

With the Emergence of Big Data, Where do Relational Technologies Fit?

Donna Burbank
Global Data Strategy Ltd.

27 October, 2016



The Business of Data

Donna Burbank



Donna is a recognised industry expert in information management with over 20 years of experience in data strategy, information management, data modeling, metadata management, and enterprise architecture. Her background is multi-faceted across consulting, product development, product management, brand strategy, marketing, and business leadership.

She is currently the Managing Director at Global Data Strategy, Ltd., an international information management consulting

company that specialises in the alignment of business drivers with data-centric technology. In past roles, she has served in key brand strategy and product management roles at CA Technologies and Embarcadero Technologies for several of the leading data management products in the market.

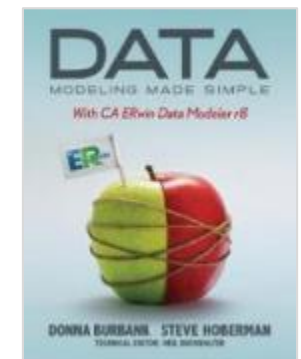
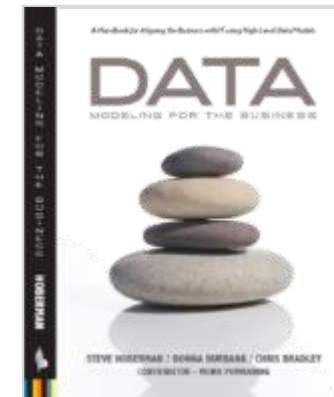
As an active contributor to the data management community, she is a long time DAMA International member and is the President of the DAMA Rocky Mountain chapter. She was also on the review committee for the Object Management Group's Information Management Metamodel (IMM) and a member of the OMG's Finalization Taskforce for the Business Process Modeling Notation (BPMN).

She has worked with dozens of Fortune 500 companies worldwide in the

Americas, Europe, Asia, and Africa and speaks regularly at industry conferences. She has co-authored two books: *Data Modeling for the Business* and *Data Modeling Made Simple with ERwin Data Modeler* and is a regular contributor to industry publications such as DATAVERSITY, EM360, & TDAN. She can be reached at

donna.burbank@globaldatastrategy.com

Donna is based in **Nederland, Colorado, USA.**



Nederland, Colorado

elevation 2700 meters



Agenda

What we'll cover today

- Business opportunities for integrating Big Data with relational technologies
- Big Data vs. relational technologies: a paradigm shift
- How traditional & Big Data technologies can work together
- Case studies & real-world applications
- Roles & organizational considerations
- Summary & questions



Big Data –A Technical & Cultural Paradigm Shift

How can we Transform our Business through Data?

Business Optimization Becoming a *Data-Driven Company*

- Making the Business More Efficient
 - Better Marketing Campaigns
 - Higher quality customer data, 360 view of customer, competitive info, etc.
 - Better Products
 - Data-Driven product development, Customer usage monitoring, etc.
 - Better Customer Support
 - Linking customer data with support logs, network outages, etc.
 - Lower Costs
 - More efficient supply chain
 - Reduced redundancies & manual effort

How do we do
what we do
better?

Business Transformation Becoming a *Data Company*

- Changing the Business Model via Data – data becomes the product
 - Monetization of Information: examples across multiple industries including:
 - *Telecom*: location information, usage & search data, etc.
 - *Retail*: Click-stream data, purchasing patterns
 - *Social Media*: social & family connections, purchasing trends & recommendations, etc.
 - *Energy*: Sensor data, consumer usage patterns, smart metering, etc.

How do we do
something
different?

Technological and Culture Shift in Society



1960

1970

1980

1990

2000

2010



Traditional Relational Technologies and “Big Data”: a Paradigm Shift

Traditional

- Top-Down, Hierarchical
- Design, then Implement
- “Passive”, Push technology
- “Manageable” volumes of information
- “Stable” rate of change
- Business Intelligence

Big Data

- Distributed, Democratic
- Discover and Analyze
- Collaborative, Interactive
- Massive volumes of information
- Rapid and Exponential rate of growth
- Statistical Analysis

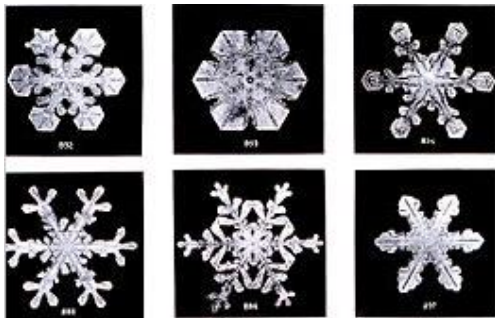
Design ➡ ***Implement***

Discover ➡ ***Analyze***

“New” Way of Looking at the World - *Emergence*

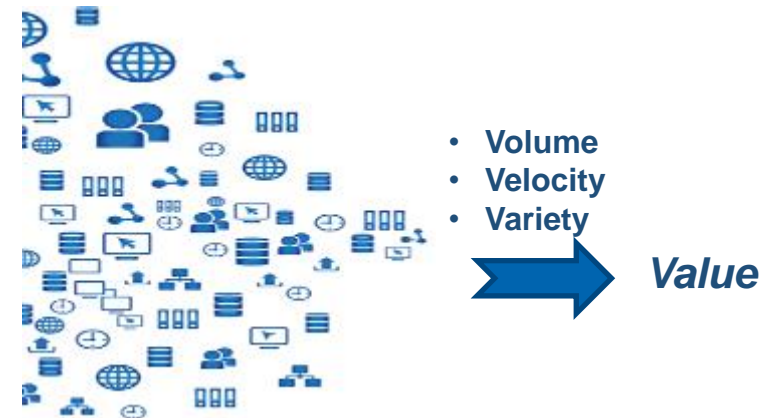
In philosophy, systems theory, science, and art, emergence is the way complex systems and patterns arise out of a multiplicity of relatively simple interactions.

- Wikipedia



What is Big Data?

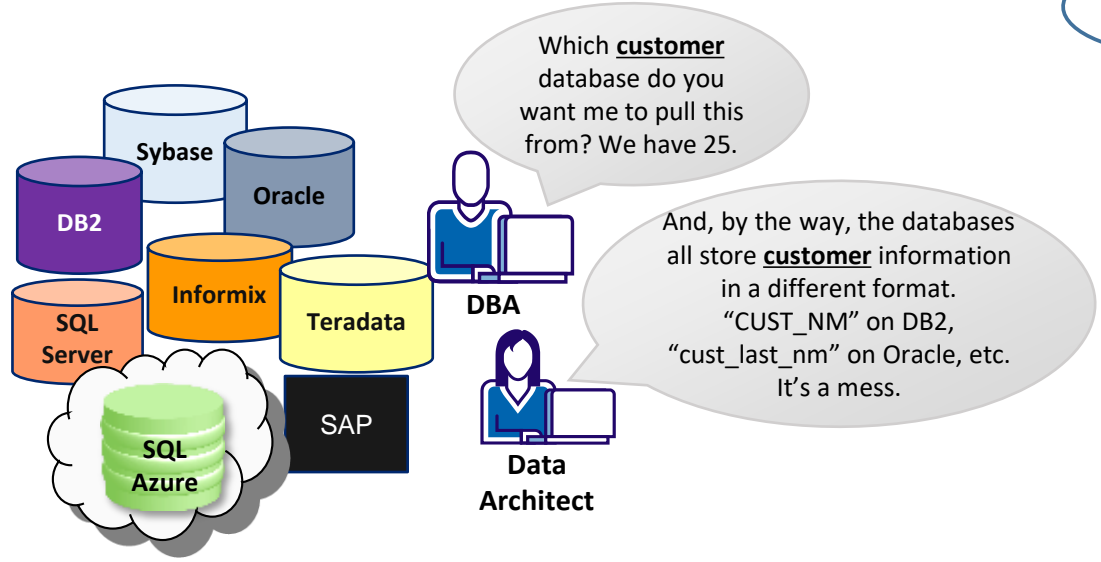
- Big Data is often characterised by the “3 Vs”:
 - **Volume:** Is there a high volume of data? (e.g. terabytes per day)
 - **Velocity:** Is data generated or changed at a rapid pace? (e.g. per second, sub-second)
 - **Variety:** Is data stored across multiple formats? (e.g. machine data, OSS data, log files)
- The ability to understand and manage these sources and integrate them into the larger Business Intelligence ecosystem can provide the ability to gain **valuable insights from data**.
 - **Social Media Sentiment Analysis** – e.g. What are customers saying about our products?
 - **Web Browsing Analytics** – Customer usage patterns
 - **Internet of Things (IoT) Analysis** – e.g. Sensor data, Machine log data
 - **Customer Support** – e.g. Call log analysis
- This ability leads to the “4th V” of Big Data – Value.
 - **Value:** Valuable insights gained from the ability to analyze and discover new patterns and trends from high-volume and/or cross-platform systems.



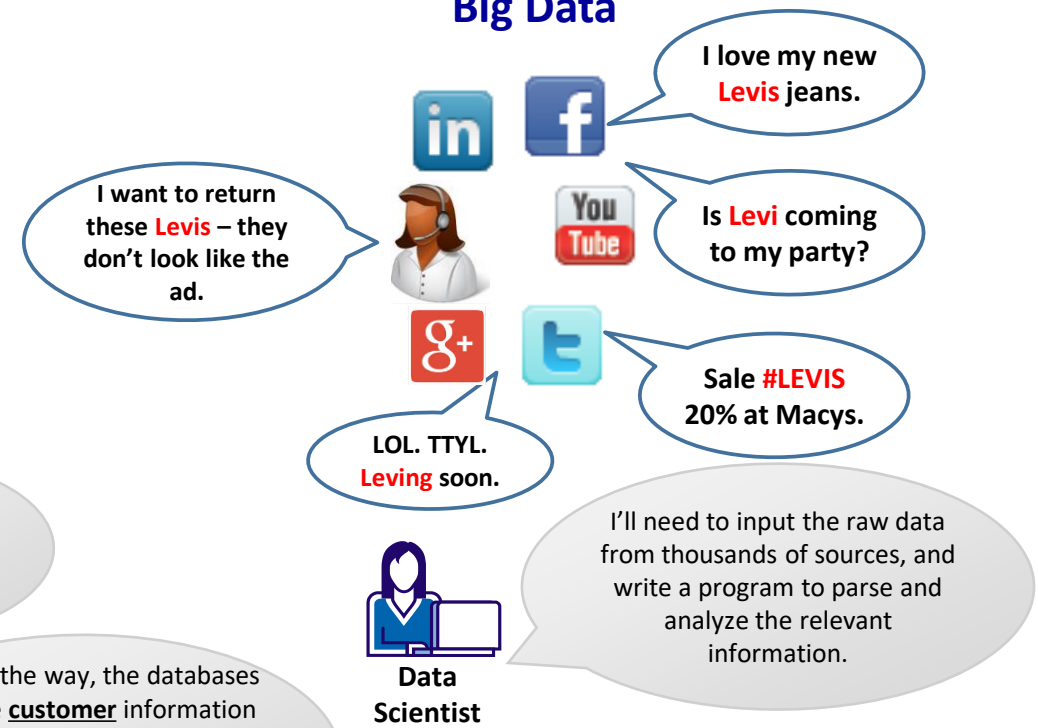
Comprehensive View of Information from All Sources



Traditional Databases

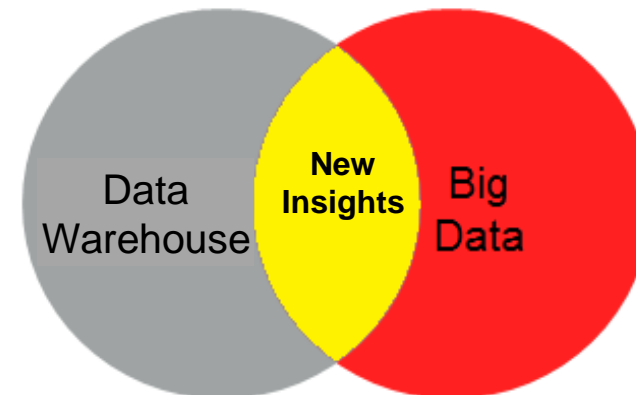


Big Data



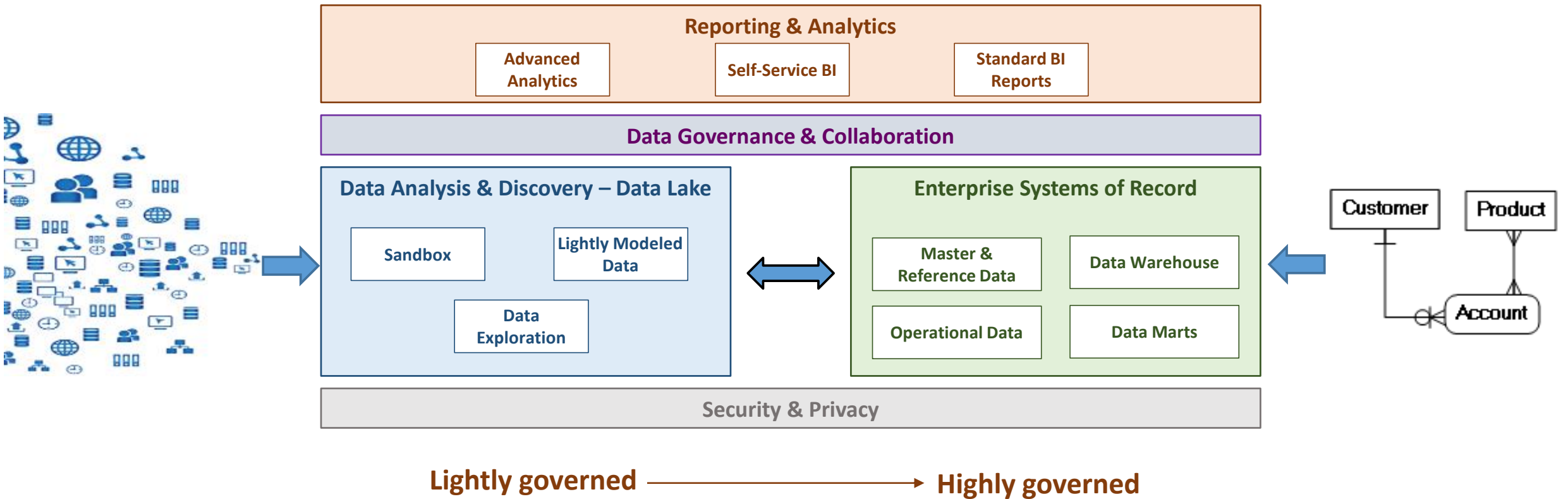
Combining DW & Big Data Can Provide Valuable Information

- There are numerous ways to gain value from data
- Relational Database and Data Warehouse systems are one key source of value
 - Customer information
 - Product information
- Big Data can offer new insights from data
 - From new data sources (e.g. social media, IoT)
 - By correlating multiple new and existing data sources (e.g. network patterns & customer data)
- Integrating DW and Big Data can provide valuable new insights.
- Examples include:
 - Customer Experience Optimization
 - Churn Management
 - Products & Services Innovation



Integrating Big Data & Traditional Data Sources

- The Big Data Lake has a different architecture & purpose than traditional data sources such as data warehouses.
- The two environments can co-exist to share relevant information.



The 5th “V” - Veracity

- Only through proper Governance, Data Quality Management, Metadata Management, etc., can organizations achieve the 5th “V” – Veracity.
 - **Veracity:** Trust in the accuracy, quality and content of the organizations’ information assets.
- i.e. The hard work doesn’t go away with Big Data

Data Science

Raw data used in Self-Service Analytics and BI environments is often so poor that **many data scientists and BI professionals spend an estimated 50 – 90% of their time cleaning and reformatting data** to make it fit for purpose.⁽⁴⁾

Source: DataCenterJournal.com

Data Lakes

The absence of commonly understood and shared metadata and data definitions is cited as one of the main impediments to the success of Data Lakes.

Source: Radiant Advisors

Data Science

Correcting poor data quality is a Data Scientist’s least favorite task, consuming on average 80% of their working day

Source: Forbes 2016

Digitization & Data Quality

71% of interviewees expect digitization to grow their business. But 70% say the biggest barrier is finding the right data; 62% cite inconsistent data

Source: Stibo Systems



Metadata & Big Data Analytics

- Modern advances in data analytics & big data storage provide a wealth of opportunities
 - But the analytics are only as good as the quality of the underlying data
 - Metadata is critical – where did the data come from? What was its intended purpose? What are the units of measure? What is the definition of key terms?
- Good data analysis is based on good data. Good data requires good metadata.



Are Metadata & Data Models Still Relevant?



From [Data Modeling for the Business](#) by
Hoberman, Burbank, Bradley, Technics
Publications, 2009

Case Study: Facebook's Data Warehouse

- Started with Big Data using Hadoop, then saw the need for a traditional Data Warehouse
 - Ken Rudin, director of analytics at Facebook presented keynote at TDWI Chicago in May 2013—video replay available.
- Needed a single source of reference for core business data
 - All data in one place- “managed chaos”
 - Define the core elements of the business, leave the rest alone
 - 100,000s of tables in Hadoop down to a few dozen in Data Warehouse
- What data warehouse was good at:
 - Operational analysis (e.g. how many users logged in by region?)
 - Faster query time (1 minute on DW, over 1 hour on Hadoop)
- What Big Data was good at:
 - Exploratory analysis (e.g. Where are users posting from—how can we infer a location if one is not listed?)

“The genius of AND and the tyranny of OR” – Jim Collins, author of
“Good to Great” -> i.e. Both solutions have their place

Case Study: Facebook's Data Warehouse

- Challenge in both Big Data and Data Warehouse solutions — **Business Definitions & Metadata**
- e.g. How many users logged in yesterday?
 - What do you mean by user?
 - Does user include mobile devices?
 - If a user posted from Spotify, is that a user?
- Sound familiar?



Case Study: Creating the 360 View of Customer

- A NY Financial Institution was looking to create a 360 View of its High Net Worth Customers
- A Big Data Analytics project performed exploratory analysis on external data sources, social media, internal unstructured data, etc.



The Risk of Bad Data

- But when they tried to match this with their Customer database, they had duplicate customers – which one was the high net worth individual?
- Errors in Customer information can cause disastrous results. The “360 View of Customer” can be hampered by:
 - **Duplicate Data:** Are “Martin G. Stykes, Phd” and “Martin Stykes” the same person?
 - **Erroneous Data:** Is Martin *Stikes* a misspelling?
 - **Ambiguous Data:** Is M. Stykes a woman married to Homer Stykes or a man married to Dr. Amber Wentworth?
 - **Poor Relationships between Data :** Does Martin have a good credit rating, or did he recently file for bankruptcy?
 - **Accessibility of Data:** The information about Mr. Styke’s Argentinian coverage can’t be accessed in North America.
 - **Timeliness of Data:** Is Martin Stykes who’s single with no children just an old record for the currently married with children Martin Stykes?
 - **Completeness of Data:** We’re missing information about Mr. Styke’s date of birth.
 - **And the list goes on...**

Family Relationships

- Married to Dr. Amber Wentworth
- Married to Homer Stykes
- Father of James P. Stykes
- Single, No Children

Community Relationships

- Board Member of Schools for Safety

1. Martin G. Stykes, Phd
2. M. Stykes
3. Martha G. Stykes, Phd
4. Martin Stikes
5. M.G. Stykes
6. Martin Stykes
7. Martin Styke
8. Martin G. Stykes
9. Etc.

Commercial Links

- Owner of Krugel Law Firm
- Owner of Mart’s Fish & Chips

SEC & Credit Data

- Excellent Credit Rating
- Filed for Bankruptcy in 2014



High Net Worth Individual?

Internet of Things (IoT)

Connecting Devices through Data

- **What is the IoT?** – The Internet of Things (IoT) is a network of physical devices that are able to share data over a network.
- Tremendous opportunities for:
 - Automation
 - Personalization
 - Interconnected information sharing

Personal Fitness Trackers



Wearable technology ownership expected to double YoY in 2016*

Smart Homes / Smart Meters



Two-thirds of consumers intend to purchase a connected home device by 2019*

Machine Monitoring & Diagnostics



IoT Metadata

- Metadata is necessary to provide context around the readings generated by IoT devices, e.g. units of measure, type of measurement, etc. .



```
<timestamp value='2016-03-07T16:24:30'>  
  <numeric name='Temperature' value='140' unit='F' />  
</timestamp>
```



```
<timestamp value='2016-03-07T16:24:30'>  
  <numeric name='HeartRate' value='140' unit='BPM' />  
</timestamp>
```

140

Is **140** the temperature of my stove or my max heart rate on my run?



Metadata & Big Data Analytics

“Our analysis shows that energy usage with Smart Meters increases by 5% for each percentage point decrease in temperature compared to a 20% increase for traditional thermostat customers.”



Metadata & Big Data Analytics

“Our analysis shows that energy usage with Smart Meters increases by 5% for each percentage point decrease in temperature compared to a 20% increase for traditional thermostat customers.”



- What was the source for the weather data?
- Were readings taking daily, monthly, weekly? Averages or actuals?
- What was the original purpose & format for the readings?
- Were temperatures in Celsius or Fahrenheit?
- Etc.



- Were readings taken by meter readings or billing amounts?
- Were readings taking daily, monthly, weekly? Averages or actuals?
- Were temperatures in Celsius or Fahrenheit?
- Meter readings for were in completely different formats. It took us weeks to standardize them.
- Etc.



- Is Usage by Address, by Individual, or by Household?
- Are households determined by residence or relationships?
- Etc.

Case Study: Consumer Energy Company

Business Transformation via Data



- For the consumer energy sector *Big Data and Smart Meters are transforming the ways of doing business and interacting with customers.*
 - Moving away from traditional data use cases of metering & billing.
 - Smart meters allow customers to be in control of their energy usage.
 - Control over energy usage with connected systems
 - Custom Energy Reports & Usage
 - Smart Billing based on usage times
- *As energy usage declines, data is becoming the true business asset for this energy company.*
 - Monetization of non-personal data is a future consideration.
- While the Big Data Opportunity is crucial, equally important are the traditional data sources
 - New Data Quality Tools in place for operational and DW data
 - Data Governance Program analyzing data in relation to business processes & roles
 - Business-critical data elements identified and definitions created



Data-Driven Business Evolution

Data is a key component for new business opportunities

Traditional Business Model

- Usage-based billing
- Issue-driven customer service

More Efficient Business Model

- More efficient billing
- Faster customer service response
- More consumer information re: energy efficiency, etc.

New Business Model

- Consumer-Driven Smart Metering
- Connected Devices, IoT
- Proactive service monitoring
- Monetization of usage data

Databases



Data Governance



Data Quality



Metadata Management



Big Data



Case Study: Telecom Company – Big Data Transformation

Becoming a “Data Company”

- UK Telecom Company is transforming its business model via Big Data
 - Executive focus was less on telecommunications, which is becoming a commodity
 - ...and more on *Data*, which is seen as a strategic asset.

Customer Value Optimization

- Customer Experience Optimization
- Customer Sentiment Analysis
- Householding & Family Identification
- Service Center Call log monitoring



Data Monetization

- Footfall Analytics & Location
 - City Planning
 - Retail Planning & Customer Patterns
- Location-based Advertising
 - Intelligent targeting

↑
New Business Model

Old Business Model
↓

Product Usage

- Usage patterns
- Click-stream analytics

- Network performance monitoring
- Network usage patterns



Roles & Organizational Considerations

Roles & Culture

DBAs

- Analytical
- Structured
- Project & Task focused
- Cautious – identifies risks
- “Just let me code!”



Data Scientist

- Looks for opportunities
- Likes to explore
- Seen as “modern”
- Seen as “hip” & “sexy”



Data Architects

- Analytical
- Structured
- “Big Picture” focused
- Can be considered “old school”
- “Let me tell you about my data model!”



Business Executive

- Results-Oriented
- Optimistic – Identifies opportunities
- “Big Picture” focused
- “I’m busy.”
- “What’s the business opportunity?”



Big Data Vendors

- It’s magic!
- It’s easy!
- No modeling needed!

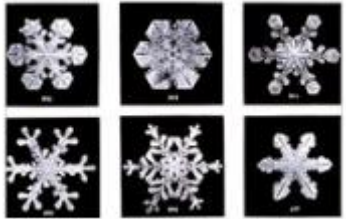


Breaking Down Organizational Siloes

- Good Communication & Governance help break down siloes and encourage information sharing.

Data Lake & Data Scientist

- Exploratory projects
- Quick wins
- Little documentation

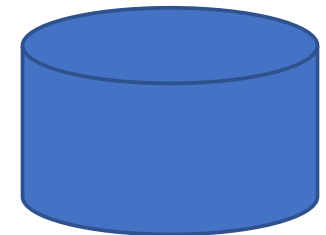


Data Warehouse & Data Architects

- Enterprise reporting
- Long term project
- Data standards & documentation

An illustration of a data warehouse table with columns and rows, featuring a green plant growing from the top right corner.

CARD	LIENNI	REGNUM	AKIN
1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16
17	18	19	20



New Operating Model: Interactions Between New & Existing Roles

Existing Roles



Privacy
Analyst



Data Architect



Data Steward



ETL Developer

Alignment



New Roles

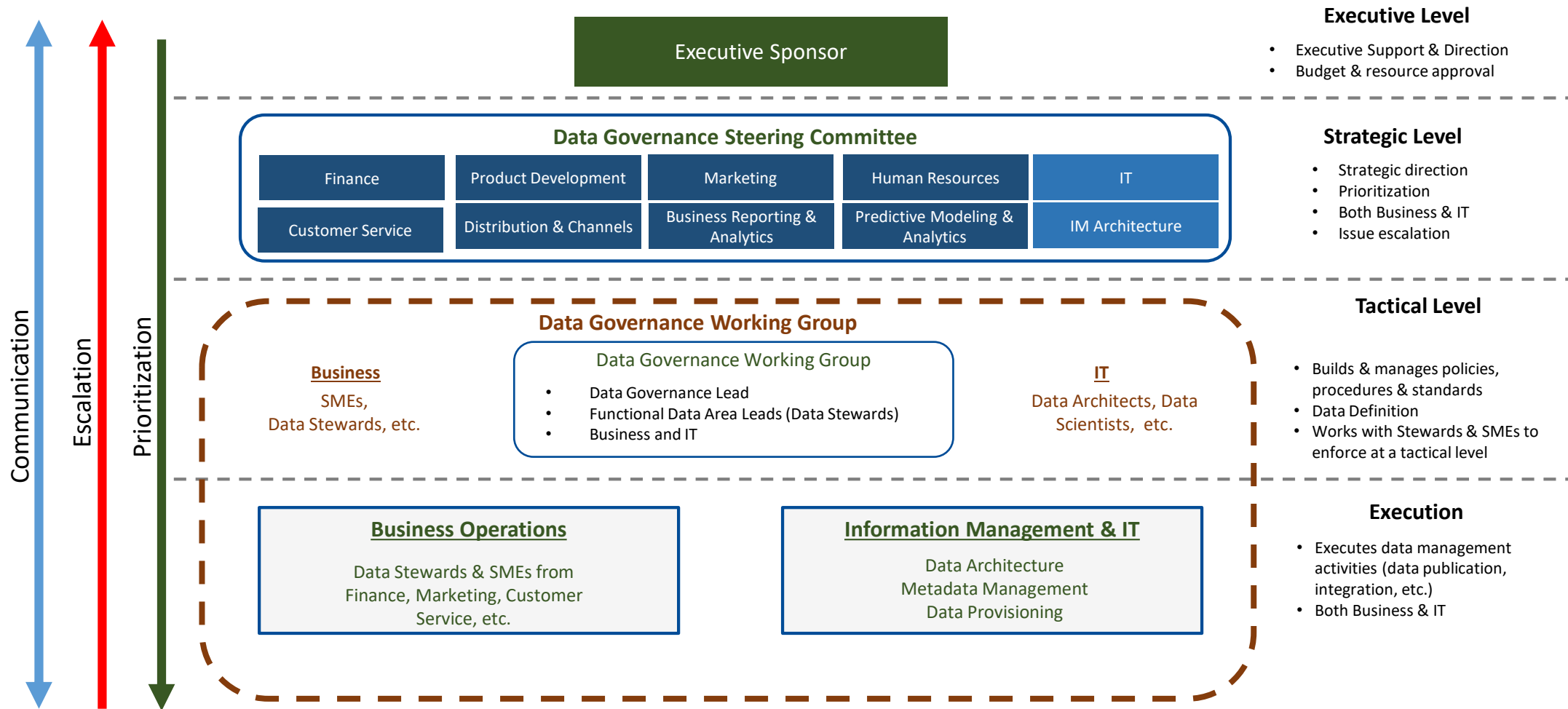


Data Scientist



Hadoop
Administrator

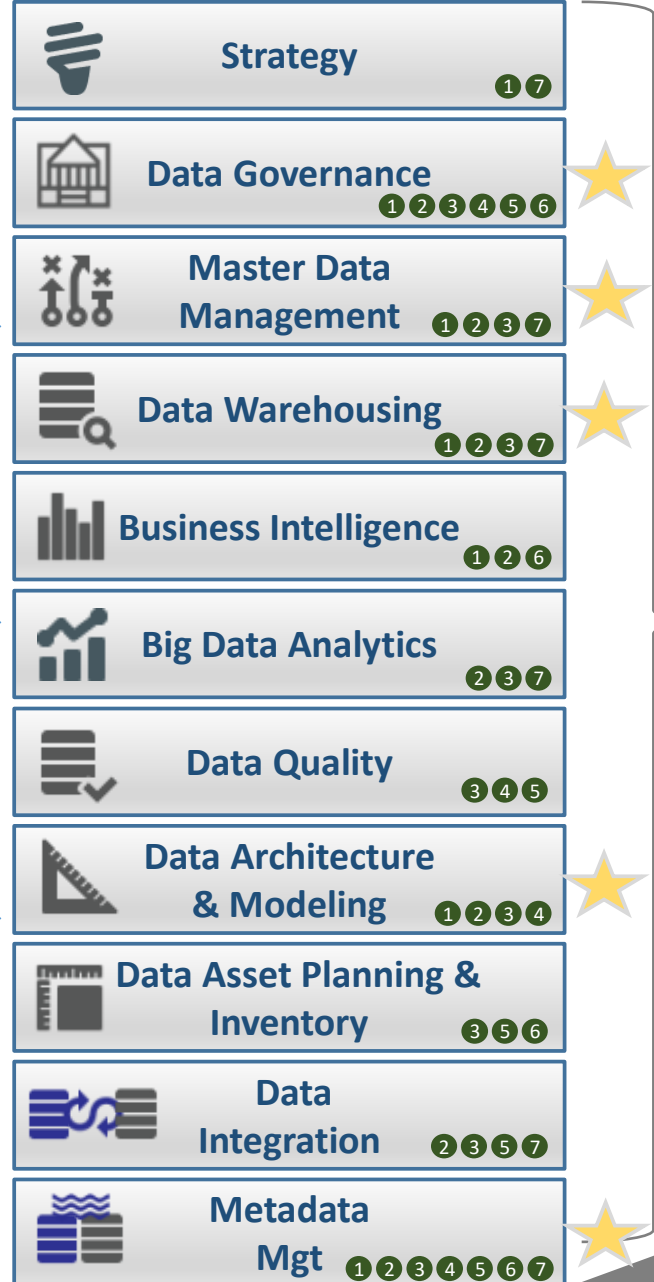
Sample Data Governance Operating Model



Business Value & Implementation

Mapping Business Drivers to Data Management Capabilities

Business-Driven Prioritization

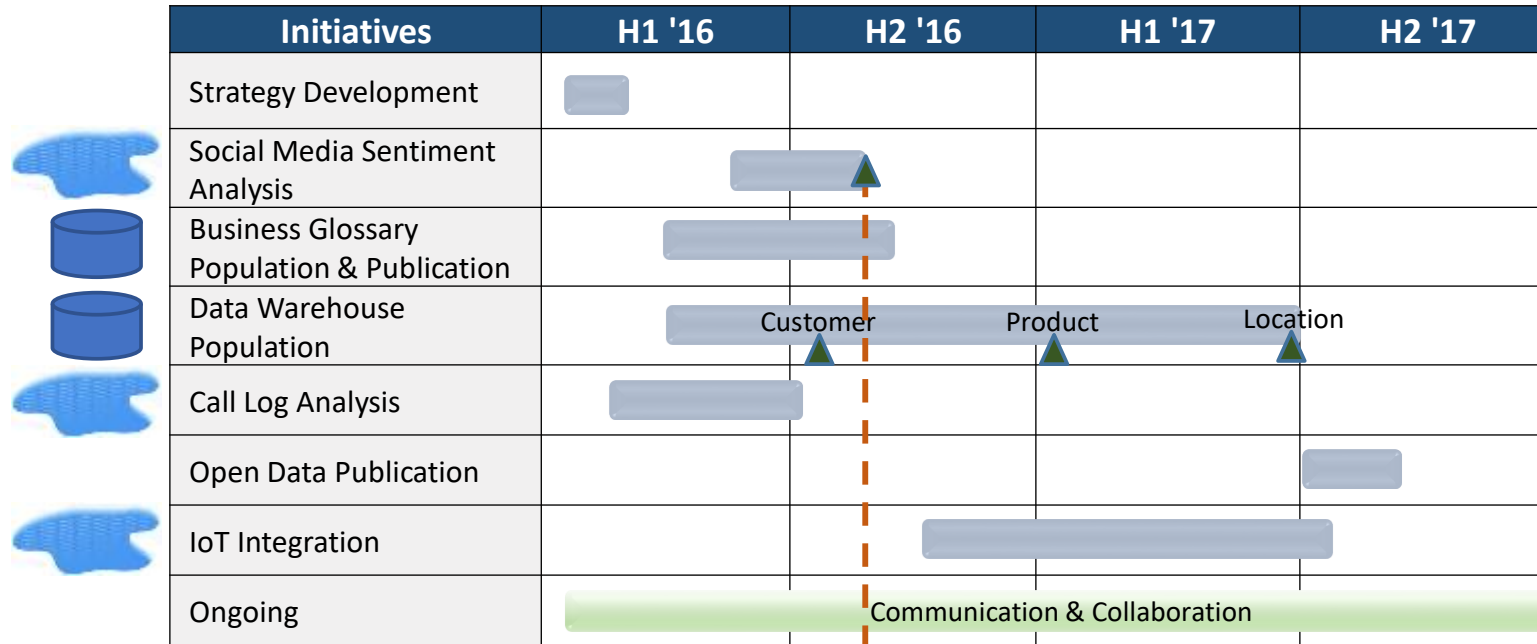


Shows "Heat Map" of Priorities

Defining an Actionable Roadmap

Maximize the Benefit to the Organization

- Develop a detailed roadmap that is both actionable and realistic
 - Show quick-wins, while building to a longer-term goal
 - Balance Business Priorities with Data Management Maturity
 - Include both Big Data exploration & Data Warehouse reporting
 - Focus on projects that benefit multiple stakeholders



Integrated Customer View

- Marketing
- Customer Support
- Sales
- Executive Team

Balance Innovation with Foundation

For Digital Transformation Success, it's important to balance Innovation with Infrastructure:

Digital Innovation

- Big Data
- IoT
- Artificial Intelligence

Foundational Technology

- Master Data Management
- Data Quality
- Architecture & Design

Summary

- We are in a period of “disruptive” technology with new opportunities
 - Rapid rate of change, massive volumes of data
 - Social changes: more participatory, engaged
 - New business models based on data
- Create a fit-for-purpose solution
 - Relational databases are still great for operational systems & data warehouses
 - Big Data offers new opportunities for analysis across large volumes of diverse data
- As with any Age of Change, the basics still apply
 - The “hard stuff” still needs to be done: analysis, metadata definition, data models, etc.
 - Governance and Operating Models are critical
 - Data models are valuable to document business requirements and technical implementation
- Have fun! This is an exciting time to be in Information Management



About Global Data Strategy, Ltd

Data-Driven Business Transformation

- Global Data Strategy is an international information management consulting company that specializes in the alignment of business drivers with data-centric technology.
- Our passion is data, and helping organizations enrich their business opportunities through data and information.
- Our core values center around providing solutions that are:
 - **Business-Driven:** We put the needs of your business first, before we look at any technology solution.
 - **Clear & Relevant:** We provide clear explanations using real-world examples.
 - **Customized & Right-Sized:** Our implementations are based on the unique needs of your organization's size, corporate culture, and geography.
 - **High Quality & Technically Precise:** We pride ourselves in excellence of execution, with years of technical expertise in the industry.

Business Strategy



Aligned With



Data Strategy



Visit www.globaldatastrategy.com for more information

Contact Info

- Email: donna.burbank@globaldatastrategy.com
- Twitter: [@donnaburbank](https://twitter.com/donnaburbank)
[@GlobalDataStrat](https://twitter.com/GlobalDataStrat)
- Website: www.globaldatastrategy.com
- Company LinkedIn: <https://www.linkedin.com/company/global-data-strategy-ltd>
- Personal LinkedIn: <https://www.linkedin.com/in/donnaburbank>

Questions?

Thoughts? Ideas?

