
STATS-DC 2012 Data Conference
July 12, 2012

Washington State's Use of the IBM Data Governance Unified Process Best Practices



Bill Huennekens – Washington State Office of Superintendent of Public Instruction, Bill.Huennekens@k12.wa.us
Eric Naiburg – IBM, ericnaiburg@us.ibm.com

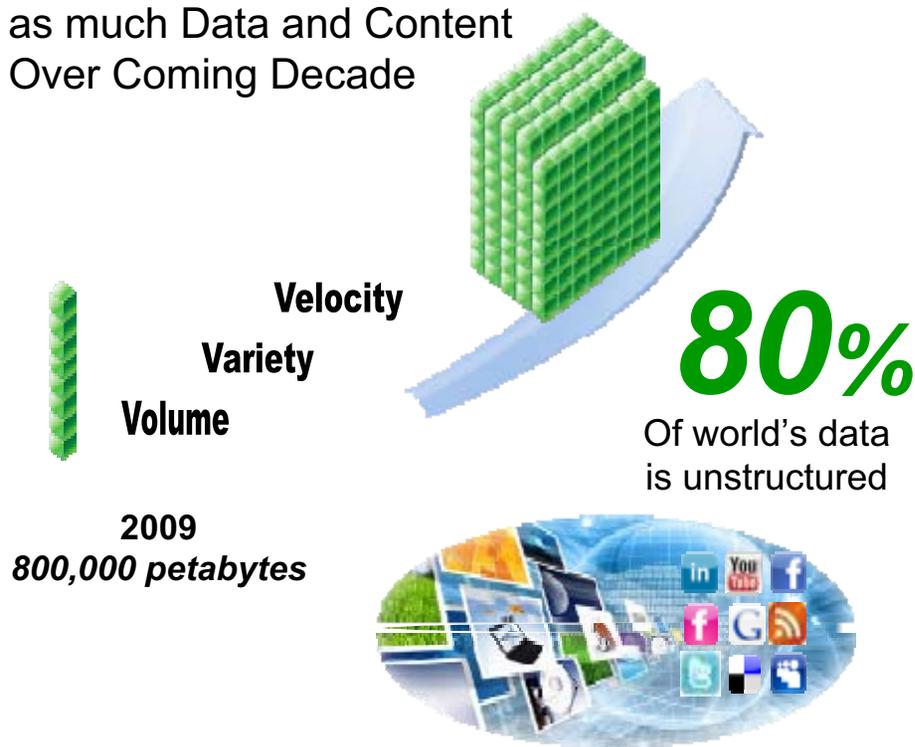
We're all information businesses now

From the state to the district

44x

as much Data and Content
Over Coming Decade

2020
35 zettabytes



... And we need reliable information ...

1 in 3

Leaders frequently make decisions based on information they don't trust, or don't have

1 in 2

Leaders say they don't have access to the information they need to do their jobs

83%

Of CIOs cited "Business intelligence and analytics" as part of their visionary plans to enhance competitiveness

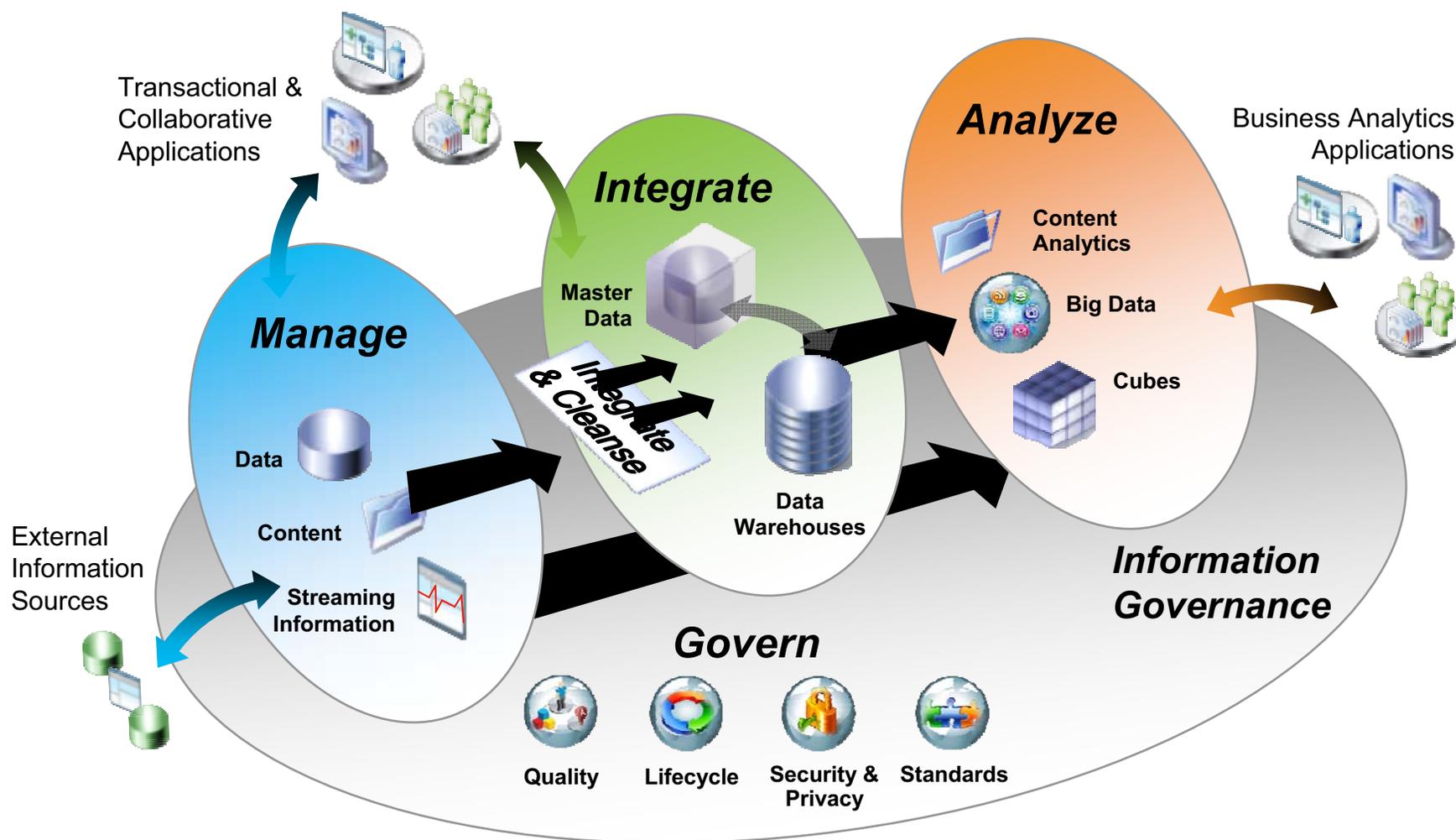
Information is a vital resource to today's education systems ...
But do you treat it as a vital resource?

- Can you deliver information ...
 - Consistently ?
 - Reliably ?
 - Cost effectively ?

- What would happen if the flow of information was disrupted?



Success requires governance across the “Information Supply Chain”



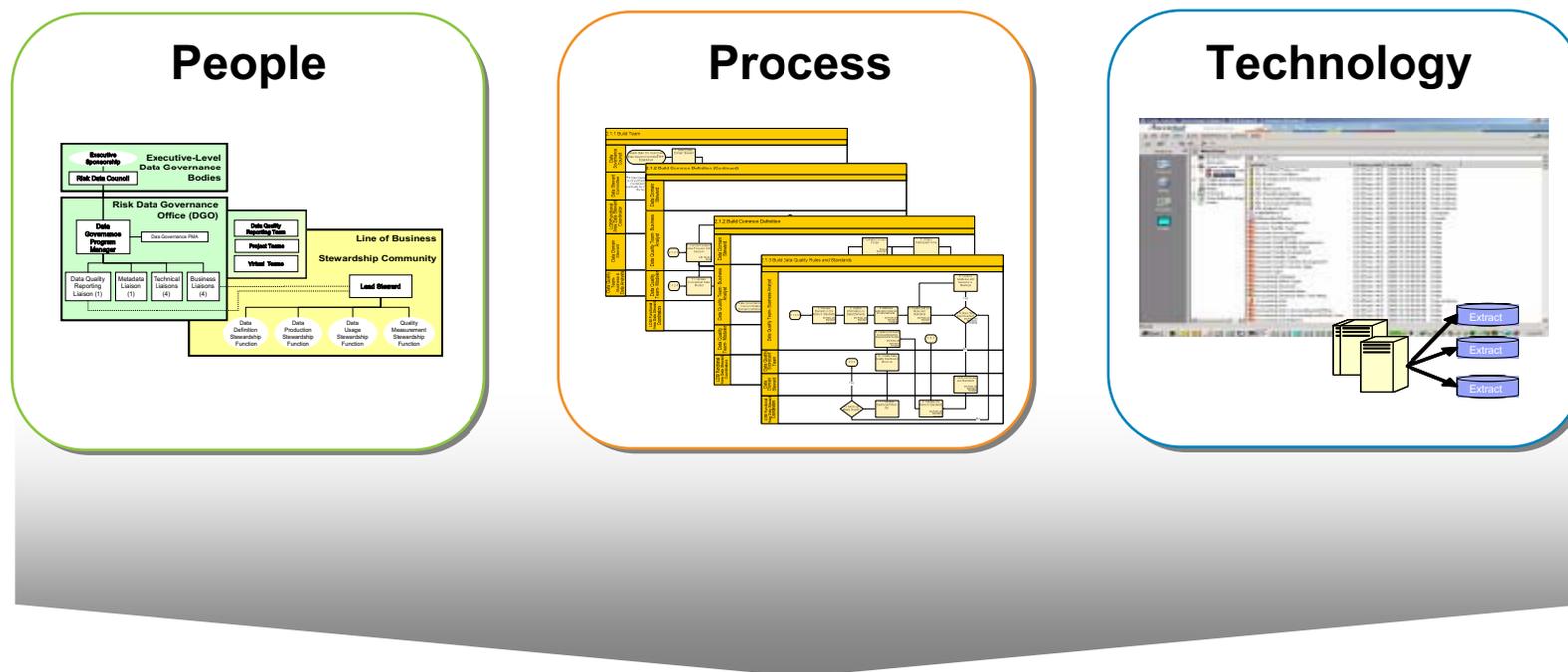
What is Information Governance?



IBM defines Information Governance as a holistic approach to managing and leveraging information for business benefits and encompasses information quality, information protection and information life cycle management

It starts with Information Governance

Proactively leveraging information to unlock value and manage risk



***Ensure information is understood and consistently defined.
Increase the use and trust of information as an enterprise asset.
Protect information, reduce risk and comply.***

... But most organizations continue to struggle

Increasing Complexity



1 trillion
Connected devices in the world

Declining Quality



\$8.2 million
Annual loss by average organization due to poor data quality

Protecting Privacy



\$204
Cost per compromised record

Ensuring Compliance



\$29.8 billion
U.S. spending on governance, risk and compliance

“

CIOs must determine whether your Information Governance strategy adequately reflects the relationship to your overall information management initiative. If the relationship is unclear, or the stated goals are different, work with the business to refactor your strategy.

”

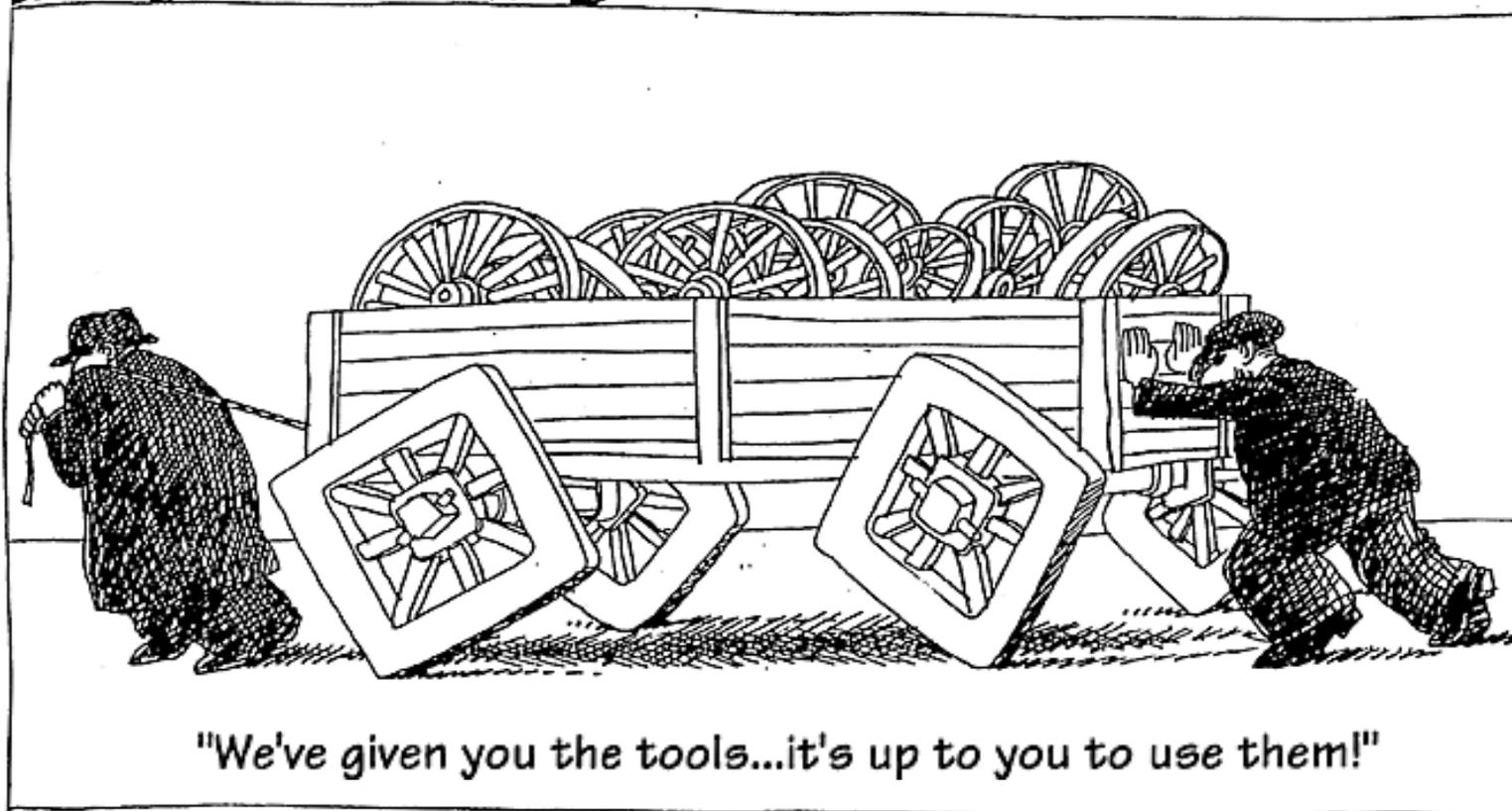
Gartner Research, Q&A: Information Governance, Anne Lapkin, Debra Logan, Jan 19, 2011

Source http://searchcompliance.techtarget.com/news/article/0,289142,sid195_gci1375707,00.html

Photo(s) are stock photos. Release for web use of all photos on file.

We give them tools, but “a fool with a tool is still a fool”

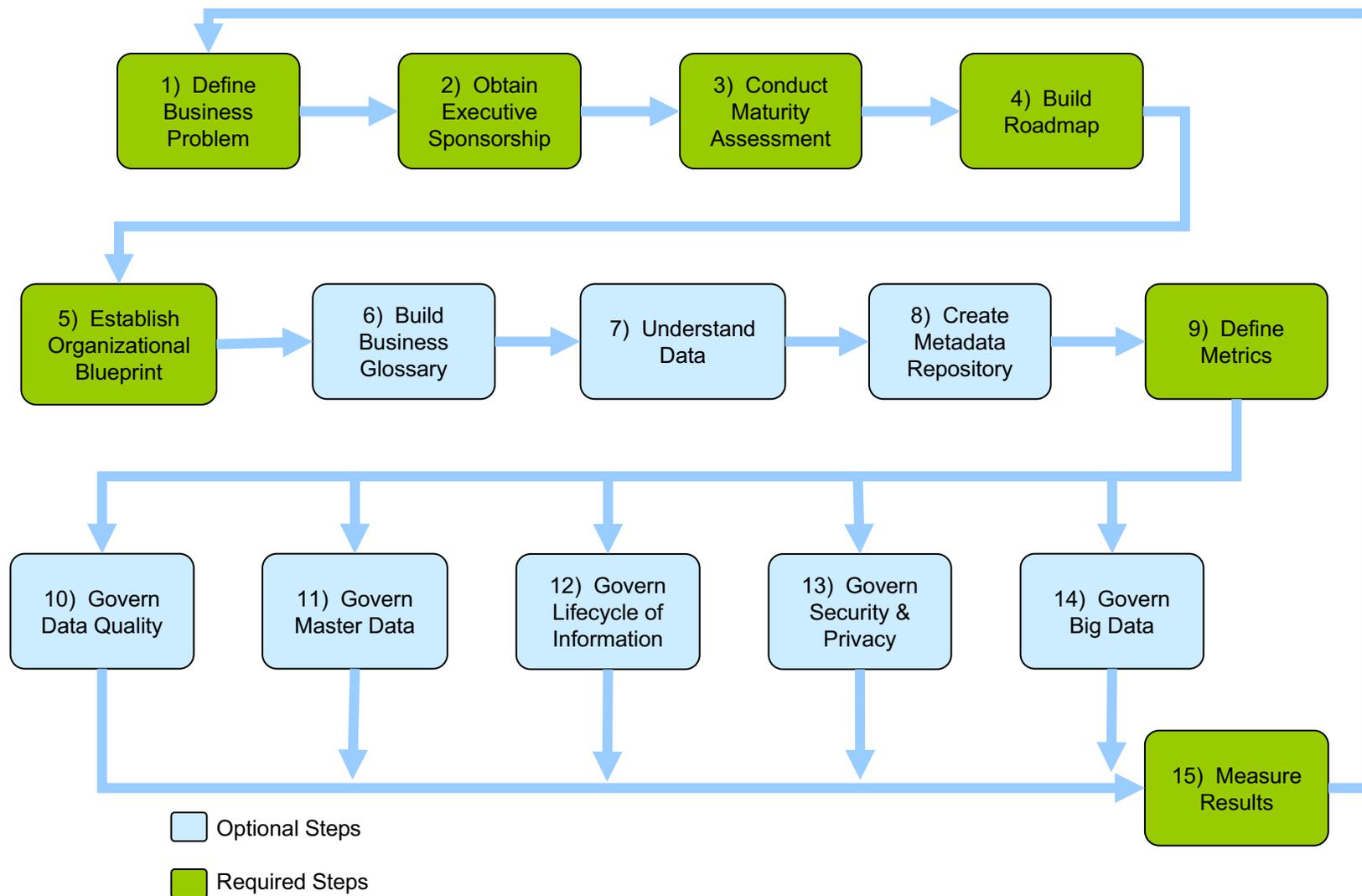
Deeper Meanings



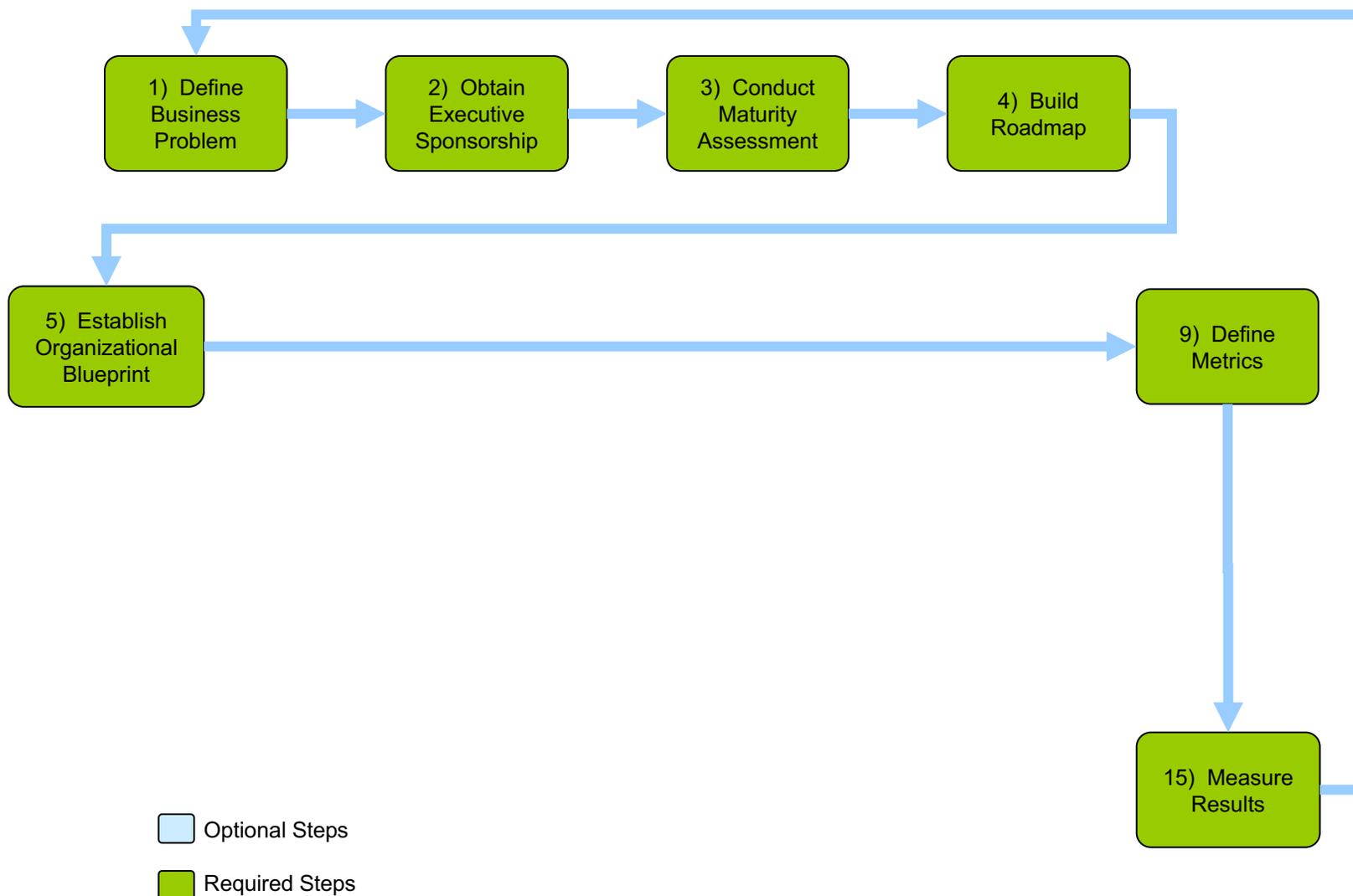
© Deeper Meanings deapermeanings.com

Good governance requires process and accountability

IBM Information Governance Unified Process



We start with the processes



Define Business Problem

The forward of the Implementation Guidelines for K12 Data Governance System outlines the business problem:

“The essential notion behind establishing a data governance system is that decisions are only as good as the data on which they are based. As OPSI transforms data into information to facilitate wise decisionmaking, users and managers of K-12 data need a clear understanding of data definitions, data and process ownership and authority, accountability, security, and reporting needs and requirements, as well as the processes and timelines around each.”

See - <http://www.k12.wa.us/K12DataGovernance/pubdocs/DataGovernanceManual.pdf>, page 6.

Obtain Executive Sponsorship

- ESHB 2261 passed during the 2009 legislative session established the K12 Data Governance Program. This portion of the legislation was supported by OSPI leadership and the implementation of the program has been supported by the leadership.

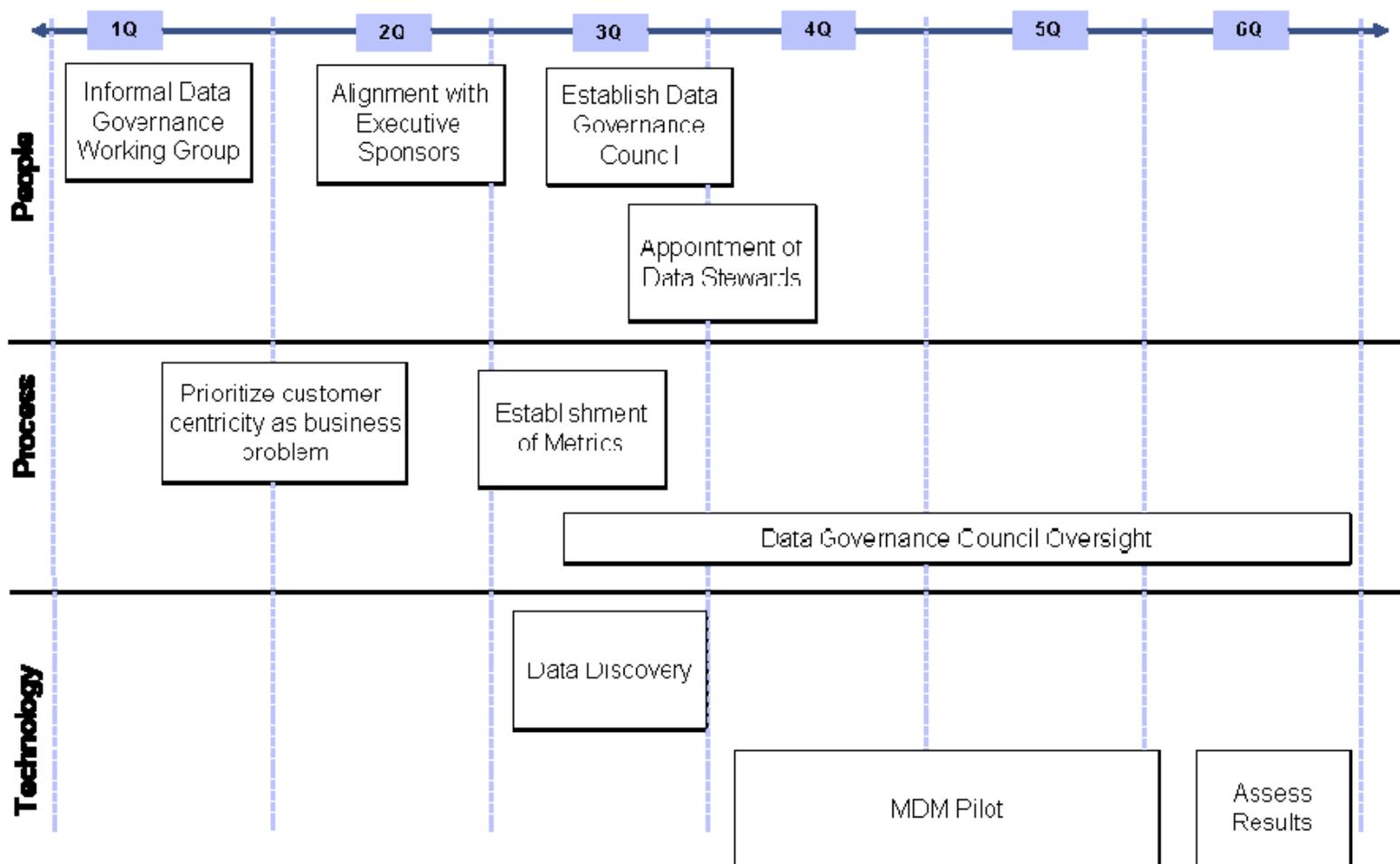
Conduct Maturity Assessment

		Section	Your Score	Desired Score	# Taken By Community	Community Average	Community Median	Take Snapshot
	Re-Take	Org Awareness & Structure	2	—	90	1.5	1.3	Clear
	Re-Take	Stewardship	2	—	77	1.8	1.5	Clear
	Re-Take	Policy	2	—	69	1.6	1.7	Clear
	Re-Take	Data Risk Management	1	—	68	1.9	1.8	Clear
	Re-Take	Value Creation	2	—	64	1.8	1.6	Clear
	Re-Take	Data Quality	2	—	83	1.9	1.7	Clear
	Re-Take	ILM	2	—	58	1.9	1.8	Clear
	Re-Take	Security	2	—	56	2.3	2.2	Clear
	Re-Take	Data Architecture	2	—	121	2.5	2.3	Clear
	Re-Take	Metadata	2	—	74	1.7	1.5	Clear
	Take	Audit	—	—	73	1.9	1.7	

Build Road Map

- This is work our agency needs to engage in now that we have documented our “as is” state and the work already underway.

Build a roadmap



Establish Organizational Blue Print

Washington State K12 Data Governance Organizational Blue Print			
Top Tier	<table border="1"> <tr> <td> <p>Data Governance Group (DGG) <i>Lead:</i> Data Governance Coordinator <i>Members:</i> http://k12.wa.us/K12DataGovernance/Members.aspx <i>Meetings:</i> Approximately six times a year, see website for details <i>Purpose:</i> See section 2 of the Data Governance Manual beginning on page 8.</p> </td> <td> <p>Data Leadership Team <i>Lead:</i> Data Governance Coordinator <i>Members:</i> CIO, Director of Student Information, Director of School Apportionment and Financial Services, Associate Director of Educator Certification <i>Meetings:</i> every 2-3 weeks as scheduled. <i>Purpose:</i> Advise on agenda and direction DGG and DMC and communicate to OSPi cabinet, legislative leadership and other government entities.</p> </td> </tr> </table>	<p>Data Governance Group (DGG) <i>Lead:</i> Data Governance Coordinator <i>Members:</i> http://k12.wa.us/K12DataGovernance/Members.aspx <i>Meetings:</i> Approximately six times a year, see website for details <i>Purpose:</i> See section 2 of the Data Governance Manual beginning on page 8.</p>	<p>Data Leadership Team <i>Lead:</i> Data Governance Coordinator <i>Members:</i> CIO, Director of Student Information, Director of School Apportionment and Financial Services, Associate Director of Educator Certification <i>Meetings:</i> every 2-3 weeks as scheduled. <i>Purpose:</i> Advise on agenda and direction DGG and DMC and communicate to OSPi cabinet, legislative leadership and other government entities.</p>
<p>Data Governance Group (DGG) <i>Lead:</i> Data Governance Coordinator <i>Members:</i> http://k12.wa.us/K12DataGovernance/Members.aspx <i>Meetings:</i> Approximately six times a year, see website for details <i>Purpose:</i> See section 2 of the Data Governance Manual beginning on page 8.</p>	<p>Data Leadership Team <i>Lead:</i> Data Governance Coordinator <i>Members:</i> CIO, Director of Student Information, Director of School Apportionment and Financial Services, Associate Director of Educator Certification <i>Meetings:</i> every 2-3 weeks as scheduled. <i>Purpose:</i> Advise on agenda and direction DGG and DMC and communicate to OSPi cabinet, legislative leadership and other government entities.</p>		
Middle Tier	<table border="1"> <tr> <td> <p>Data Management Committee <i>Lead:</i> Data Governance Coordinator <i>Members:</i> See section 4 of the Data Governance Manual beginning on page 17. <i>Meetings:</i> Monthly as Scheduled <i>Purpose:</i> See page 18 of the Data Governance Manual.</p> </td> </tr> </table>	<p>Data Management Committee <i>Lead:</i> Data Governance Coordinator <i>Members:</i> See section 4 of the Data Governance Manual beginning on page 17. <i>Meetings:</i> Monthly as Scheduled <i>Purpose:</i> See page 18 of the Data Governance Manual.</p>	
<p>Data Management Committee <i>Lead:</i> Data Governance Coordinator <i>Members:</i> See section 4 of the Data Governance Manual beginning on page 17. <i>Meetings:</i> Monthly as Scheduled <i>Purpose:</i> See page 18 of the Data Governance Manual.</p>			
Final Tier	<table border="1"> <tr> <td> <p>CEDARS Weekly Business Meeting <i>Lead:</i> Director of Application Development <i>Members:</i> Student data owner and stewards, educator data stewards, application development director, help desk staff and other data stewards and IT staff as needed. <i>Meetings:</i> Every Tuesday 10-12 <i>Purpose:</i> Work to resolve operation issues with the CEDARS data collection system.</p> </td> </tr> </table>	<p>CEDARS Weekly Business Meeting <i>Lead:</i> Director of Application Development <i>Members:</i> Student data owner and stewards, educator data stewards, application development director, help desk staff and other data stewards and IT staff as needed. <i>Meetings:</i> Every Tuesday 10-12 <i>Purpose:</i> Work to resolve operation issues with the CEDARS data collection system.</p>	
<p>CEDARS Weekly Business Meeting <i>Lead:</i> Director of Application Development <i>Members:</i> Student data owner and stewards, educator data stewards, application development director, help desk staff and other data stewards and IT staff as needed. <i>Meetings:</i> Every Tuesday 10-12 <i>Purpose:</i> Work to resolve operation issues with the CEDARS data collection system.</p>			
Stakeholder Tier	<table border="1"> <tr> <td> <p>Dual Credit Report Team <i>Lead:</i> Director of SI, Secondary ED Supervisor <i>Members:</i> OSPi staff and various post secondary partners <i>Meetings:</i> 2/3 times a year <i>Purpose:</i> Discuss business rules, timelines, issues, questions, and concerns regarding the acquisition and analysis of dual credit participation data from multiple sources (SBCTC, COP, HECB, ERDC, & CEDARS)</p> </td> <td> <p>Graduation a Team Effort Data Subcommittee <i>Lead:</i> Data Governance Coordinator and LEA Colleague <i>Members:</i> Various SEA and LEA Staff <i>Meetings:</i> As needed. <i>Purpose:</i> Make recommendations and seek solution for using data to assist with the graduation goals.</p> </td> </tr> </table>	<p>Dual Credit Report Team <i>Lead:</i> Director of SI, Secondary ED Supervisor <i>Members:</i> OSPi staff and various post secondary partners <i>Meetings:</i> 2/3 times a year <i>Purpose:</i> Discuss business rules, timelines, issues, questions, and concerns regarding the acquisition and analysis of dual credit participation data from multiple sources (SBCTC, COP, HECB, ERDC, & CEDARS)</p>	<p>Graduation a Team Effort Data Subcommittee <i>Lead:</i> Data Governance Coordinator and LEA Colleague <i>Members:</i> Various SEA and LEA Staff <i>Meetings:</i> As needed. <i>Purpose:</i> Make recommendations and seek solution for using data to assist with the graduation goals.</p>
<p>Dual Credit Report Team <i>Lead:</i> Director of SI, Secondary ED Supervisor <i>Members:</i> OSPi staff and various post secondary partners <i>Meetings:</i> 2/3 times a year <i>Purpose:</i> Discuss business rules, timelines, issues, questions, and concerns regarding the acquisition and analysis of dual credit participation data from multiple sources (SBCTC, COP, HECB, ERDC, & CEDARS)</p>	<p>Graduation a Team Effort Data Subcommittee <i>Lead:</i> Data Governance Coordinator and LEA Colleague <i>Members:</i> Various SEA and LEA Staff <i>Meetings:</i> As needed. <i>Purpose:</i> Make recommendations and seek solution for using data to assist with the graduation goals.</p>		

“

Remember that when Information Governance roles are created, they should be explicitly made part of the job, and the individuals assigned to them must be accountable for their performance in those roles. It's not good enough to say 'we know you have a day job, but we'd like you to be responsible for this information as well.'

”

Gartner Research, Q&A: Information Governance, Anne Lapkin, Debra Logan, Jan 19, 2011

Establish Organizational Blue Print

Data Management Committee Facilitator\Coordinator: Data Governance Coordinator		
<i>Program Area</i>	<i>Data Owners</i>	<i>Data Stewards</i>
Overall Student Data	Director of Student Information	
Student Demographics		Student Information Coordinator
Assessment		Student Information Coordinator
Graduation and Dropout Status		Research Analyst, Student Information
Learning and Teaching		Learning and Technology Program Director
Early Childhood		Early Learning Project Coordinator
Special Education		Special Education Data Management Coordinator

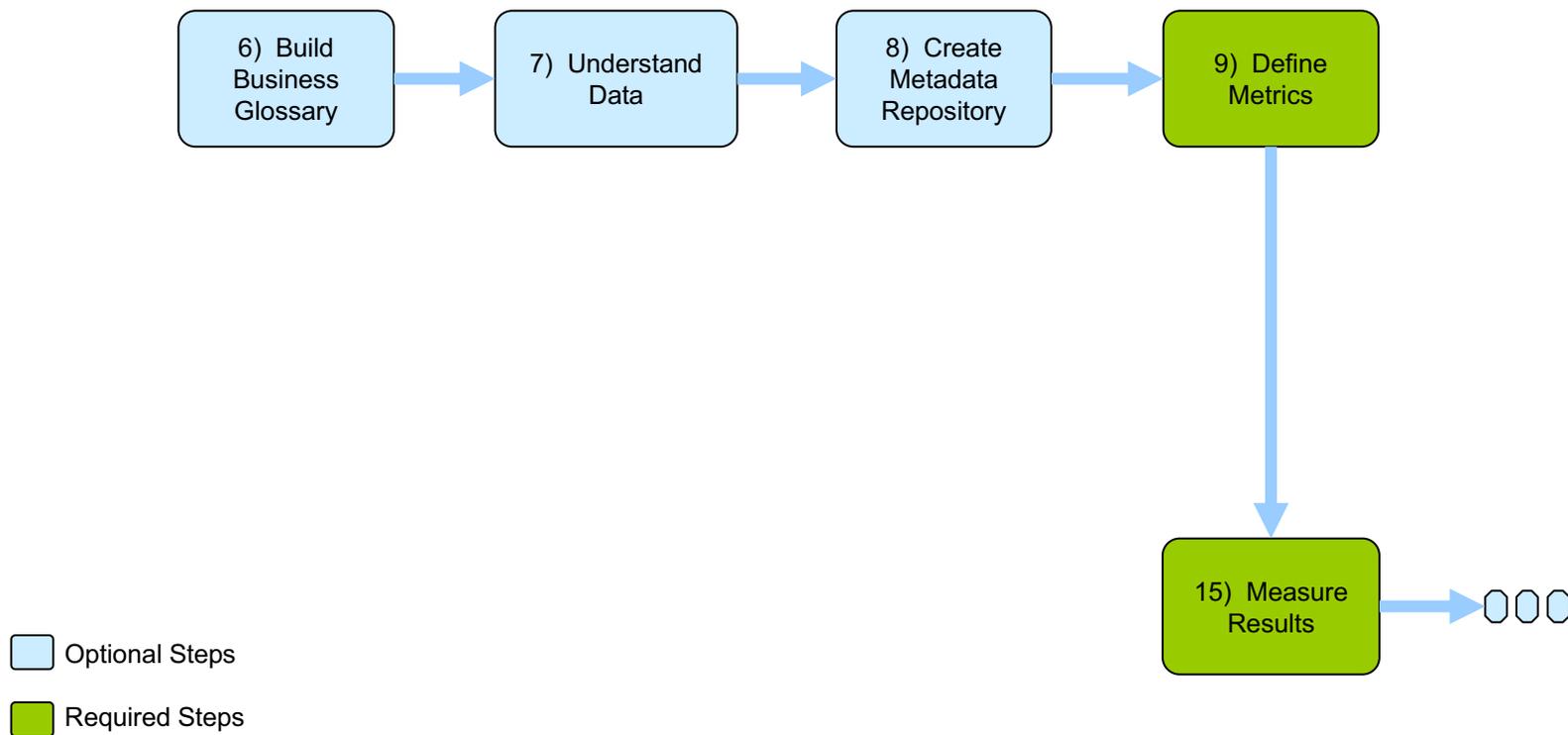
Define Metrics and Measure Results

- This is our next step.
- Following slide is the recommendation from IBM and the workshop they conducted with us.

Getting Started with Data Governance Metrics

- Metrics can be developed across the software development lifecycle and across data collection, data integration, and data dissemination activities
 - Data Collection
 - Data enters the SLDS through source systems such as CEDARS, eCert, the Certificate database, and financial systems
 - Analyzing the sources will provide assessment of baseline data quality, help scope effort needed to integrate data in the Longitudinal Data Warehouse, and identify data issues that can be fixed at the source (system of record)
 - Data Integration
 - Metrics should be identified to track delivery of data from school districts, measure pass/fail of data records validation, and track issues by district that indicate communications and training needs
 - The profiling tool can assess the integrity of the data warehouse, ensure primary and foreign keys exist for data linkage, identify missing values, and measure business rules adherence
 - Data Dissemination
 - Measure the usage / consumption of information from the data warehouse
 - Track calls into the support desk by data domain, type of issue, and data quality impacts

Put the right standards in place



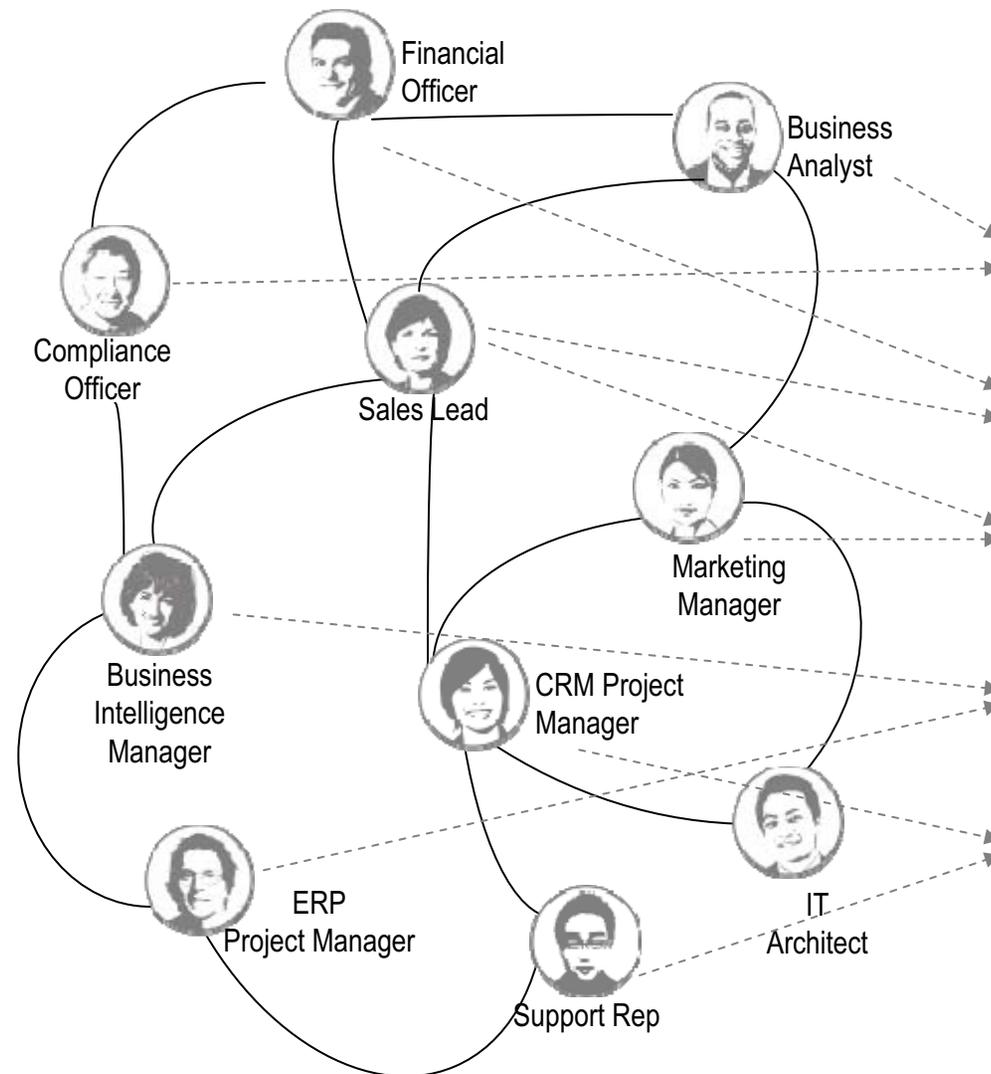
Miscommunication happens



©2002 Bill Keene, Inc. *The Family Circus*®
From the *Denver Post* 5/17/02

Photo(s) are stock photos. Release for web use of all photos on file.

Define a common vocabulary

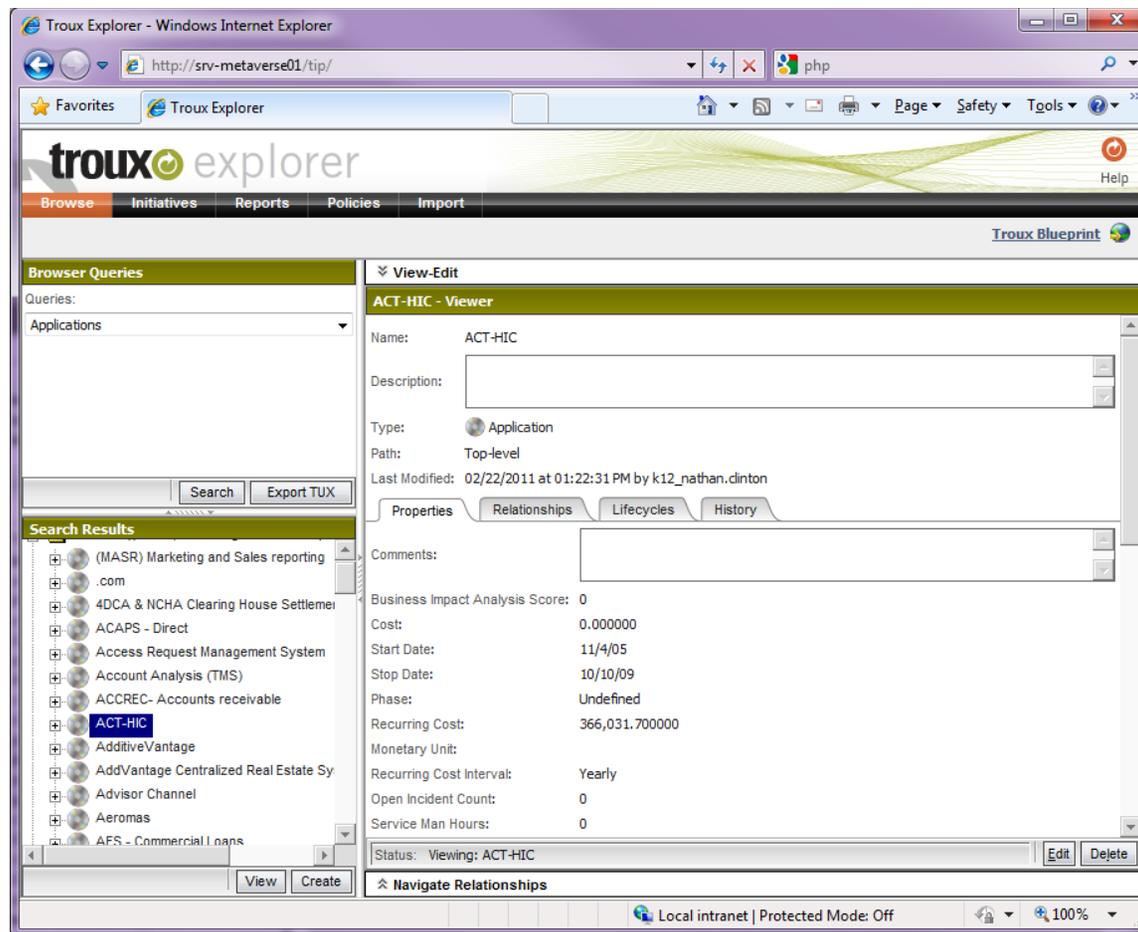


For example, define

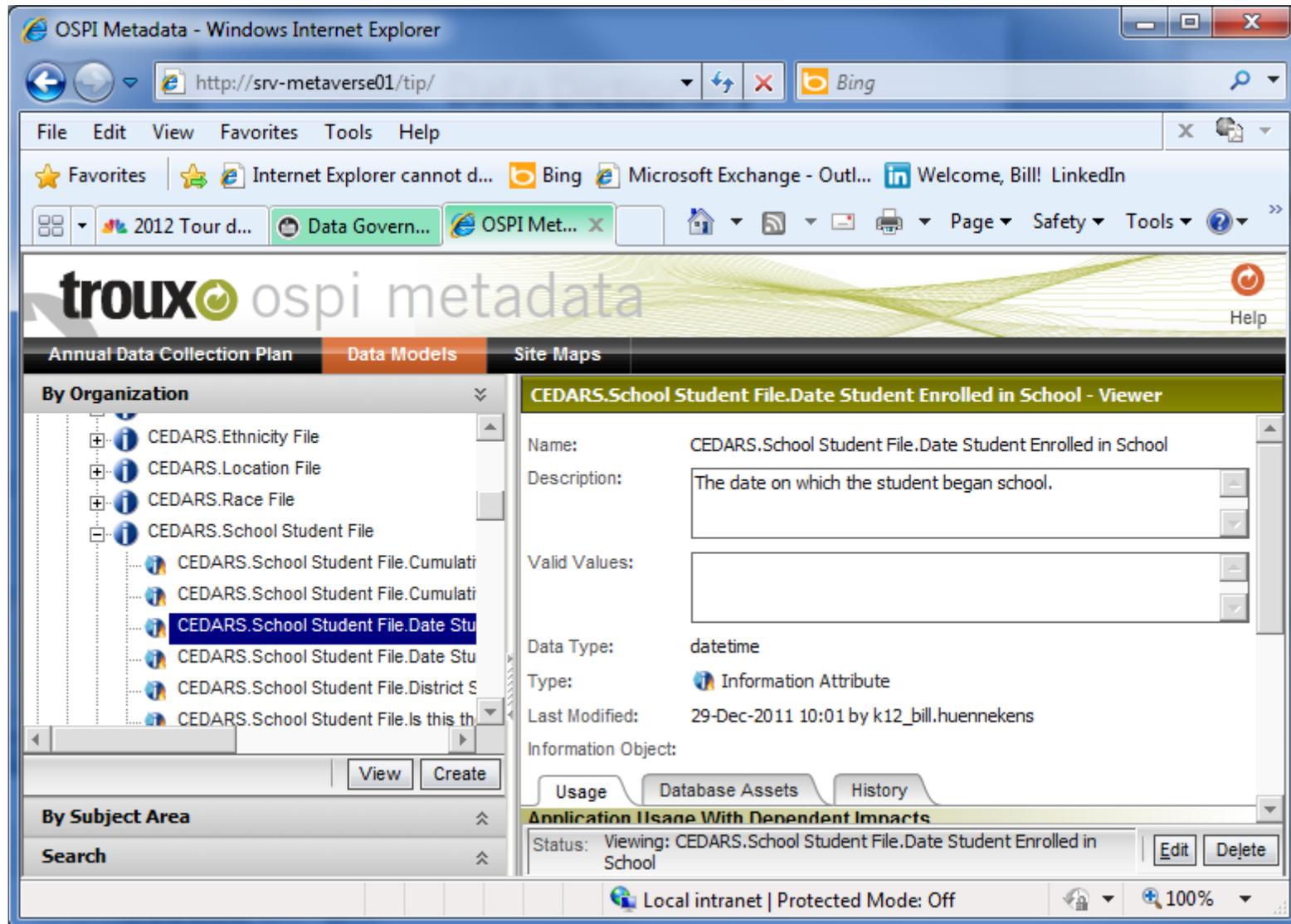
“Active Subscriber”

- **Mobile user who has used “any” service in the mobile network**
- **User who paid for the service at least 1 time in the past 90 days.**
- **Mobile user who has a phone plan, but not SMS**
- **Only post-paid customers, not pre-paid customers**
- **User who makes at least 1 call over the period of 90 days**

Using Enterprise Architecture and Metadata tool for data dictionary, understanding the data and as a metadata repository

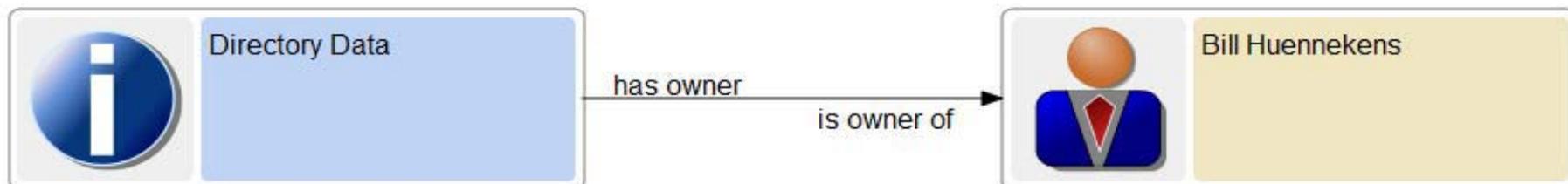


Data Dictionary\Metadata – Understand Data

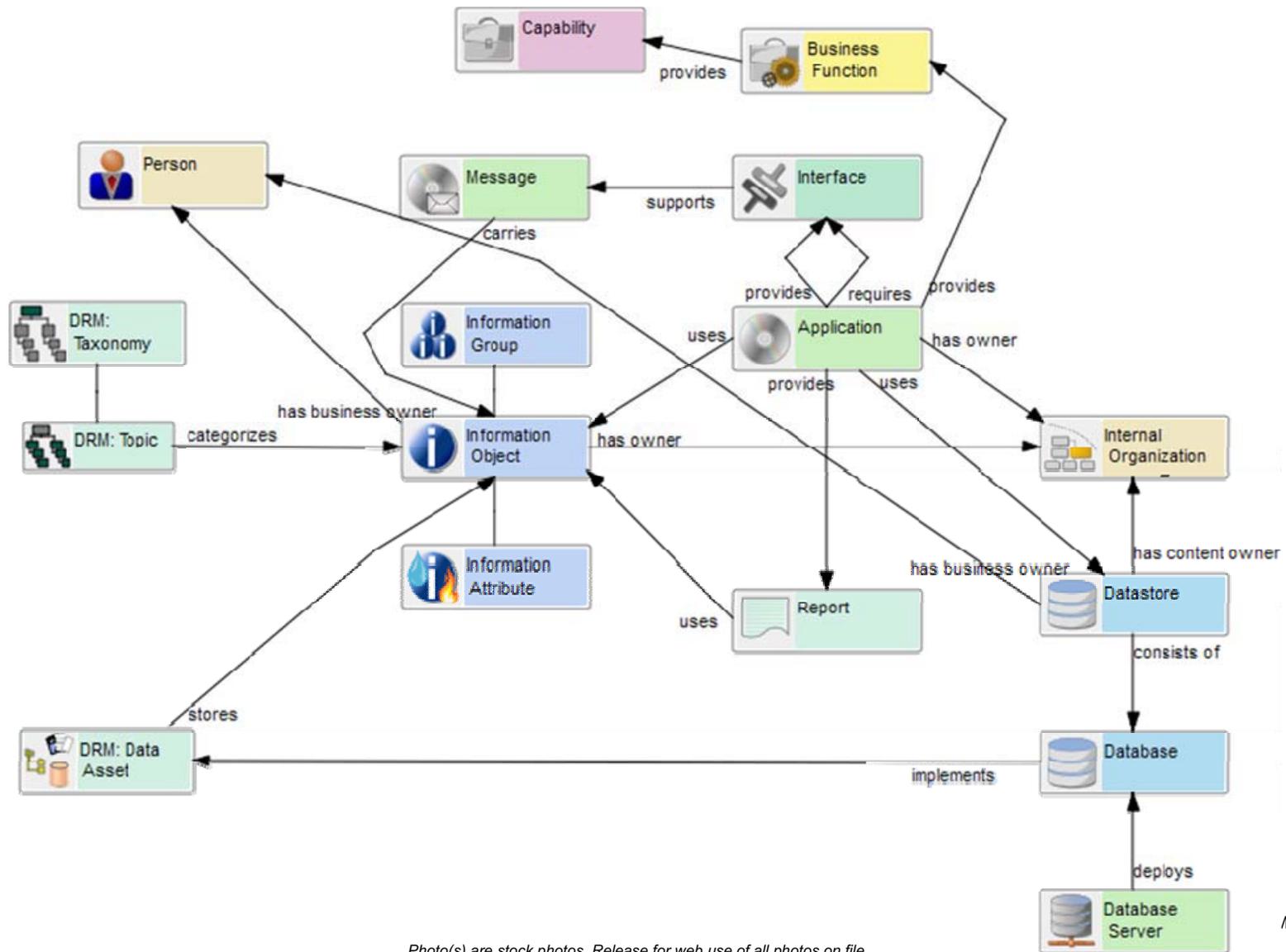


Understand the Data

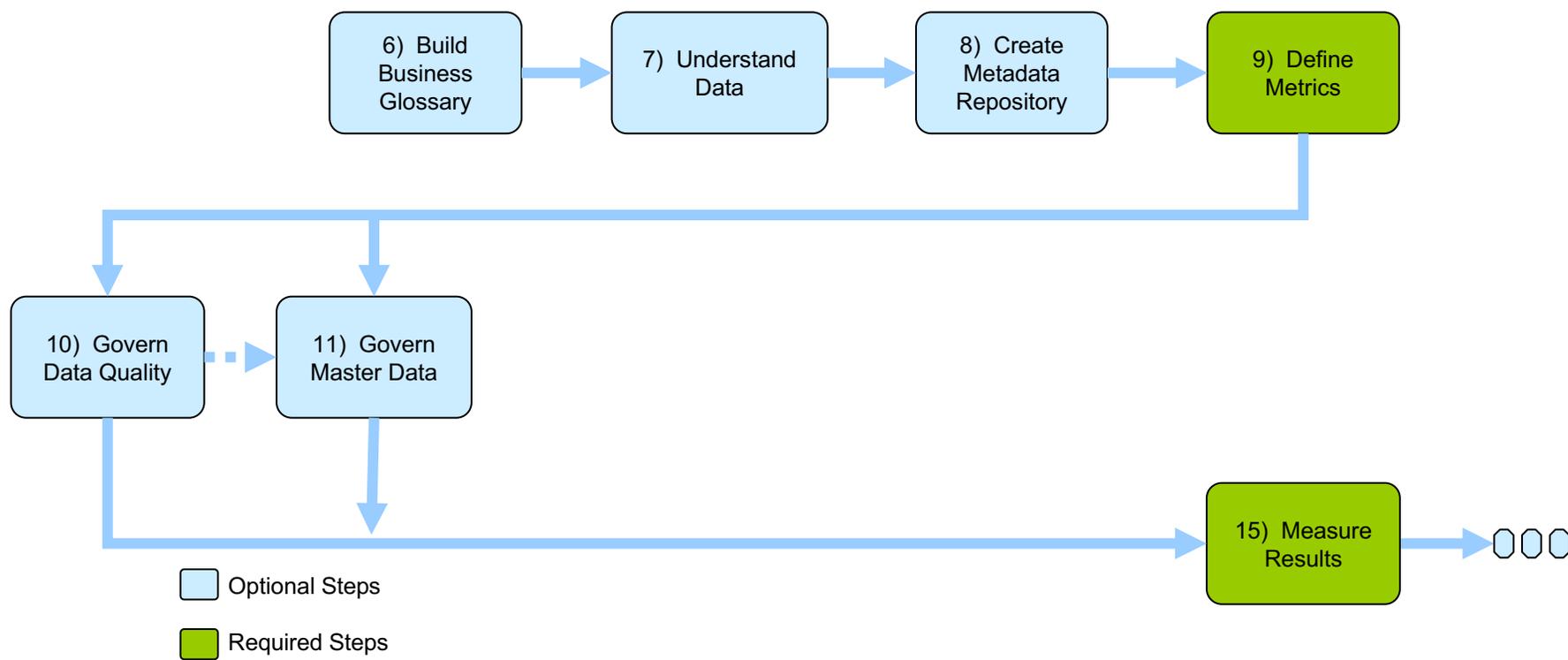
Information and Objects



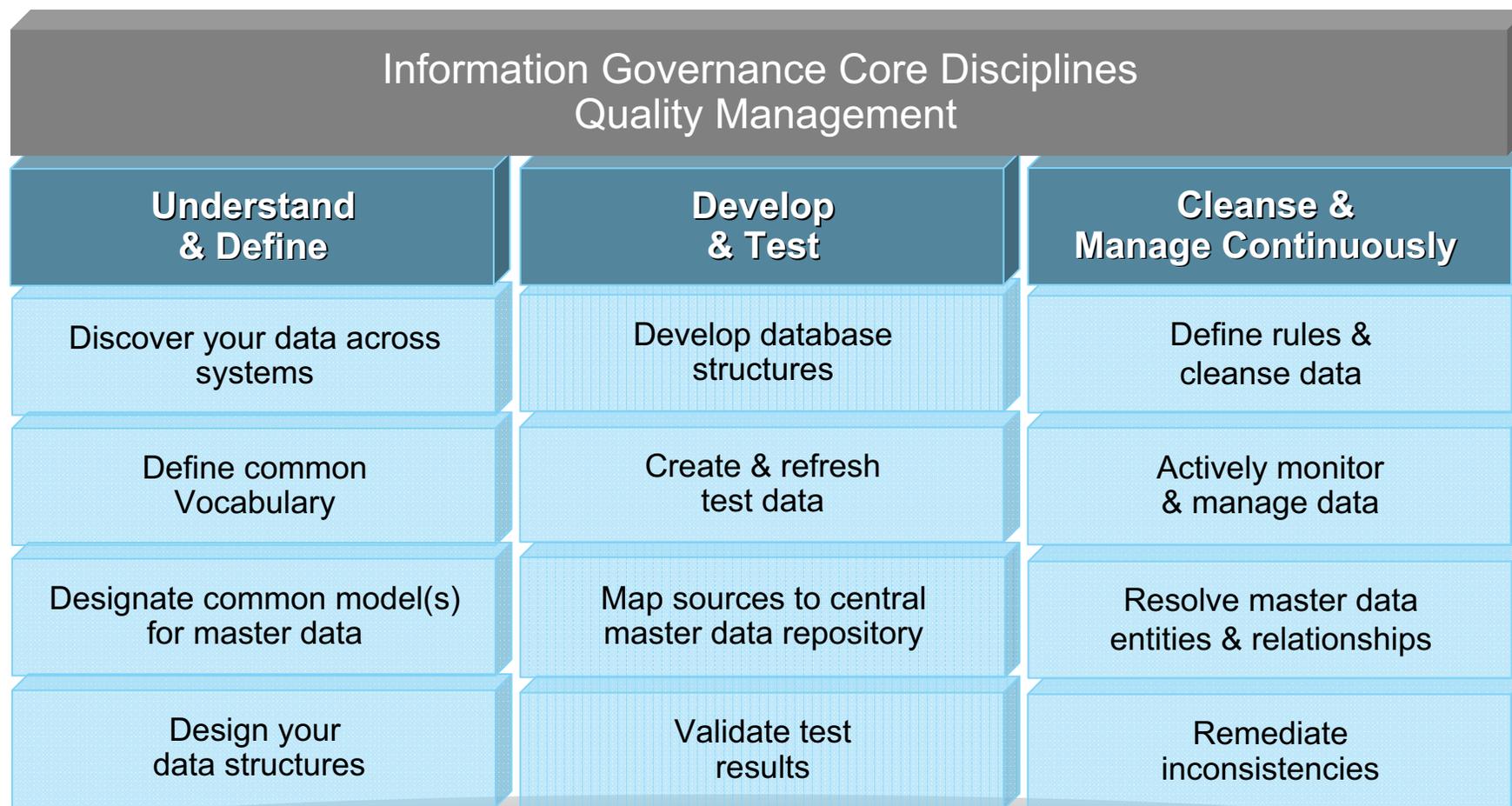
Understand Data - Information Metamodel



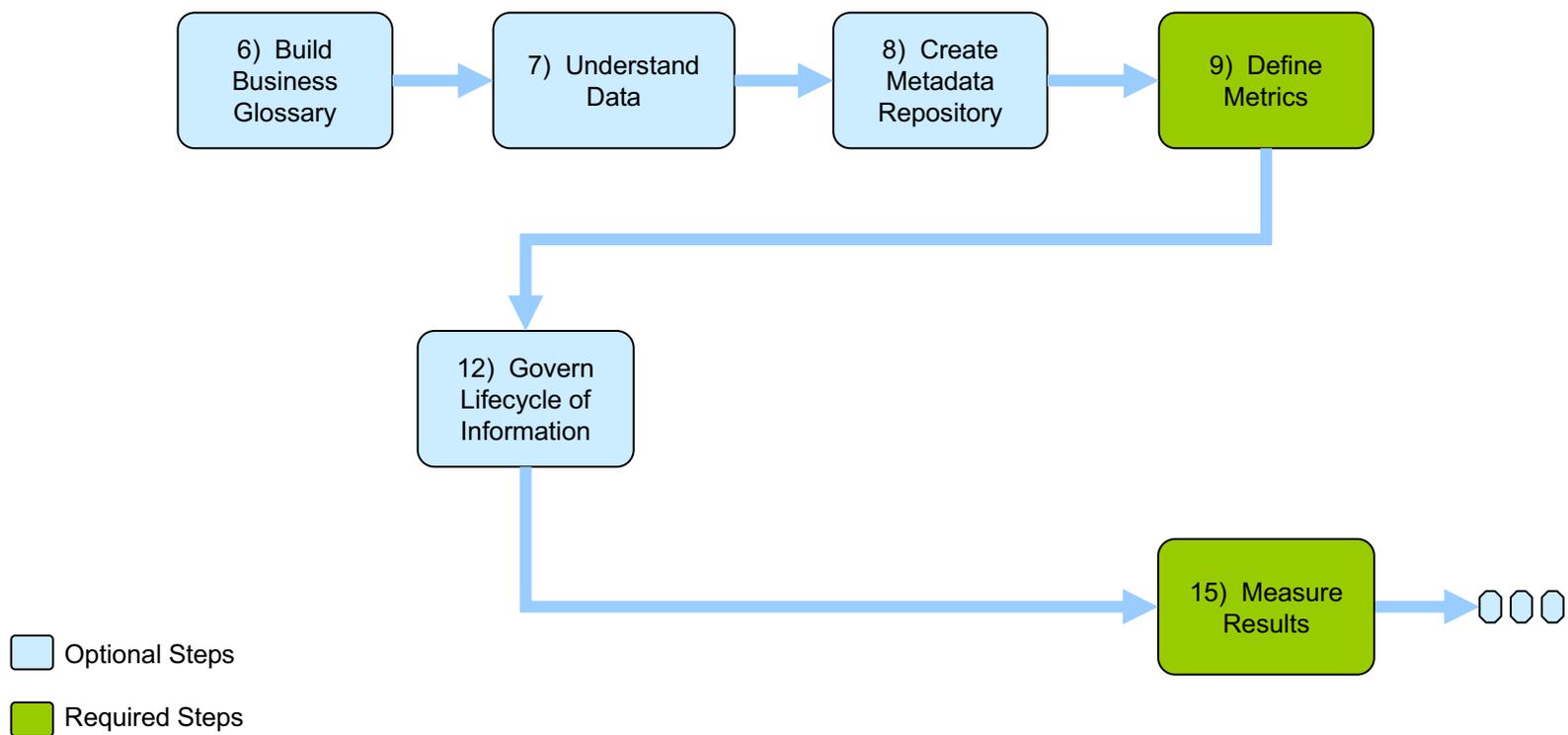
Master Data Governance



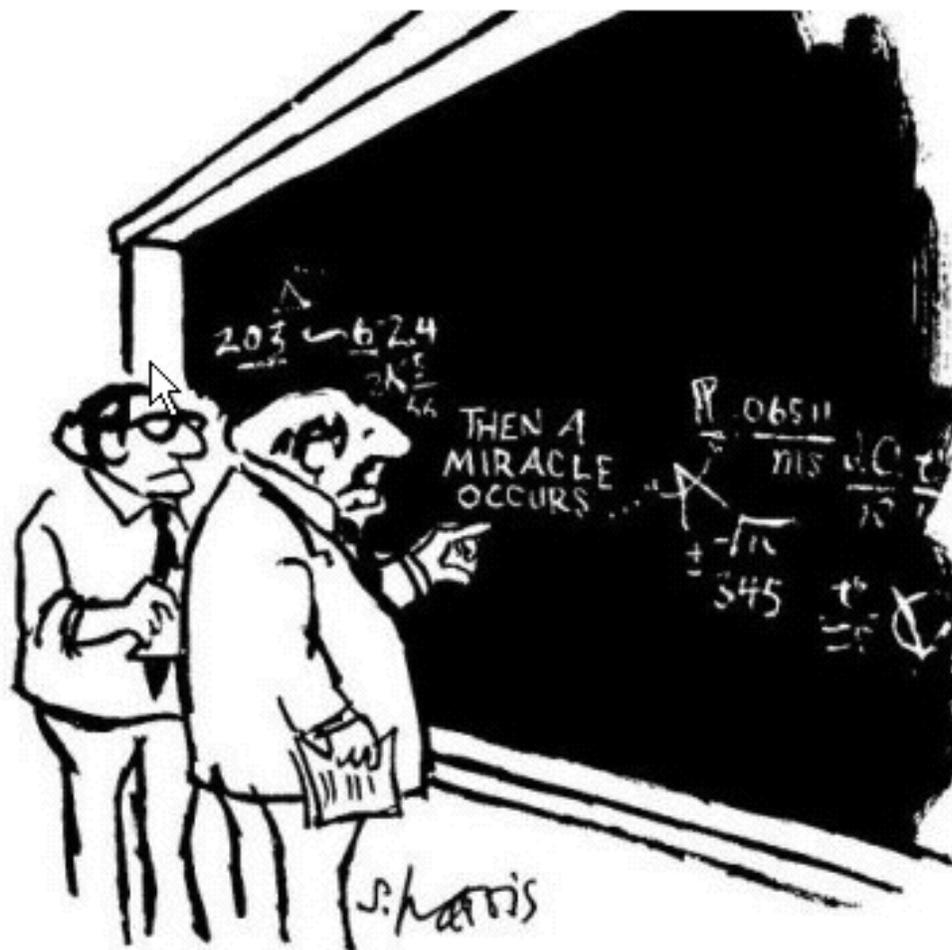
Requirements to govern the quality of data



Govern Lifecycle of Information

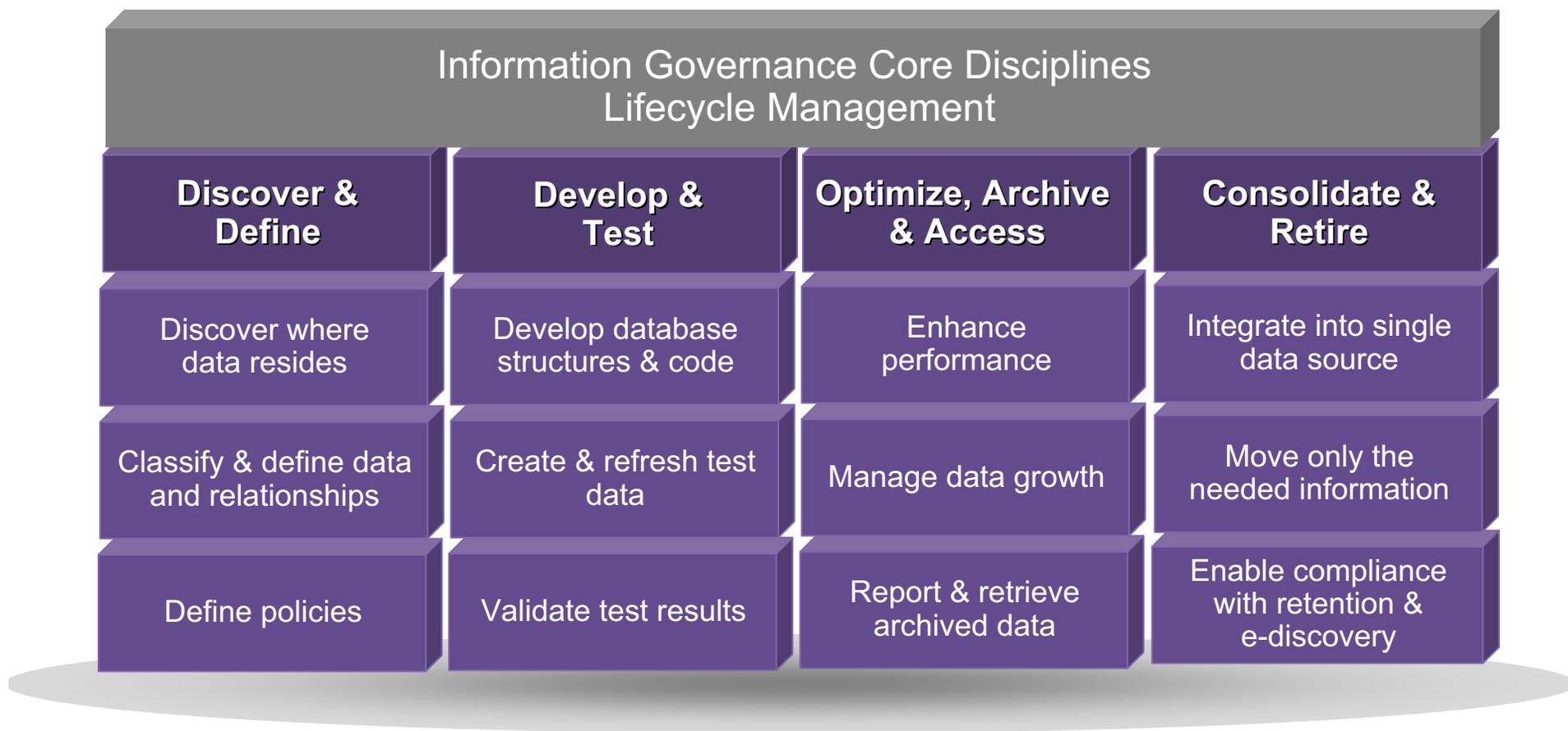


Common strategy for managing data growth

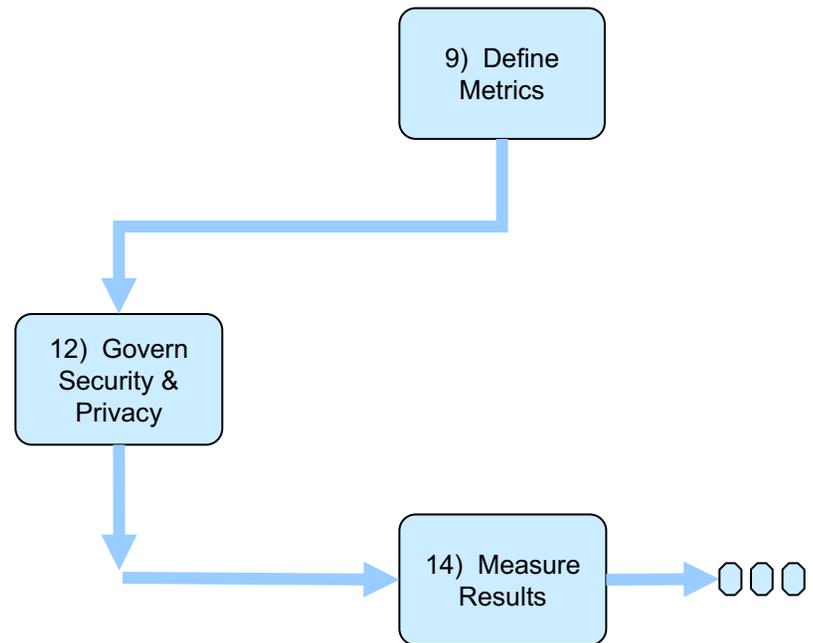


"I THINK YOU SHOULD BE MORE EXPLICIT HERE IN STEP TWO."

Streamlining the Governance of Data Across its Lifecycle

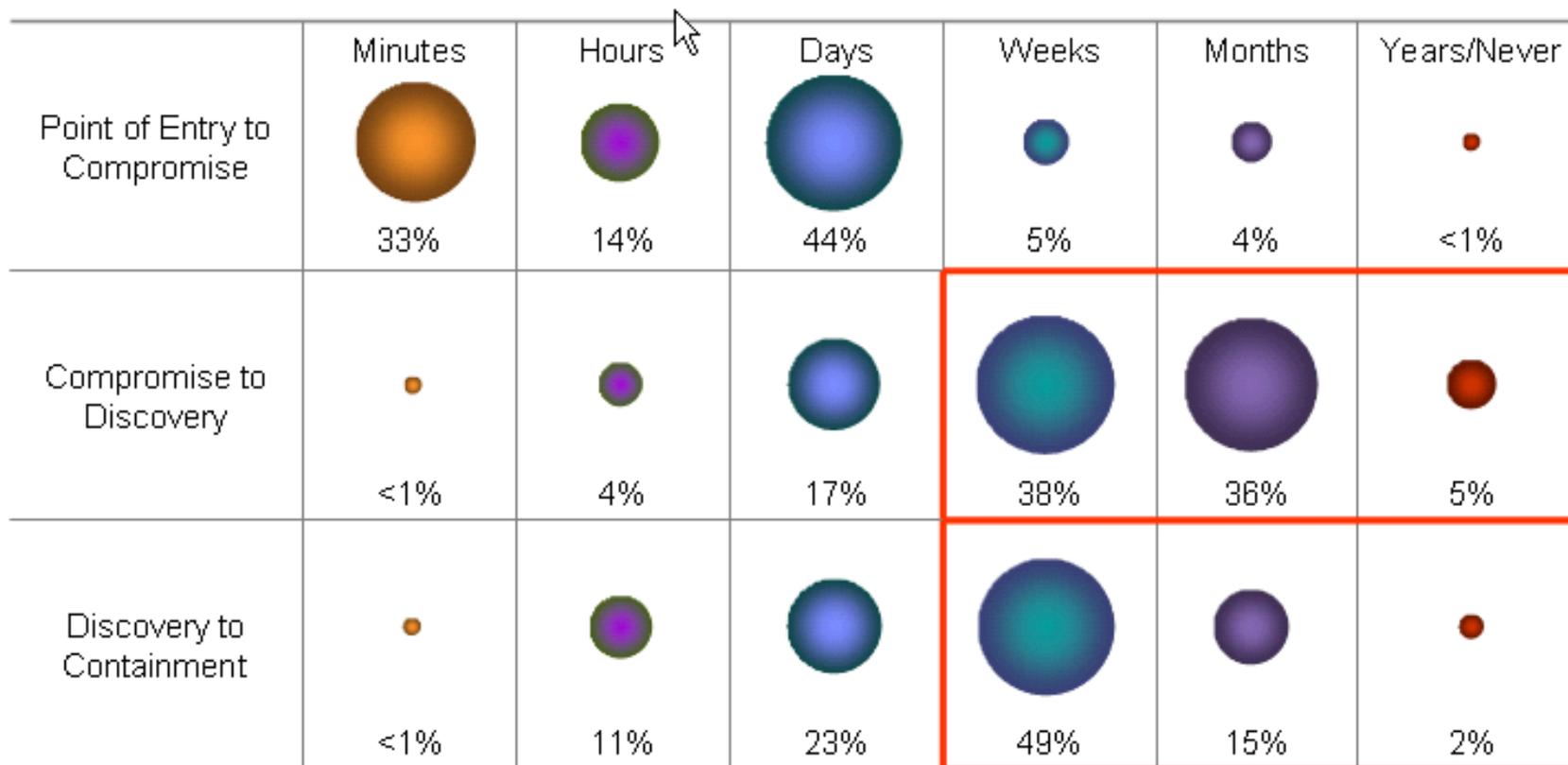


Govern Security and Privacy



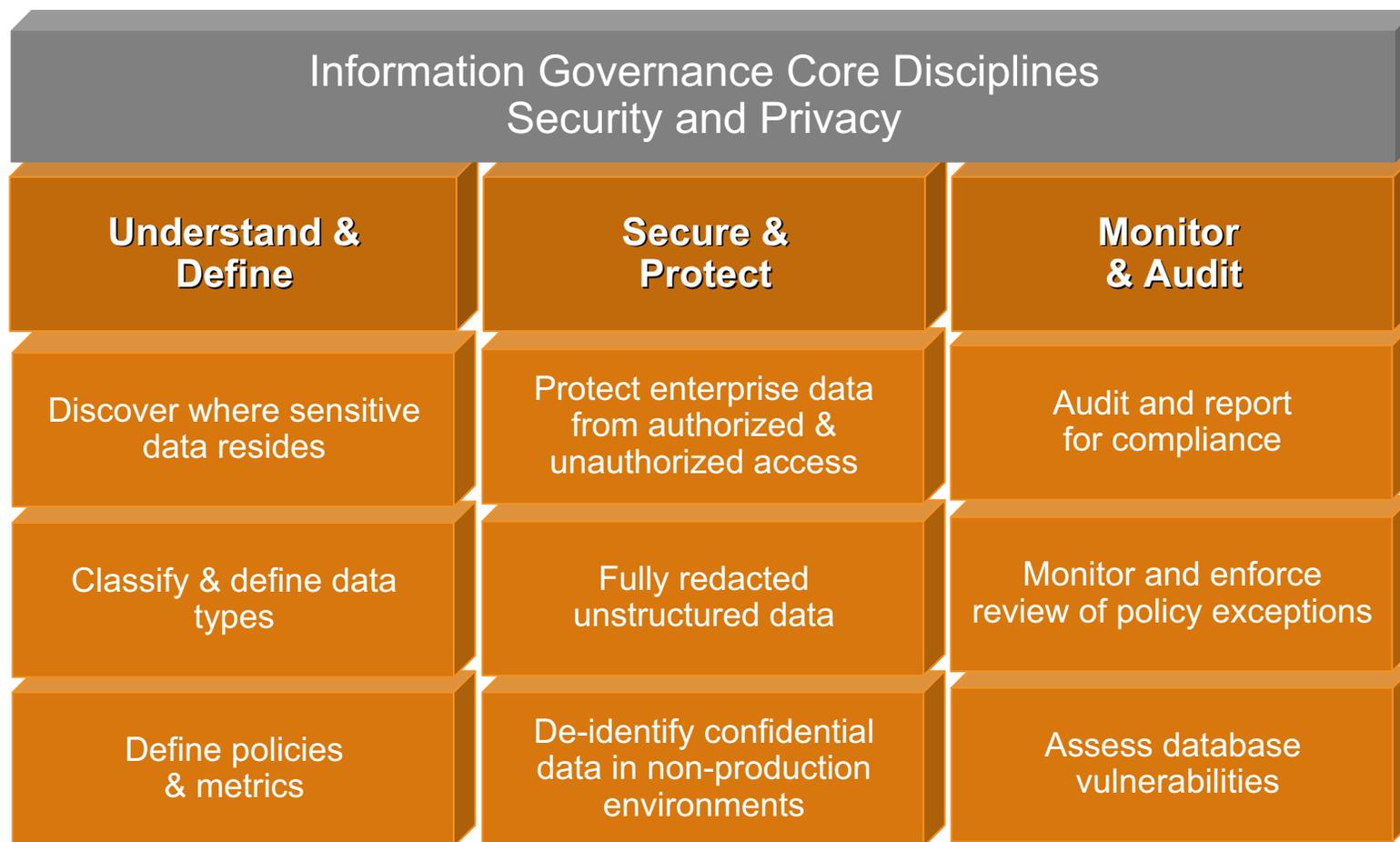
Do you know you have been breached?

Time span of events by percent of breaches

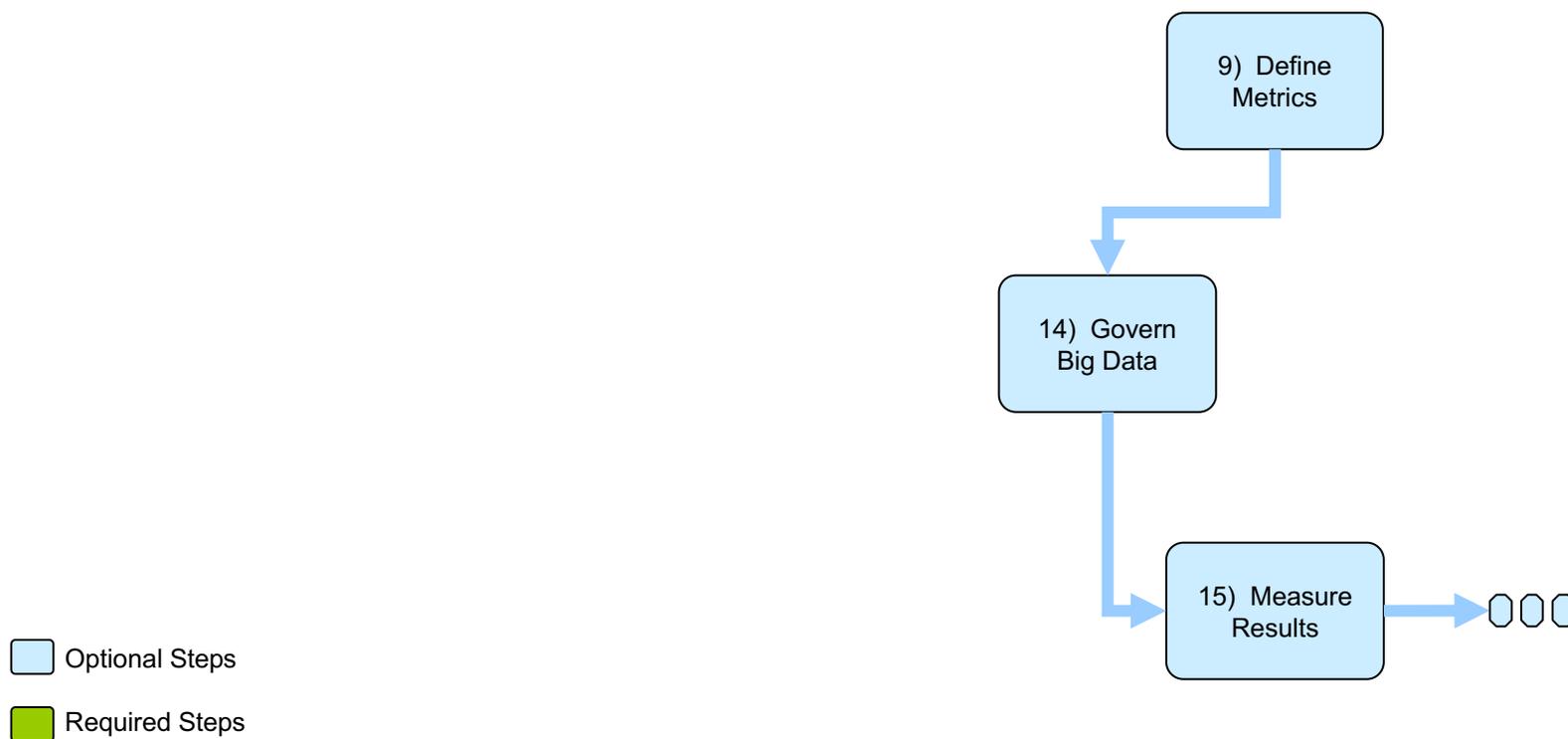


http://www.verizonbusiness.com/resources/reports/rp_data-breach-investigations-report-2011_en_xg.pdf

Requirements to govern the security and privacy of data



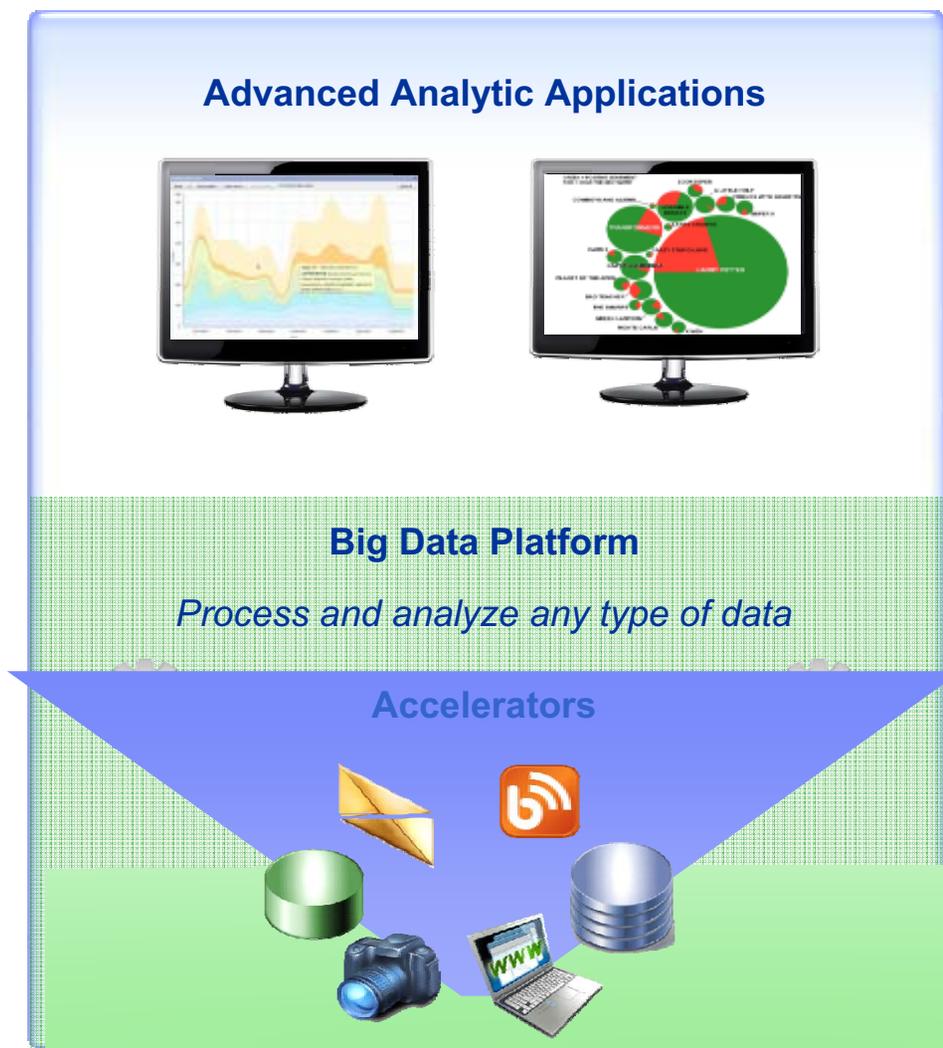
Govern Big Data



New analytic applications require a big data platform

Clients are realizing new possibilities by analyzing all of the available data

- Multi-channel customer analysis to prevent customer churn
- Social media analytics to predict next best product offer
- Securities transaction analysis to detect fraud
- Video and audio monitoring and analysis to determine potential security events
- Real-time meter feeds analysis to predict power consumption



Results from Good Information Governance

- **Understand your information**
 - Know what exists
 - How is it related
 - Ensure common understanding and definitions
- **Contain costs**
 - Manage costs with continuous growth
 - Retain information without growing retention costs
- **Maximize value from your information**
 - Make decisions that you can trust
 - Increase revenues
 - Reduce costs
- **Secure and Protect**
 - Keep information safe from internal and external threats
 - Know who is accessing what information and why
- **Comply with regulatory requirements**
 - Retention
 - Security
 - Filings
 - Audits



Thank you

