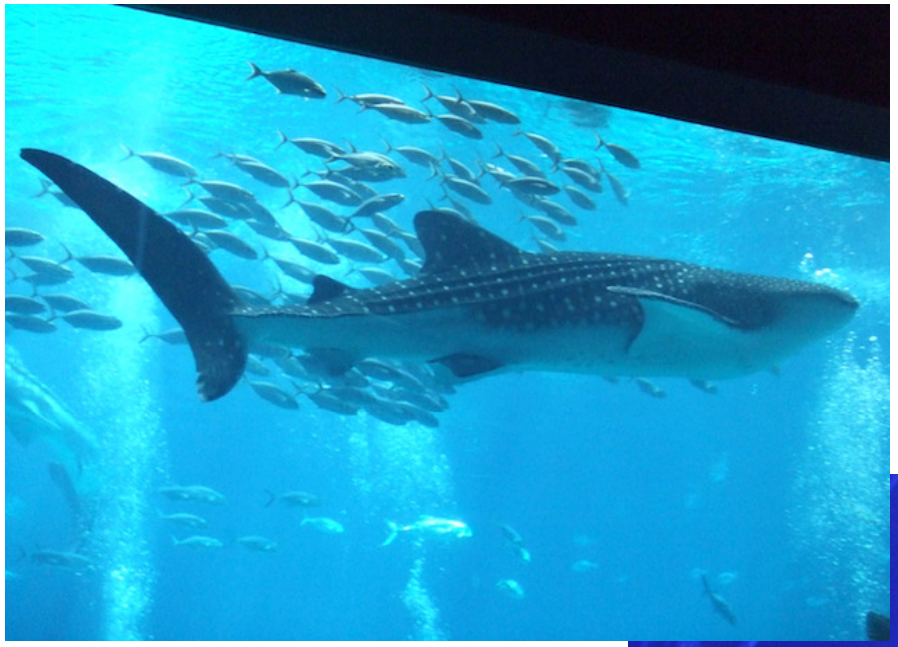


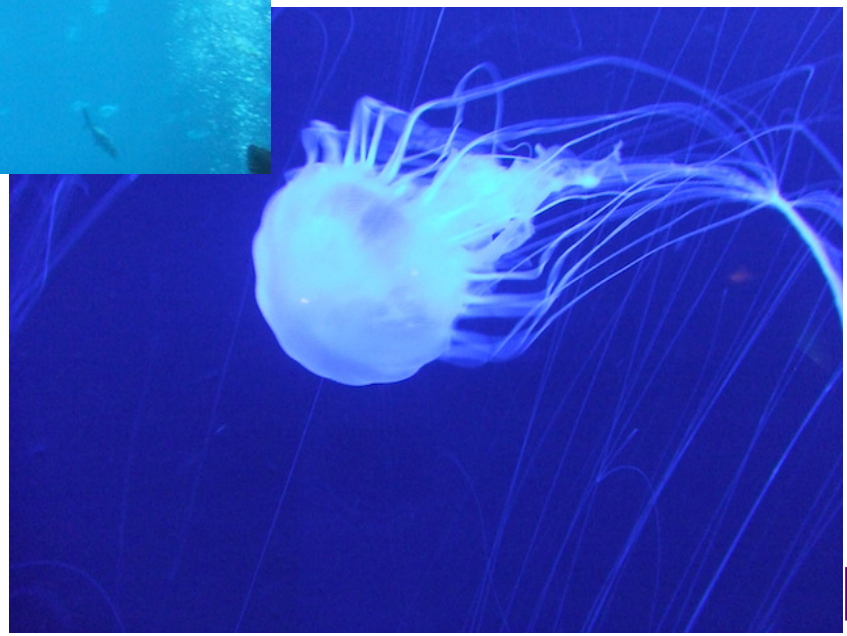
Enterprise Data World 2012 in Atlanta, Georgia

Selena Smeaton

Atlanta

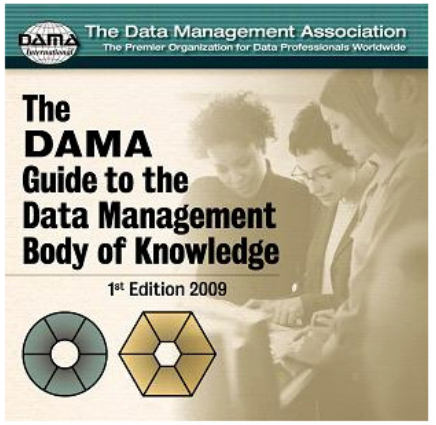
Atlanta





knoware
The Knowledge Warehouse Limited

- 5 day schedule of education, workshops, tutorials, presentations, discussions and case studies
- Endorsed by DAMA International (Data Management Int.)
- Multiple tracks:



knoware
The Knowledge Warehouse Limited

Themes

- Big Data
- Governance and Framework Development
- Measurement and Value

Key Note: Big Data

Eric Kavanagh
Host, DM Radio
CEO, Bloor Group
www.bloorgroup.com

Neil Raden
VP & Principal Analyst
Constellation Research

April Reeve
Advisory Consultant
EMC Consulting
http://infocus.emc.com/april_reeve/

Anjul Bhambhri
Vice President of Big Data Products
IBM

Robin Bloor
Principal Analyst
The Bloor Group

Paul Pedersen
Deputy CTO
10gen
www.10gen.com

John Haddad
Director Enterprise Data Integration
Informatica
www.informatica.com

Paul Barth
Managing Partner and Founder
NewVantage Partners

knoware
The Knowledge Warehouse Limited

Key Note: Big Data

- On perceptions of big data:
  - "It's not the bigness of data that is important" but the knowledge that can be gleaned from it
  - "It's a solution looking for a problem". Go back to defining the problem. Don't start with the technology or the tool

- On challenges of using big data:
  - "There needs to be a better understanding and broader use of metadata"
  - "Who will make sense of the data?"
  - "Open minds to big data analytics without overhyping the technology"

Key Note: Big Data

- On Information/Analytics/Big Data as a Service:
  - "Big data AAS will be centralised and in the could, virtualised and built project by project"
  - "I live in a world of secret and private data" Information AAS may not implement well in some industries
  - "Up to now, engagement have been between one organisation and another. We will see a rise in data aggregators"

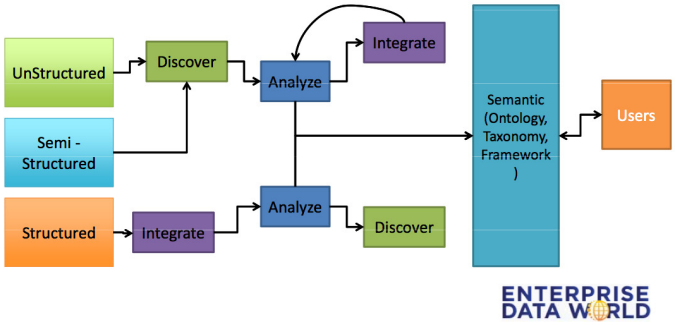
- Download the MP3 of the panel: http://www.information-management.com/dmradio/BigPanelBigData-EDW2012-DMRadio.php

Big Data: What other people said!

Metadata

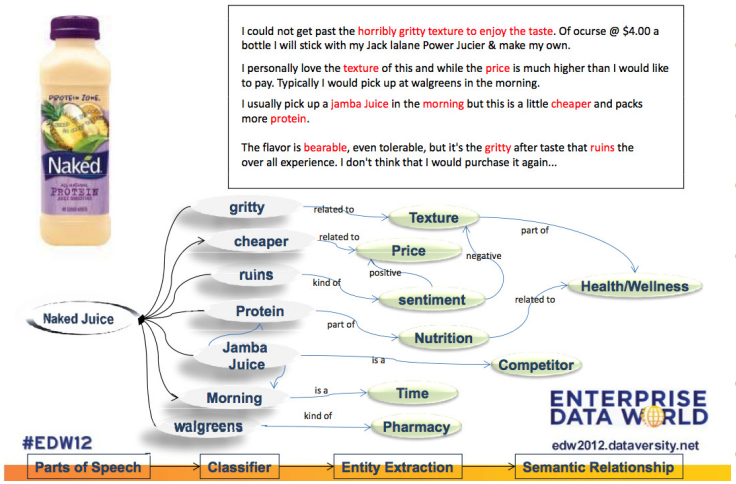
- The key to the castle in integrating Big Data is metadata
- Whatever the tool, technology and technique, if you do not know your metadata, your integration will fail
- Semantic technologies and architectures will be the way to process and integrate the Big Data.
- Business domain experts can identify large data patterns by association relationships with small metadata.

New Architecture



Semantic Integration

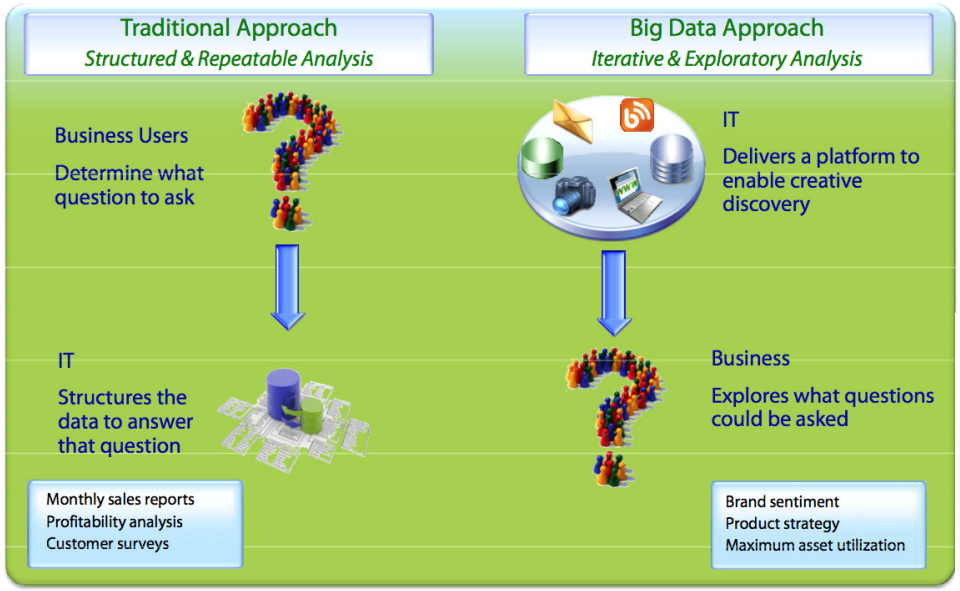
Semantic Integration Example



- Perform entity extraction from unstructured texts using advanced computational linguistics and natural language processing
- Organize unstructured data into meaningful, relevant and quantifiable knowledge automatically
- Building the data graph of the web via semantic understanding of content
- Connects people with relevant information without the need for human supervision
- Identify people, companies, organizations, topics, places and keywords; understands the relationship in the news flow, and puts them in context through interactive visualization workbench for refinement
- Support semantic reasoning and semantic query
- Link unstructured data to enterprise structure data via semantic transformation

Big Data: Requires new paradigms

Merging the Traditional and Big Data Approaches

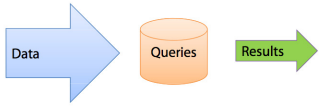


SciSpike www.scispike.com Copyright © SciSpike 2012

Working With Streaming Data: A New Paradigm

- Conventional processing: **static data**

- Real-time processing: **streaming data**



SciSpike www.scispike.com Copyright © SciSpike 2012

Applications for Big Data Analytics

SciSpike www.scispike.com Copyright © SciSpike 2012

Addressing the Key Requirements

1. **Platform for V³ – Variety, Velocity, Volume**
   - Variety - manage data & content "As Is"
   - Handle any velocity - low-latency streams and large volume batch
   - Volume - huge volumes of at-rest or streaming data
2. **Analytics for V³**
   - Analyze Sources in their native format - text, data, rich content
   - Analyze all of the data - not just a subset
   - Dynamic analytics - automatic adjustments and actions
3. **Ease of Use for Developers and Users**
   - Developer UIs, common languages & automatic optimization
   - End-user UIs & visualization
4. **Enterprise Class**
   - Failure tolerance, Security and Privacy
   - Scale Economically
5. **Extensive Integration Capabilities**
   - Integrate wide variety of sources
   - Leverage enterprise integration technologies

SciSpike www.scispike.com Copyright © SciSpike 2012

knoware
The Knowledge Warehouse Limited

Themes

- Big Data
- Governance and Framework Development
- Measurement and Value

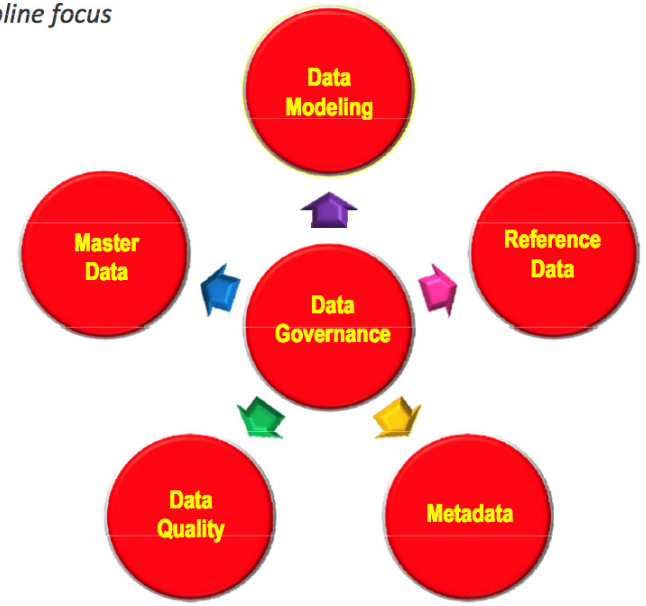
Walgreens: Initiating an Enterprise Information Framework

- Walgreens, USA
- 240, 000 staff
- 784 million prescriptions filled in 2012
- Each store carries ~18,000 different items
- 8000+ stores, plus mail service, home services and health centres
- = a lot of data. Really.

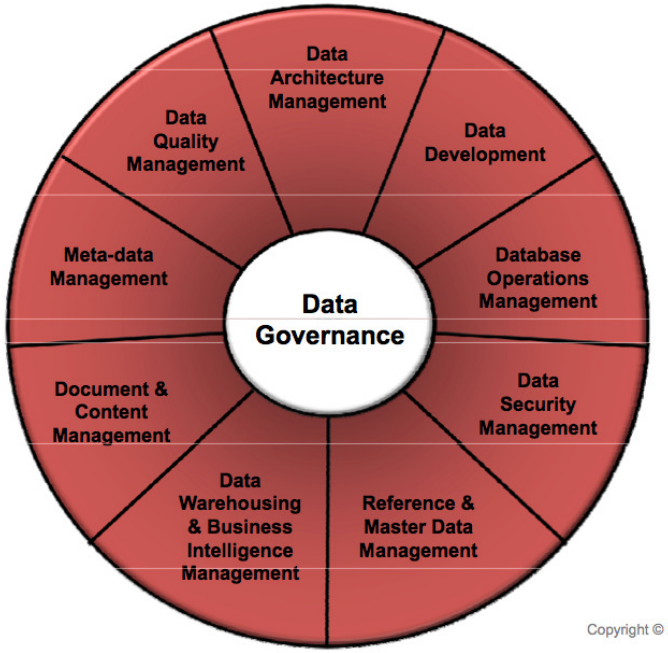
knoware
The Knowledge Warehouse Limited

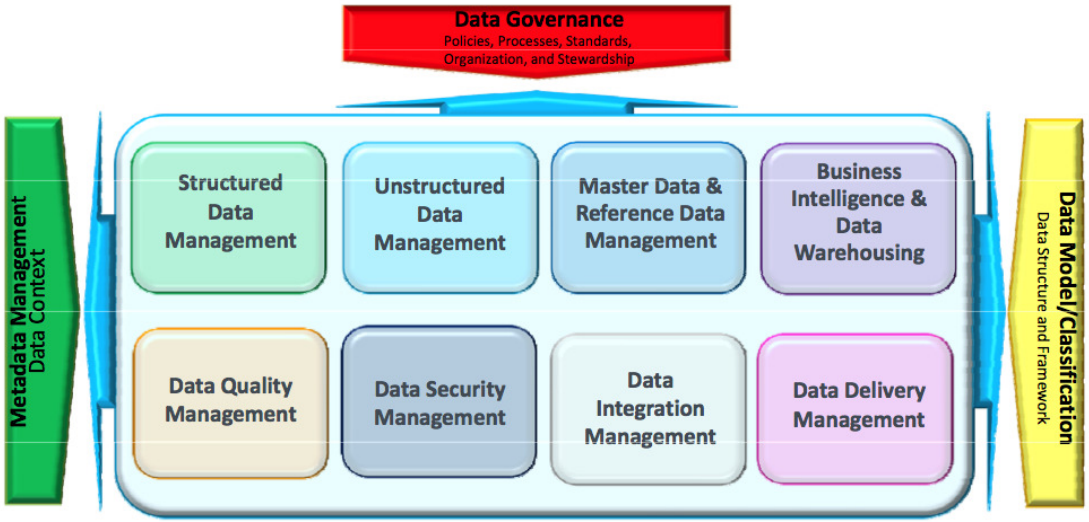
Walgreens: EIM

EIM discipline focus



Context with the DMBOK Framework





knoware
The Knowledge Warehouse Limited

Walgreens: EIM

Enterprise Information Management Framework

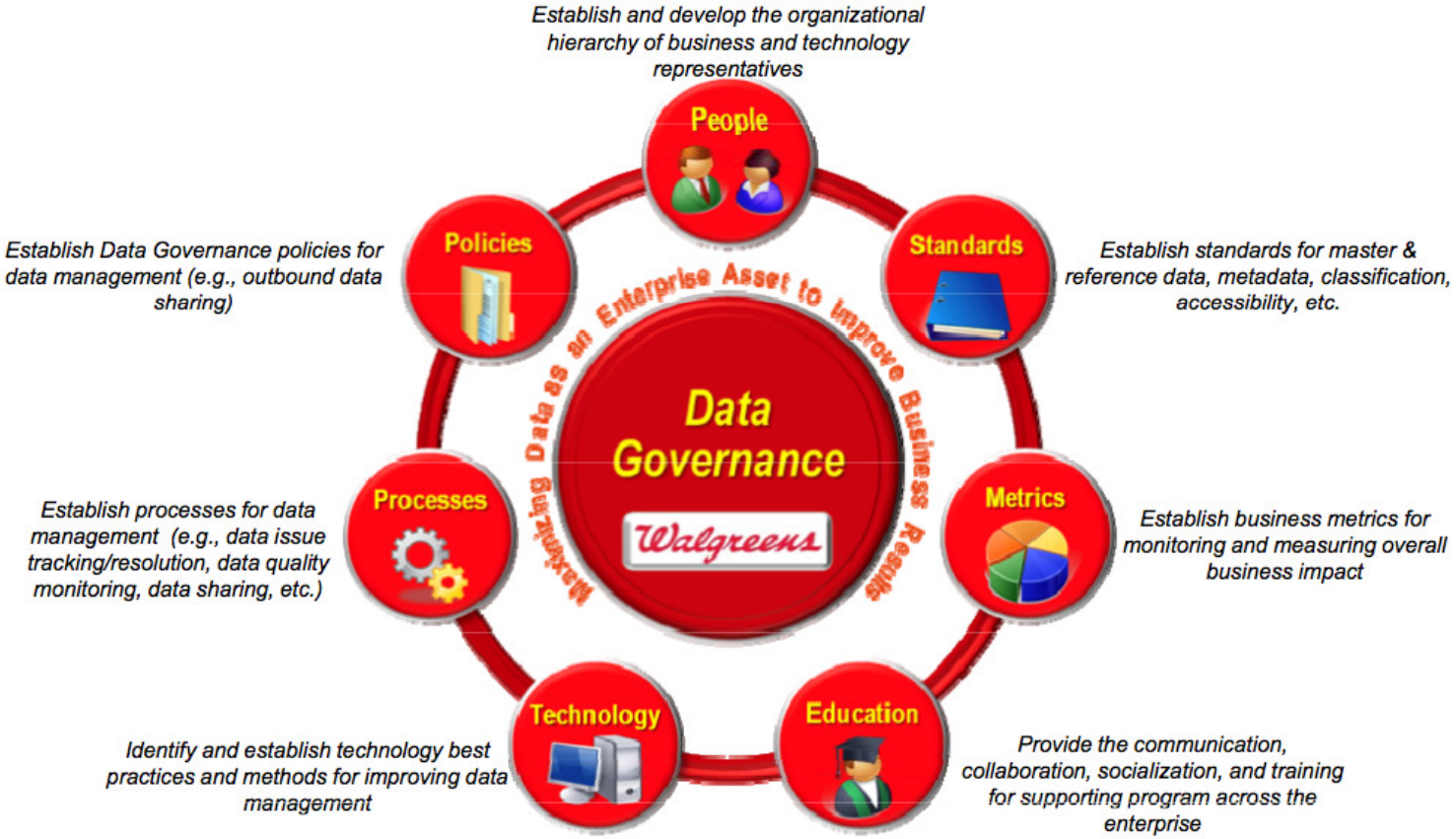
Challenges for EIM Programs

- No two EIM programs are the same. Differences include:
  - Organizational, political, and cultural differences in each organization.
  - Level of business and executive support.
  - Level of enterprise realization, support and seed funding.
    - ***Everyone wants enterprise until they have to pay for it.***
  - Data management maturity level.
- Adopt the best practices to your organization, while aligning to your unique data management state, business strategy and goals.
- Maintain a long-term enterprise focus while implementing incrementally and iteratively.

Walgreens: EIM

Data Governance

Program Components



Walgreens Enterprise Data Governance Logo