

Exorcising the Seven Deadly Data Sins

Peter Aiken, PhD

datablueprint.com

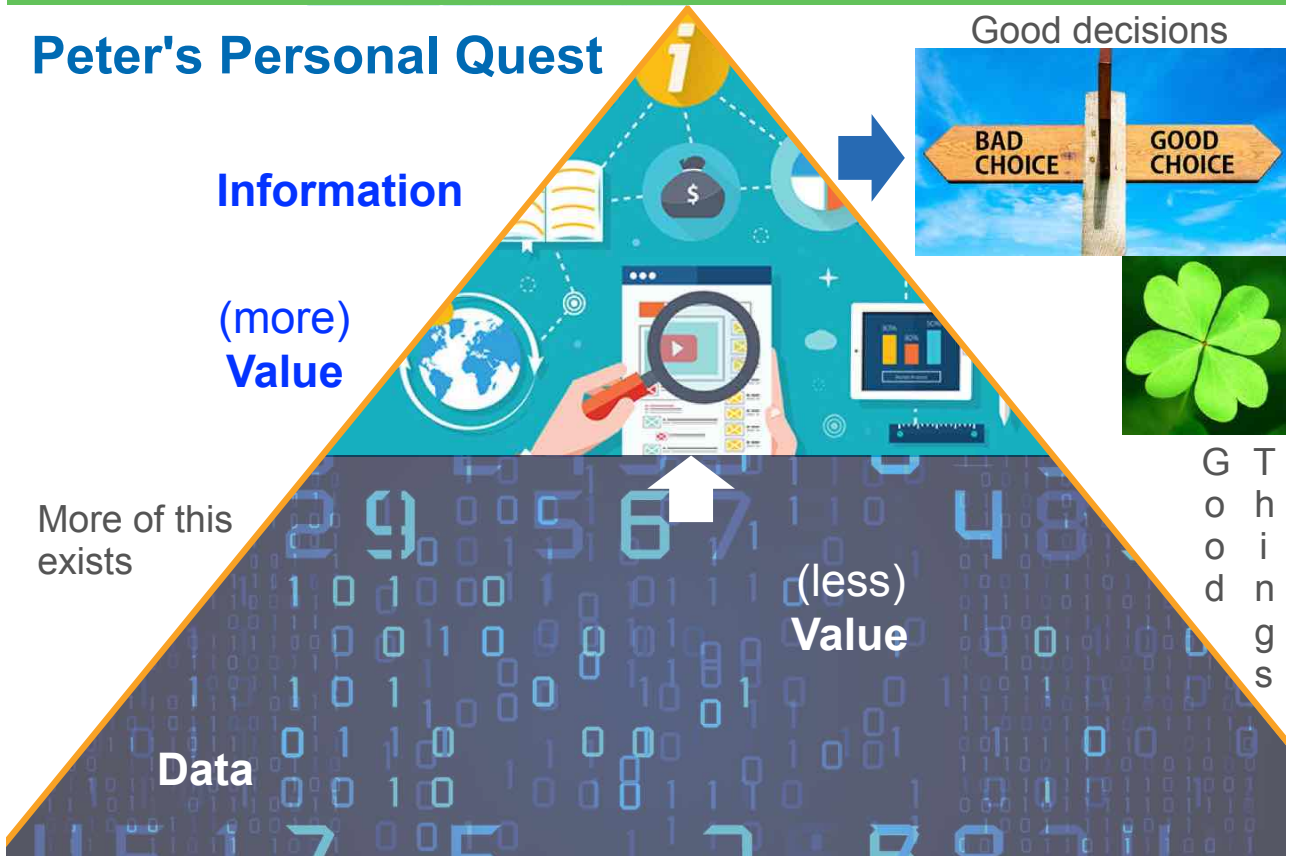
Peter Aiken, Ph.D.

- I've been doing this a long time
- My work is recognized as useful
- Associate Professor of IS (vcu.edu)
- Founder, Data Blueprint (datablueprint.com)
- DAMA International (dama.org)
- 10 books and dozens of articles
- Experienced w/ 500+ data management practices
- Multi-year immersions
 - US DoD (DISA/Army/Marines/DLA)
 - Nokia
 - Deutsche Bank
 - Wells Fargo
 - Walmart
 - ...





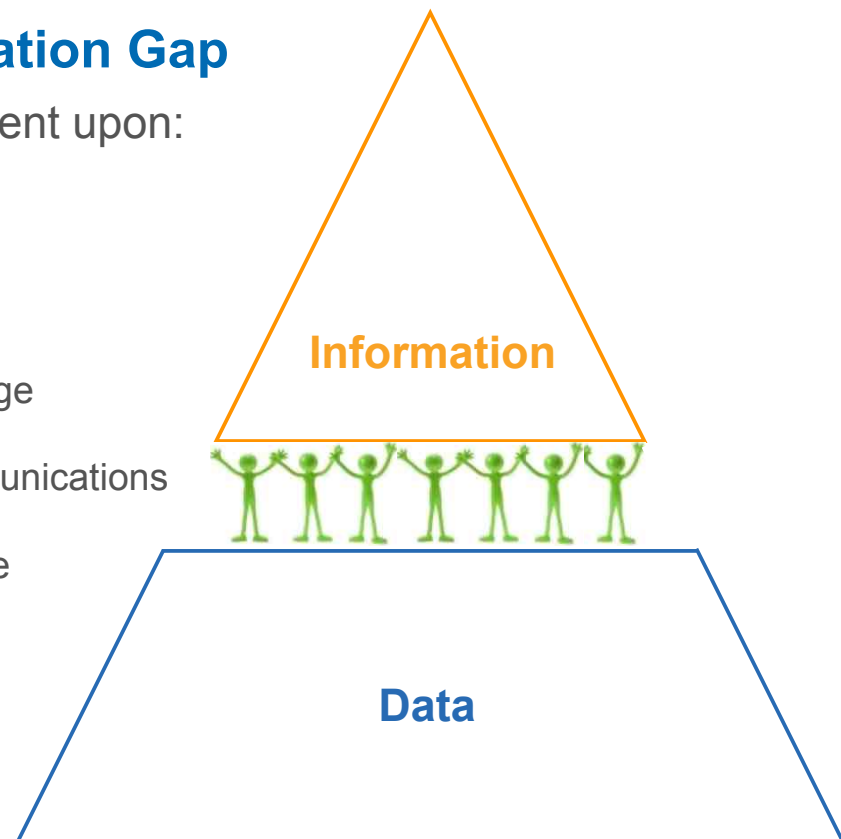
Peter's Personal Quest





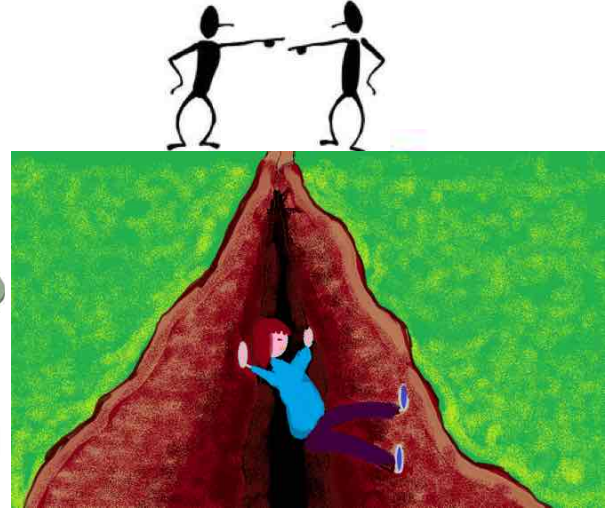
Data / Information Gap

- Overly dependent upon:
 - Human-beings
 - Wetwear
 - Tribal knowledge
 - Informal communications
 - Non-repeatable practices



Confusion

- IT thinks data is a business problem
 - "If they can connect to the server, then my job is done!"
- The business thinks IT is managing data adequately
 - "Who else would be taking care of it?"



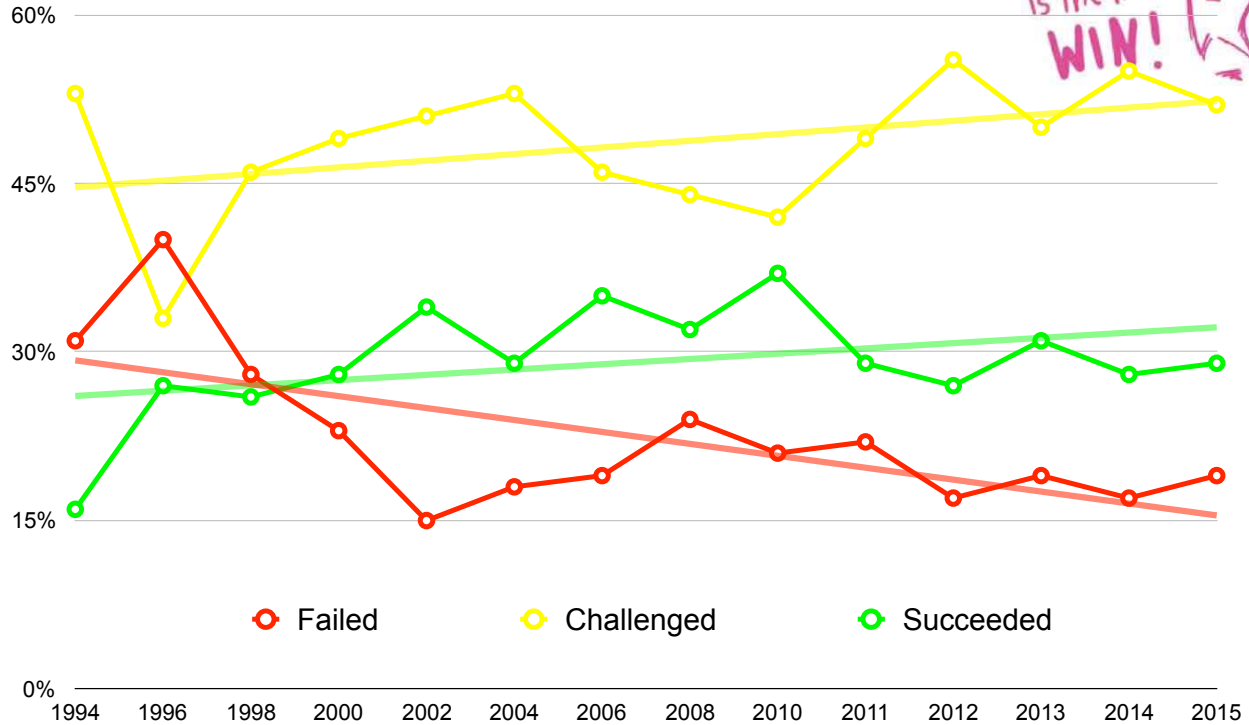
Communication/Perception Gap

	What	How	Where	Who	When	Why
Scope			Management Focus			
Business						
System	"Techie" Focus					
Technology						
Component						
Operations						

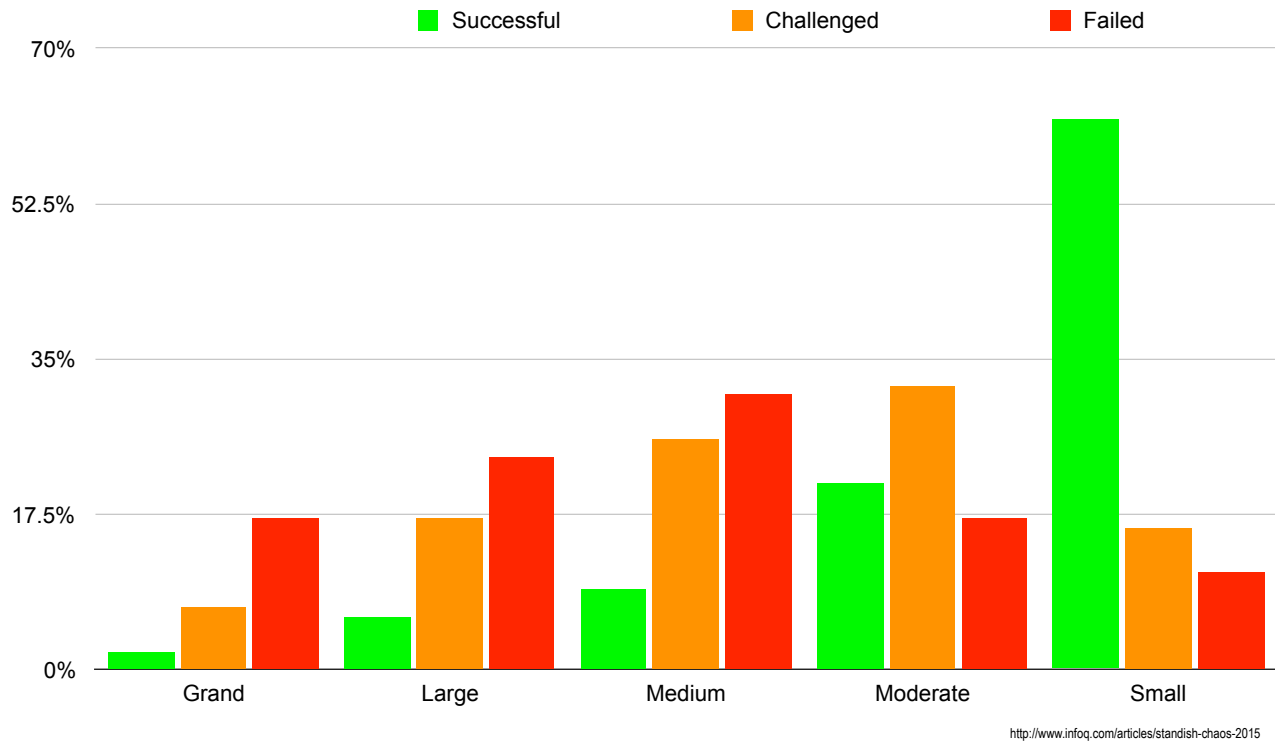
Courtesy of Jerry Rosenbaum

IT Project Failure Rates (1994-2015)

FAIL is the new WIN!



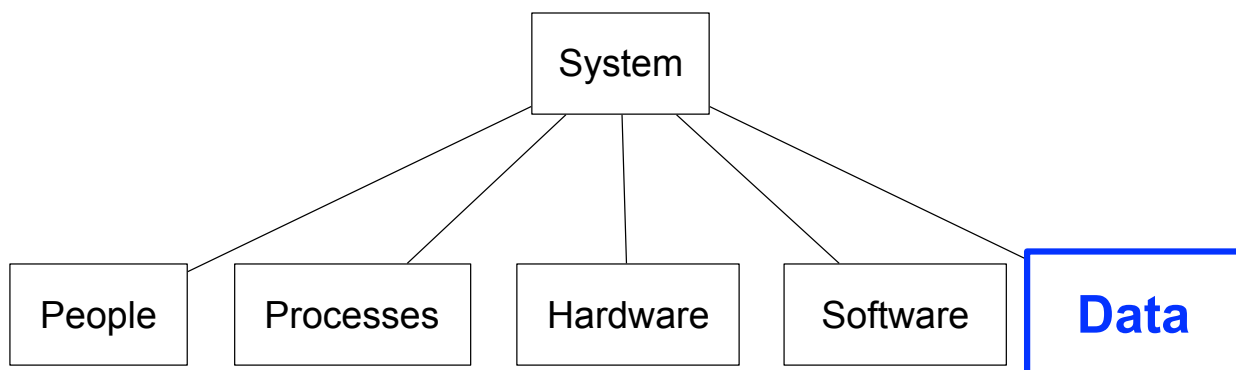
Chaos Resolution by Project Size (2011-2015)



System

- A set of detailed methods, procedures, and routines established or formulated to carry out a specific activity, perform a duty, or solve a problem.
- An organized, purposeful structure regarded as a whole and consisting of interrelated and interdependent elements (components, entities, factors, members, parts, etc.). These elements continually influence one another (directly or indirectly) to maintain their activity and the existence of the system, in order to achieve the goal of the system.

<http://www.businessdictionary.com/definition/system.html#ixzz23T7LyAjJ>





As articulated by Micheline Casey



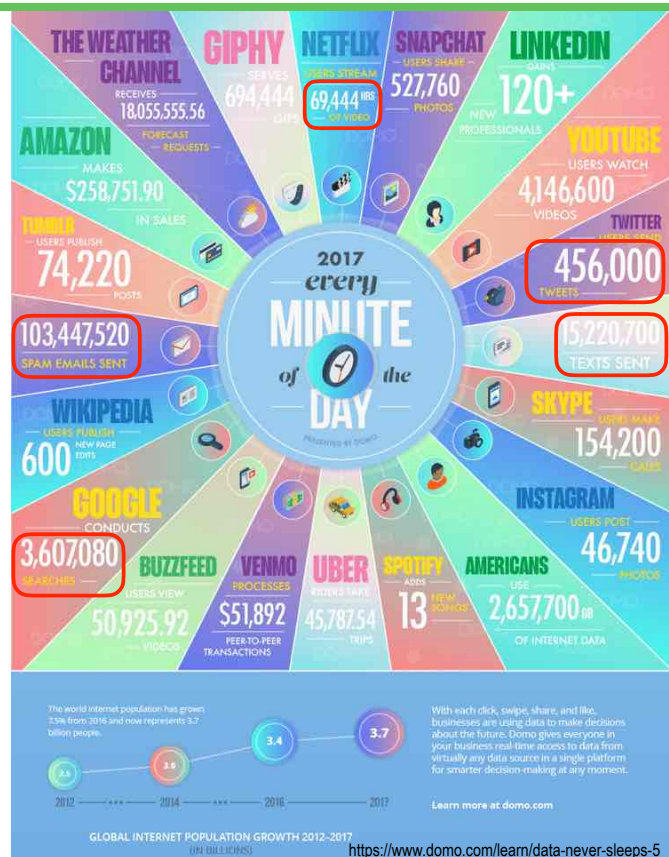
UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 13

How much data, by the minute!

For the entirety of 2017, every minute of every day:

- (almost) Seventy thousand hours of Netflix
- (almost) a half million tweets
- 15+ million texts
- 3.5+ million google searches
- 103+ million email spams



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 14

Data Assets Win!

Asset: A resource controlled by the organization as a result of past events or transactions and from which future economic benefits are expected to flow [Wikipedia]

- Today, data is the most powerful, yet underutilized and poorly managed organizational asset
- Data is your
 - Sole
 - Non-depletable
 - Non-degrading
 - Durable
 - Strategic
- Asset
 - Data is the new oil!
 - Data is the new (s)oil!
 - Data is the new bacon!
- As such, data deserves:
 - It's own strategy
 - Attention on par with similar organizational assets
 - Professional ministraton to make up for past neglect

	Data Assets	Financial Assets	Real Estate Assets	Inventory Assets
Non-depletable	Available for subsequent use	Can be used up		Can be used up
Non-degrading	✓	✓	Can degrade over time	Can degrade over time
Durable	Non-taxed		✓	✓
Strategic Asset	✓	✓	✓	✓



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 15

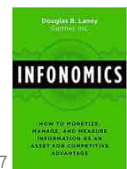
Information is the new oil – (No it's not!)

Information is:

- Non rivalrous
 - A good is considered **non-rivalrous** or non-rival if, for any level of production, the cost of providing it to a marginal (additional) individual is zero.
- Non depleting
 - Does not require replenishment
- Regenerative
- Nearly unlimited
- Low inventory and transportation/ transmission costs
- More difficult to control and own
- Eco friendly
- Impossible to clean-up if you spill it

CONFIDENTIAL AND PROPRIETARY
 I © 2017 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner and ITxpo are registered trademarks of Gartner, Inc. or its affiliates.

<https://www.amazon.com/Infonomics-Monetize-Information-Competitive-Advantage/dp/1138090387>



UNLOCKING BUSINESS VALUE

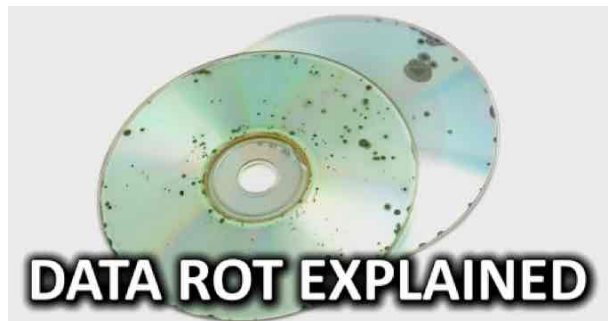
Copyright 2018 by Data Blueprint Slide # 16

Separating the Wheat from the Chaff



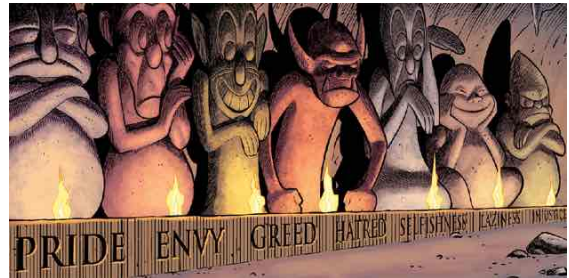
- Data that is better organized increases in value
- Poor data management practices are costing organizations money/time/effort
- 80% of organizational data is **ROT**

- Redundant
- Incomplete
- Obsolete
- Trivial



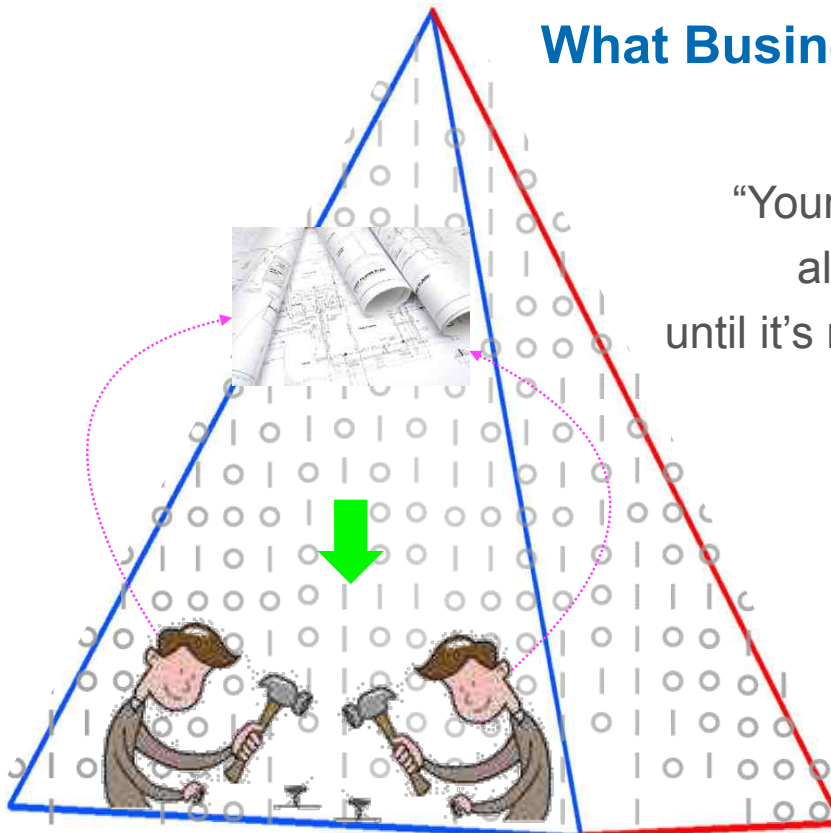
Exorcising the Seven Deadly Data Sins

- 7 Failing To Address Cultural And Change Management Challenges



What Business are you in?

“Your Organization is all about Data, until it’s not about just Data”



Chief Data Officer Combat

- Recasting the executive team. make full use of the most valuable assets

清华大学数据科学研究中心与清华大学数据产业联合研究中心



The Case for the Chief Data Officer

Recasting the C-Suite to Leverage Your Most Valuable Asset

首席数据官实战：

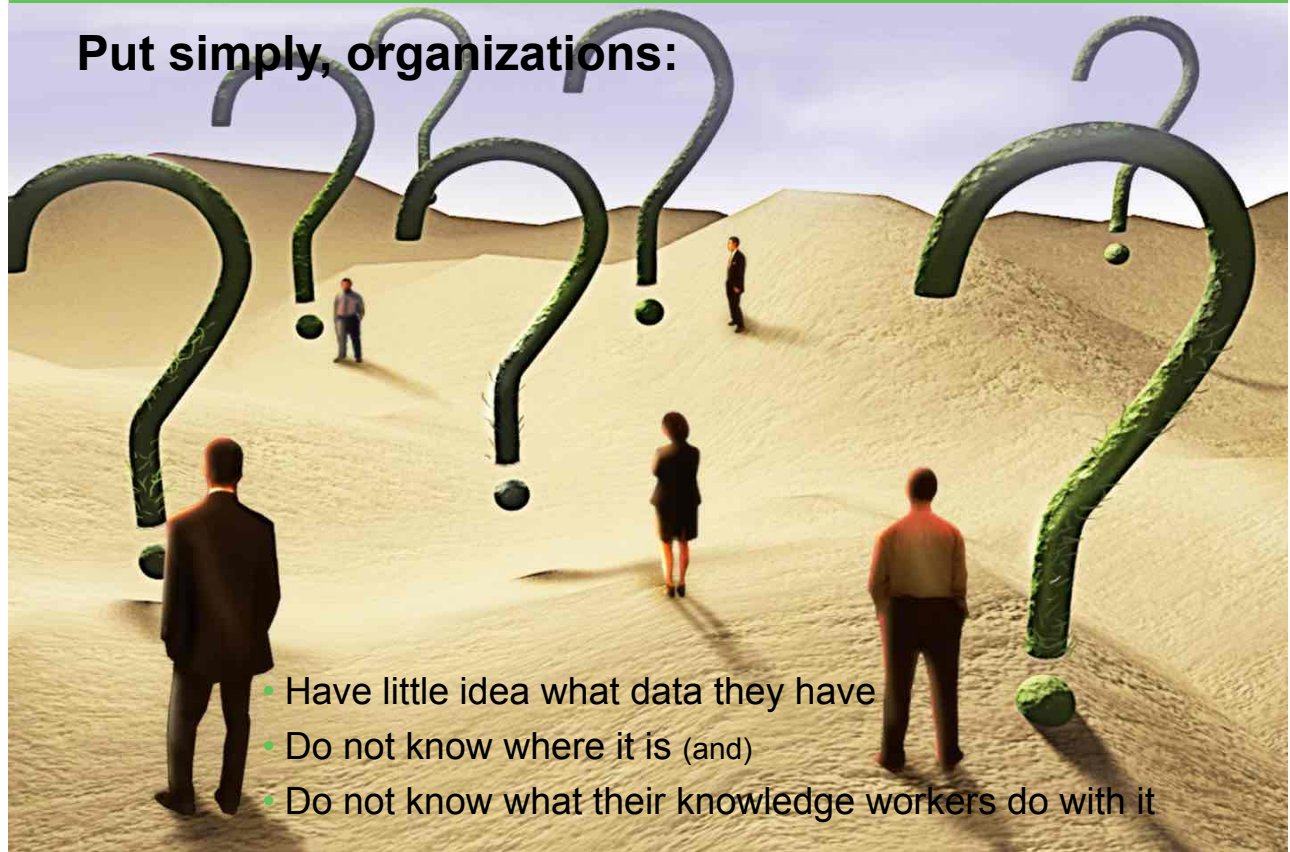
重组高管团队，充分利用最有价值资产

【美】Peter Allen, William Gorman 著 / 刘松武等译



清华大学出版社

Put simply, organizations:



- Have little idea what data they have
- Do not know where it is (and)
- Do not know what their knowledge workers do with it

Old Beer Accounting



This records a purchase of "best" beer from a brewer, c. 2050 BC from the Sumerian city of Umma in Ancient Iraq

http://en.wikipedia.org/wiki/File:Alulu_Beer_Receipt.jpg



The Hymn to Ninkasi, inscribed on a nineteenth-century BC tablet, contains a recipe for Sumerian beer.

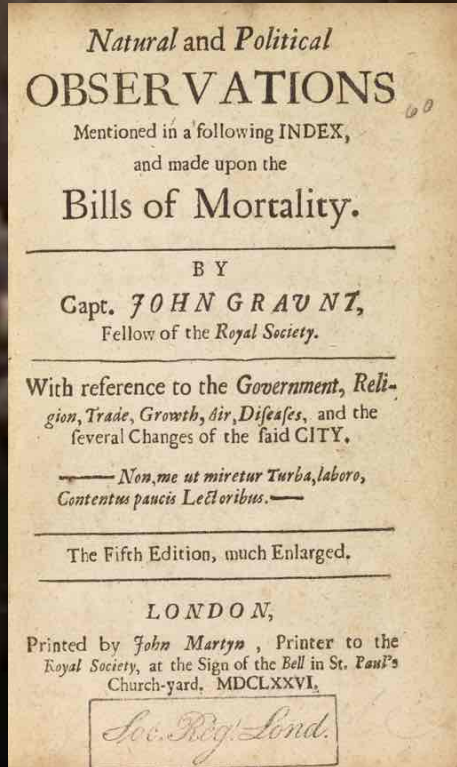
The first references to beer dates to as early as 6,000 BC. The very first recipe for beer is found on a 4,000-year-old Sumerian tablet containing the Hymn to Ninkasi, a prayer to the goddess of brewing.

<http://www.neatorama.com/2009/02/18/neatolicious-fun-facts-beer/#!kN0hf>

Bills of Mortality by Captain John Graunt



Plague in 1665.



Natural and Political OBSERVATIONS

Mentioned in a following INDEX,
and made upon the
Bills of Mortality.

BY

Capt. JOHN GRAUNT,
Fellow of the Royal Society.

With reference to the Government, Reli-
gion, Trade, Growth, Air, Diseases, and the
several Changes of the said CITY.

— Non me ut miretur Turba, laboro,
Contentus paucis Lecloribus. —

The Fifth Edition, much Enlarged.

LONDON,

Printed by John Martyn, Printer to the
Royal Society, at the Sign of the Bell in St. Paul's
Church-yard, MDCLXXVI.

See Reg. Lond.

Bills of Mortality

Bills of Mortality
The Diseases and Mortalities this Week

London 40 From the 17 of September to the 26, 1665

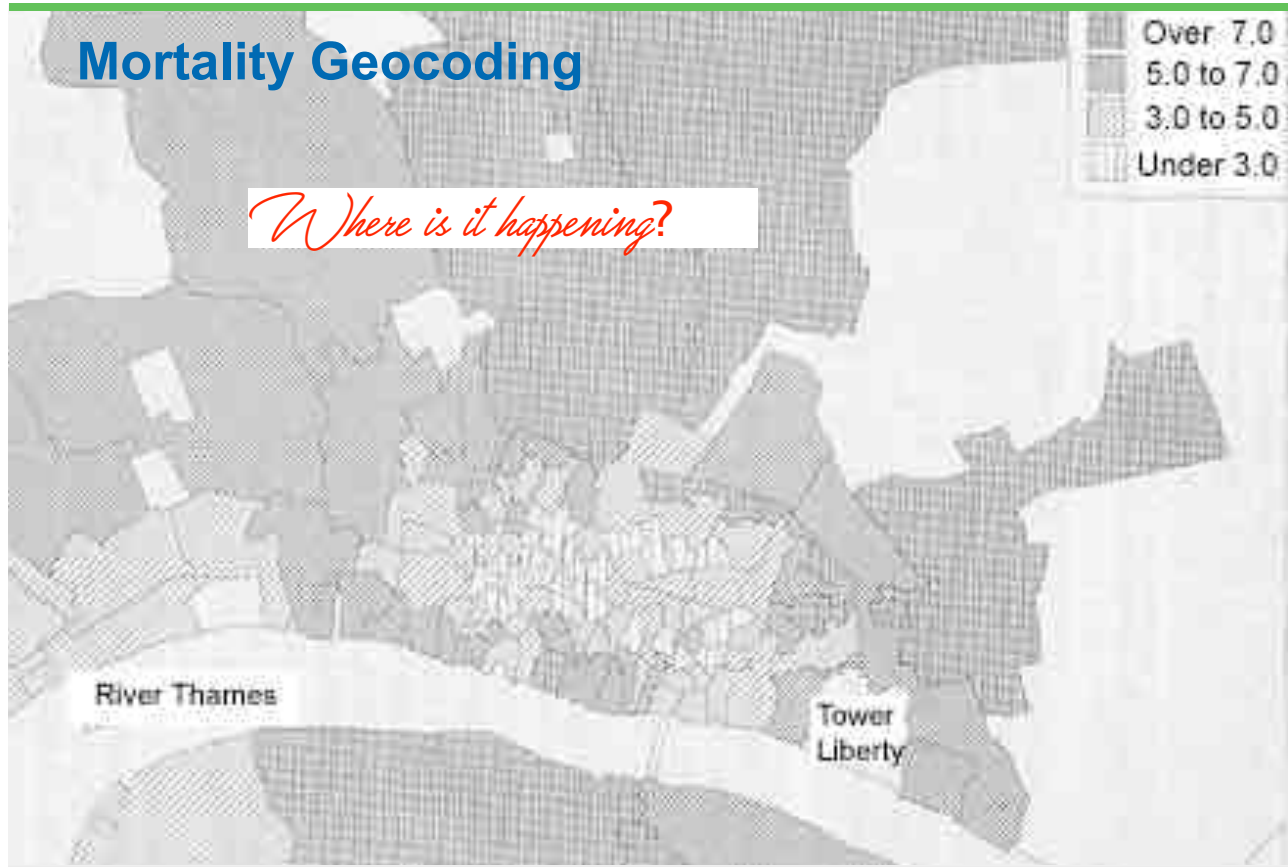
<p>A Bontive 5 Aged 42 Apoplexie 1 Bleeding 2 Burne on his Bed by a Canale at St. Giles Cripple lane 1 Cankers 4 Children 41 Cholmes 18 Consumption 134 Convulsion 64 Cough 2 Deopthe 35 Fetere 309 Flux and Small-pox 5 Frighted 3 Gawt 1 Grief 3 Griping in the Guts 11</p> <p>Christned Males 852 Females 817 In all 1769</p> <p>Interred in the Buriall this Week 607 Parishes Interred of the Plague 4</p>	<p>Janthine 5 Incontinencie 11 Forrains 14 Killed by a fall from the Bell-tower Alhallwas the Great 4 Kinge-wild 4 Leishargy 1 Puffe 3 Plague 71 Rickets 15 Killing of the Leghes 17 Scouring 1 Scurrey 1 Soles 1 Spotted Fleaver 10 Stillborn 10 Struck 1 <u>Swelling of the stomack</u> 9 Strangury 1 Suddensy 1 Surfeis 10 Teeth 12 Thrush 3 Tympany 1 Itchick 11 Vomiting 5 Winde 2 <u>Wormes</u> 11</p> <p>Christned Males 4051 Females 4102 In all 8153</p> <p>Plague 718</p>
--	---

*The Affixe of Bread set forth by Order of the Lord Mayor and Court of Aldermen
 A penny Wheaten Loaf to contain Nine Ounces and a half, and three*



UNLOCKING BUSINESS VALUE

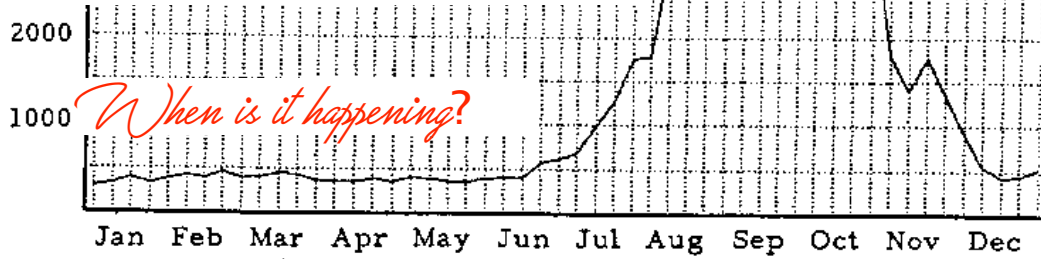
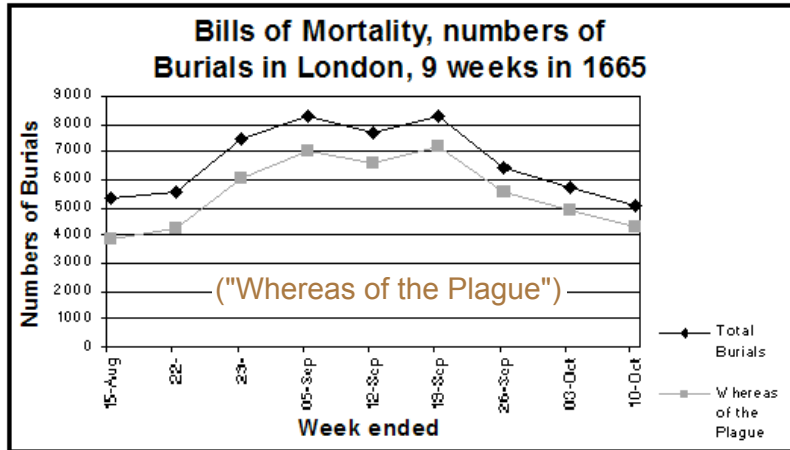
Mortality Geocoding



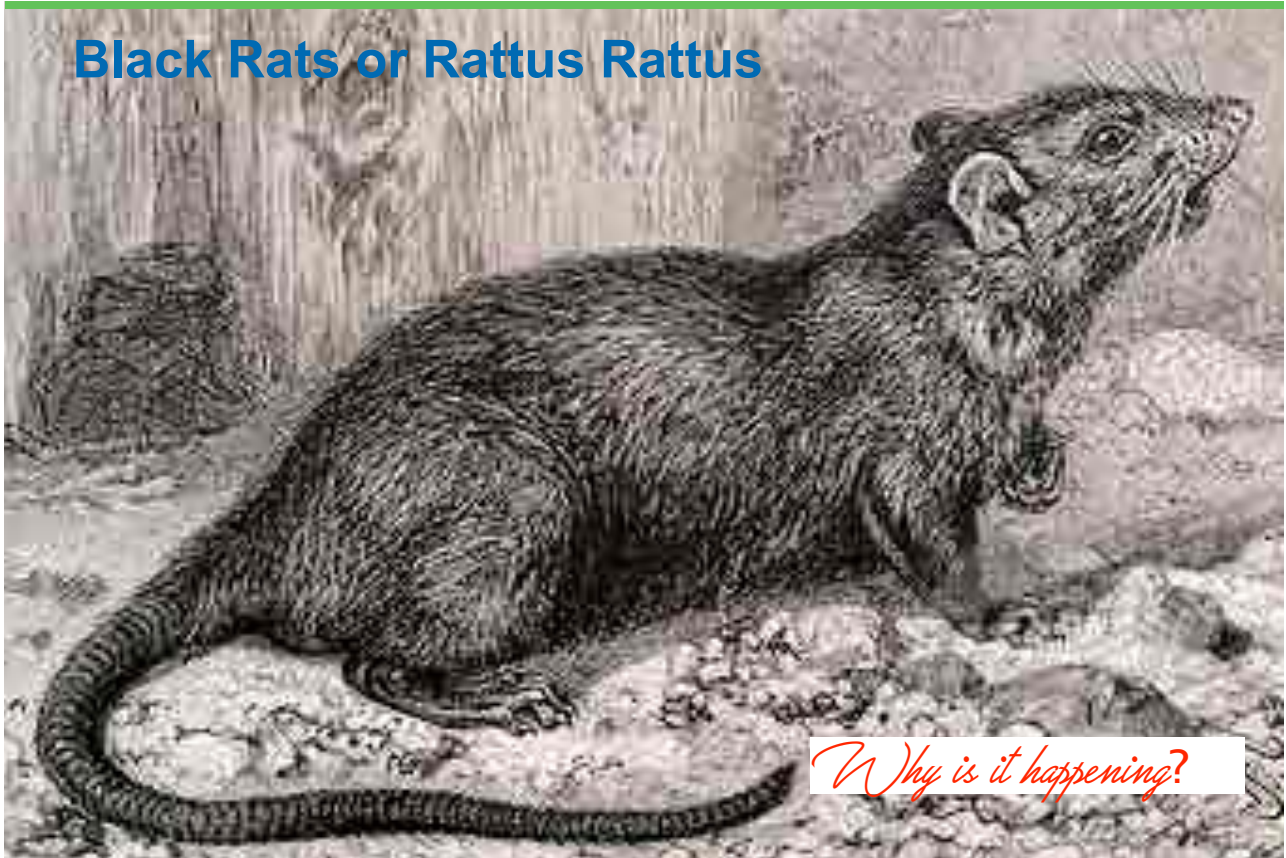
UNLOCKING BUSINESS VALUE

Plague Peak

8000



Black Rats or Rattus Rattus



Why is it happening?

What Will Happen?

What will happen?



200 years on

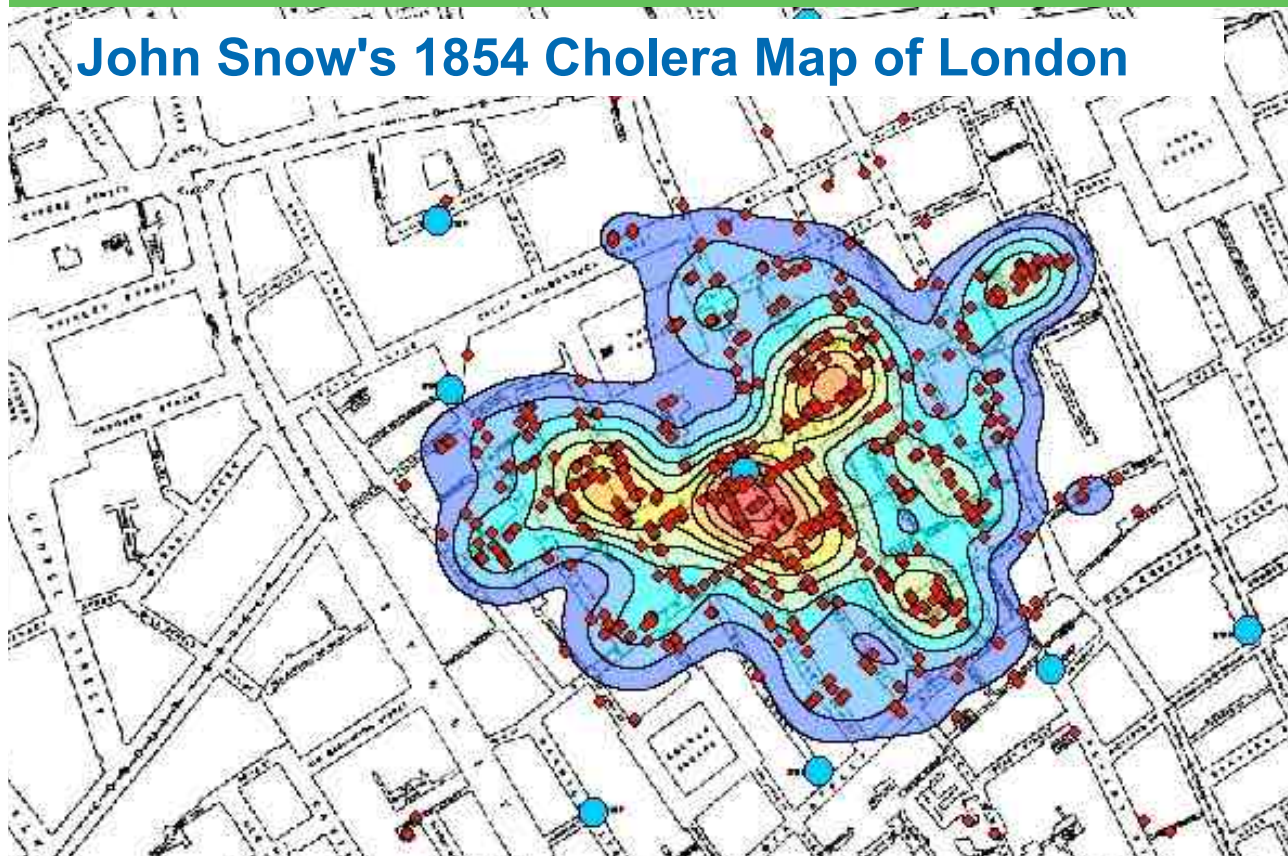
200
YEARS

Lord Francis Bacon, 1st Viscount St. Alban

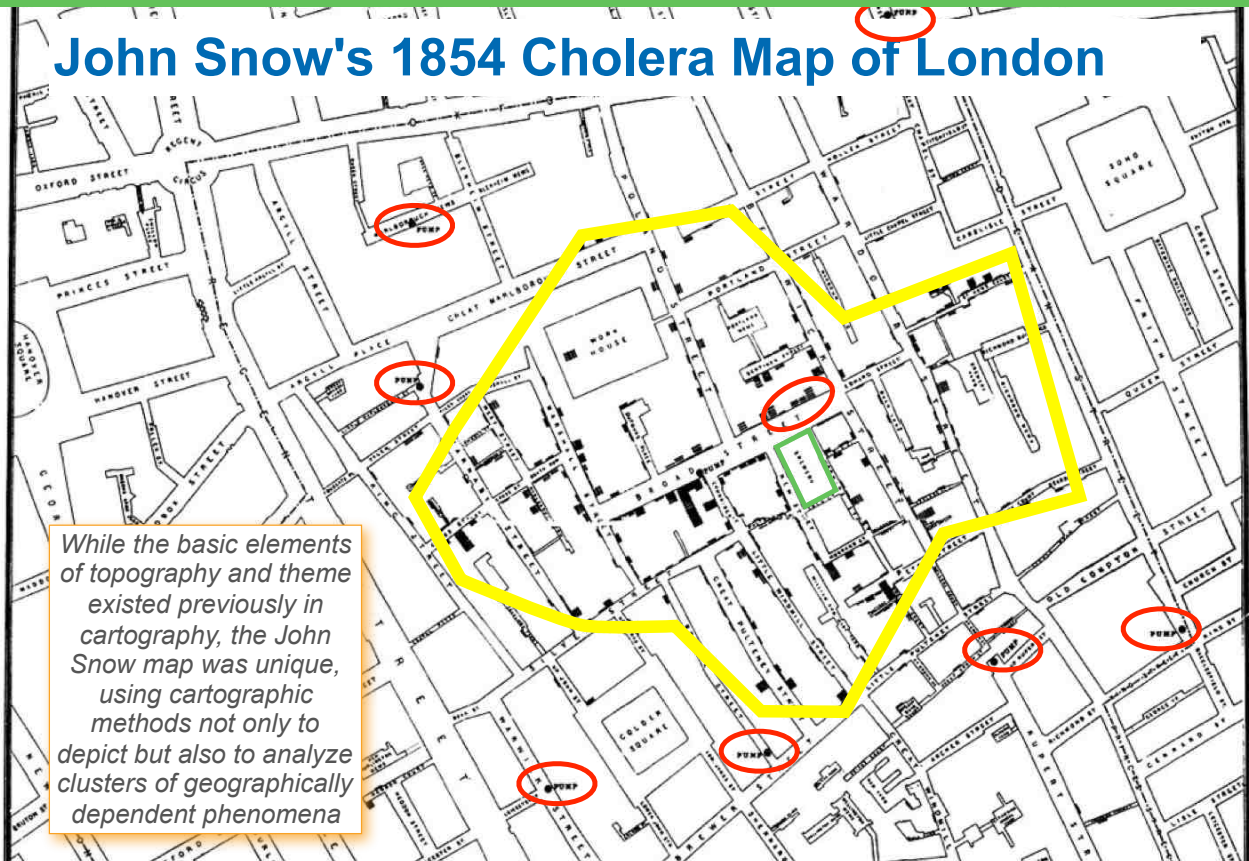
- Father of empiricism
- Popularized inductive methodologies for scientific inquiry
- Inspiration for the founding of the Royal Society in 1660



John Snow's 1854 Cholera Map of London



John Snow's 1854 Cholera Map of London



While the basic elements of topography and theme existed previously in cartography, the John Snow map was unique, using cartographic methods not only to depict but also to analyze clusters of geographically dependent phenomena



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide #

60 years on

60 YEARS



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 34

Formalizing Data Systems

- *Defend the Realm: The authorized history of MI5* by Christopher Andrew
- World War I
- 1914
- At war with much of Europe
- 14,000,000 Germans living in the United Kingdom
- How to efficiently and effectively manage information on that many individuals?
- Using index cards?



30 years this time

30 YEARS

Hedy Lamarr

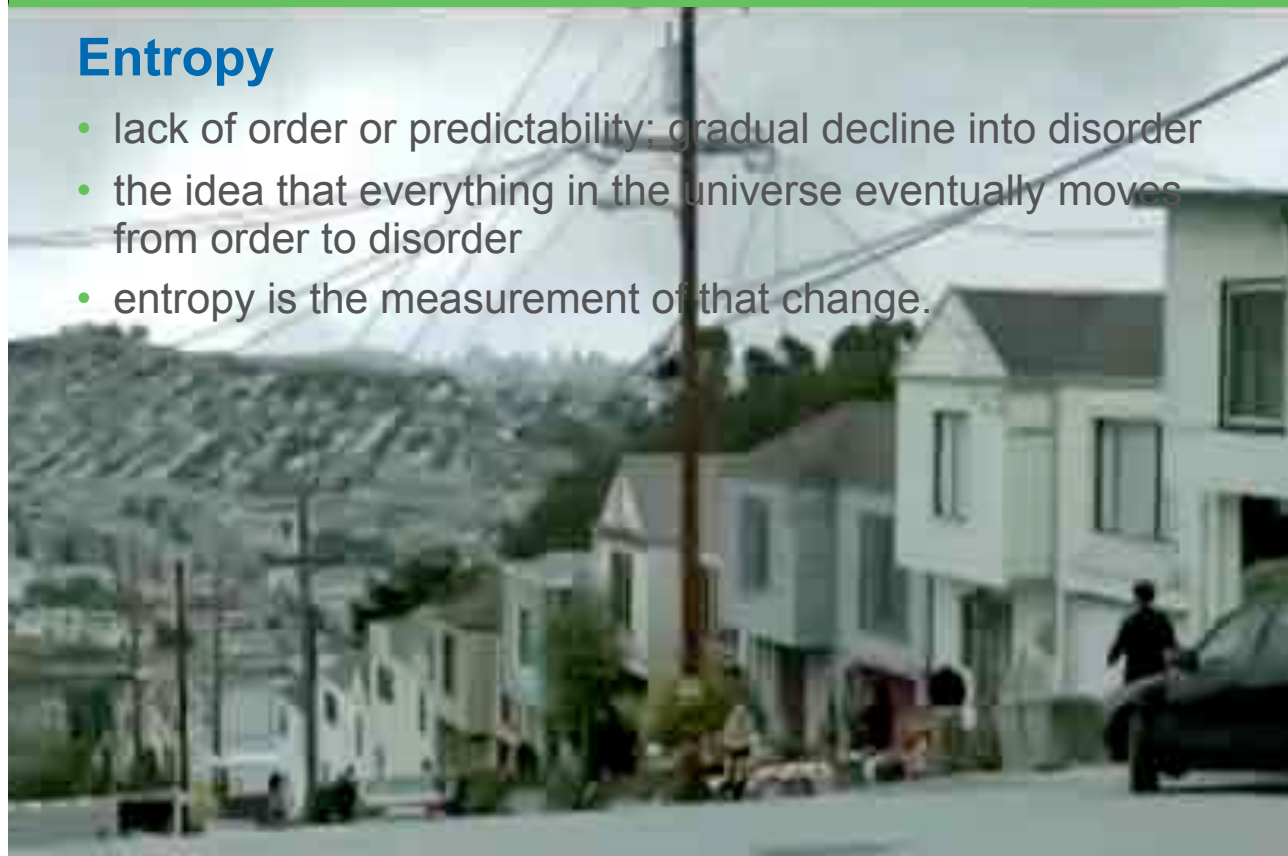
- Google celebrated her 101st Birthday on 11/8/2015
 - Tablets for fizzy drinks
 - Improved stop light design
- Invention of “frequency hopping” radio
 - By jumping from one radio frequency to another rapidly, only a receiver that shares the key can find the transmission
 - Prevent interference with the radio guidance controls of torpedoes
- U.S. Patent 2,292,387 (w/ George Antheil)
- Associated traffic analysis
 - Looking at other elements of a communication when you don't know the actual content
 - Time/duration of a message
 - Location of transmitters
 - Detect specific operator “fists”
 - Identify the operator and you could identify a specific ship or military unit, locate it with direction finding, and then track its activity over time

<https://theconversation.com/how-wwi-codebreakers-taught-your-gas-meter-to-snitch-on-you-29924>



Entropy

- lack of order or predictability; gradual decline into disorder
- the idea that everything in the universe eventually moves from order to disorder
- entropy is the measurement of that change.



50 years on



US DoD Reverse Engineering Program Manager

- *"Your first project is to keep me from having to testify to a Congressional Hearing!"* (Belkis Leon-Hong former ASD-C3I)
- Problem:
 - 37 systems paid personnel within DoD
 - How many were needed?
 - How many potential losers?
 - What do you mean by employee?
- Process modeling
 - Inconclusive results
- Data reverse engineering - definitive
 - One legged engineer, working in waist deep waters, underneath rotating helicopter blades, on overtime



US DoD Reverse Engineering Program Manager

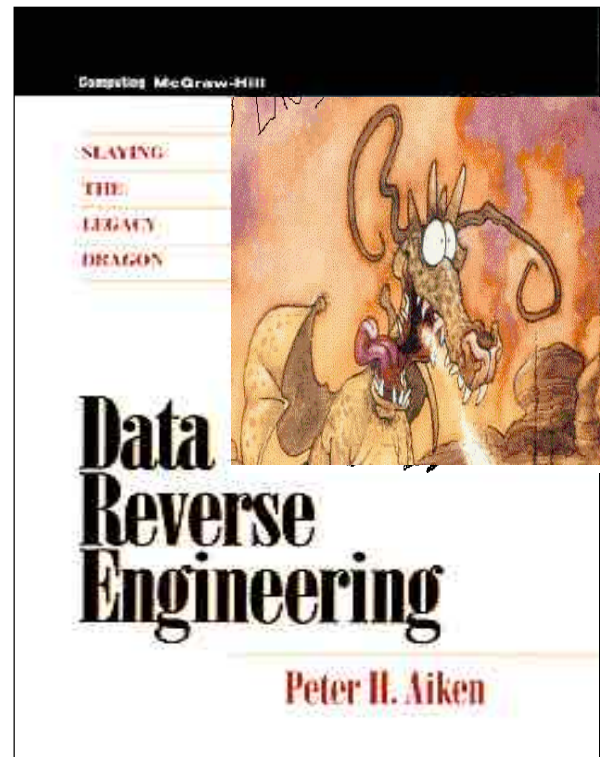
Specific case study lessons

- Lack of governance prevented management from occurring
- This was clearly not yet a technical problem
 - Technologists should not be the primary determinant of business concepts such as (employee)
 - It needs a technical resolution
 - Sequencing is critical
- An objective selection was made
 - Agreement on the specific decision
 - and about the process



US DoD Reverse Engineering Program Manager

- *Data Reverse Engineering* (1995)
- Amazon Best Sellers Rank:
 - **#3,835,772**
 - in Books
 - **#9,736**
 - in Books > Computers & Technology > Databases & Big Data
 - **#30,149**
 - in Books > Business & Money > Management & Leadership > Management (accessed April 29, 2017)



40 years this time



Change the **status quo!**

- *Keep in mind that the appointment of a CDO typically comes from a high-level decision. In practice, it can trigger an array of problematic reactions within the organization including:*
 - Confusion,
 - Uncertainty,
 - Doubt,
 - Resentment and
 - Resistance.
- *CDOs need to rise to the challenge of changing the status quo if they expect to lead the business in making data a strategic asset.*
 - from *What Chief Data Officers Need to Do to Succeed* by Mario Faria

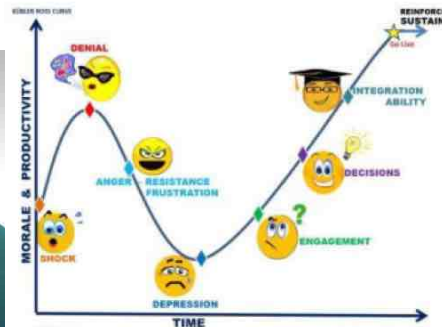


<https://www.forbes.com/sites/gartnergroup/2016/04/11/what-chief-data-officers-need-to-do-to-succeed/>

Change Management



INDIVIDUAL RESPONSE TO CHANGE



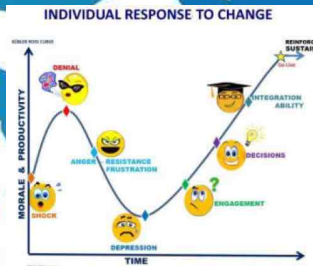
Diagnosing Organizational Readiness

adapted from the *Managing Complex Change* model by Dr. Mary Lippitt, 1987

	+	Skills	+	Incentive	+	Resources	+	Action Plan	=	Confusion
Vision	+		+	Incentive	+	Resources	+	Action Plan	=	Anxiety
Vision	+	Skills	+		+	Resources	+	Action Plan	=	Gradual Change
Vision	+	Skills	+	Incentive	+		+	Action Plan	=	Frustration
Vision	+	Skills	+	Incentive	+	Resources	+		=	False Starts
Vision	+	Skills	+	Incentive	+	Resources	+	Action Plan	=	Change

Culture is the biggest impediment to a shift in organizational thinking about data!

Change Management & Leadership



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 47



[My Author Page](#) [My Binders](#) [SIGN OUT:](#)
Dr. Peter Aiken

EXPERIENCE: Succeeding at Data Management—BigCo Attempts to Leverage Data

Full Text: [PDF](#)

Author: [Peter Aiken Virginia Commonwealth University/Data Blueprint](#)

Published in:

Journal
Journal of Data and Information Quality (JDIQ) - Challenge Papers,
Regular Papers and Experience Paper [JDIQ Homepage archive](#)
Volume 7 Issue 1-2, June 2016
Article No. 8
ACM New York, NY, USA
[table of contents](#) doi>[10.1145/2893482](#)

2016 Article
• Research
• Refereed

 Open Access

Bibliometrics
• Downloads (6 Weeks):
161
• Downloads (12 Months):
705
• Downloads (cumulative):
705
• Citation Count: 0

Tools and Resources

- [Buy this Article \(PRINT\)](#)
- [Recommend the ACM DL to your organization](#)
- [Request Permissions](#)
- TOC Service:**
 [Email](#) [RSS](#)
- [Save to Binder](#)
[View My Binders](#)
- Export Formats:**
[BibTeX](#) [EndNote](#)
[ACM Ref](#)

- Free Case Study Download

- <http://dl.acm.org/citation.cfm?doid=2888577.2893482>

or

<http://tinyurl.com/PeterStudy>

or scan the QR Code at the right



UNLOCKING BUSINESS VALUE

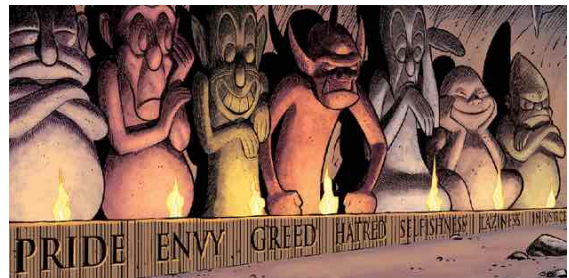
Copyright 2018 by Data Blueprint Slide # 48

Exorcising the Seven Deadly Data Sins

6 Not Sequencing Data Strategy Implementation

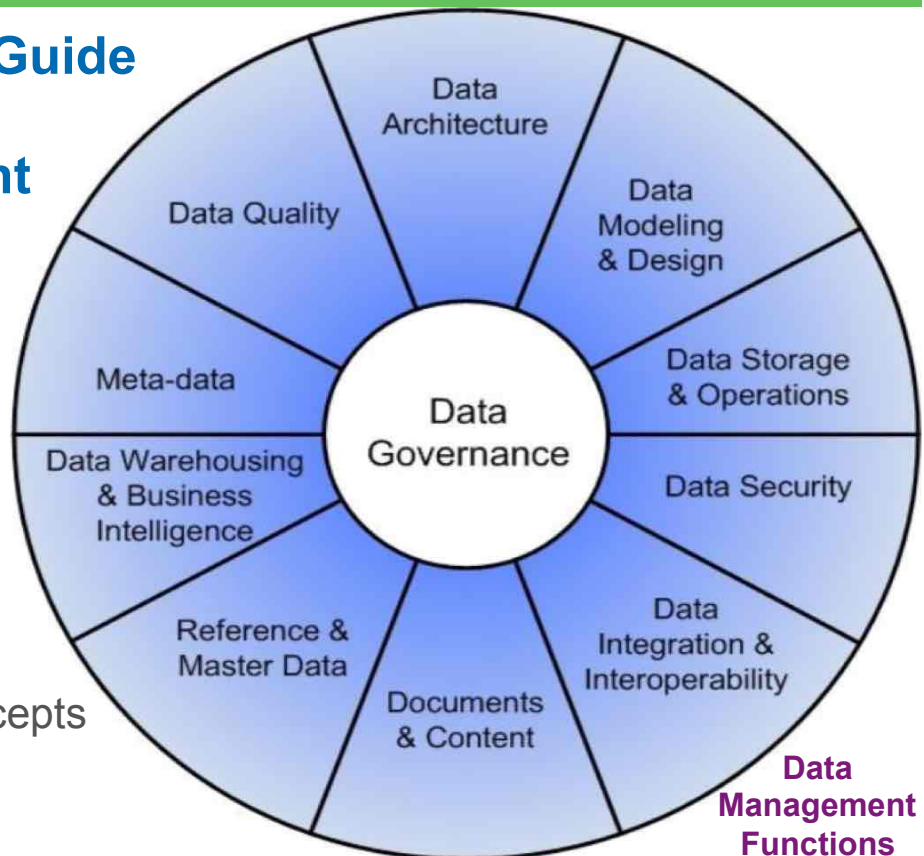


7 Failing To Address Cultural And Change Management Challenges



The DAMA Guide to the Data Management Body of Knowledge

- Good enough to criticize
 - All models are wrong
 - Some models are useful
- Missing two important concepts
 - Optionality
 - Dependency



Data Management Functions

Our barn had to pass a foundation inspection



- Before further construction could proceed
- No IT equivalent



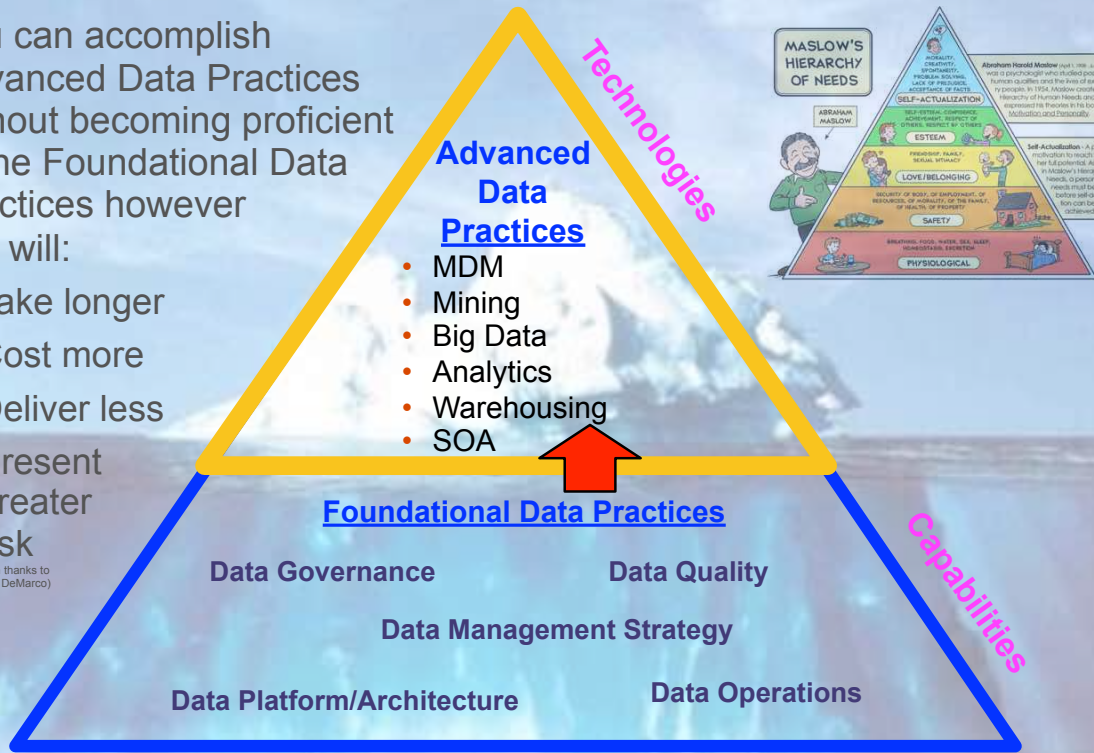
Separating the Wheat from the Chaff

Data Management Practices Hierarchy

You can accomplish Advanced Data Practices without becoming proficient in the Foundational Data Practices however this will:

- Take longer
- Cost more
- Deliver less
- Present greater risk

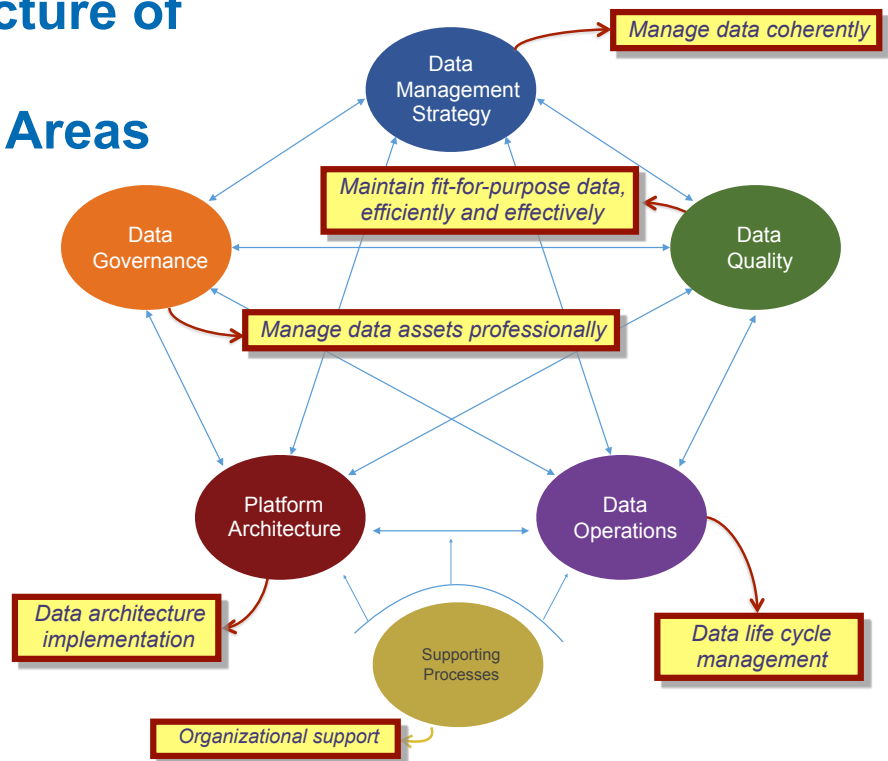
(with thanks to Tom DeMarco)



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 53

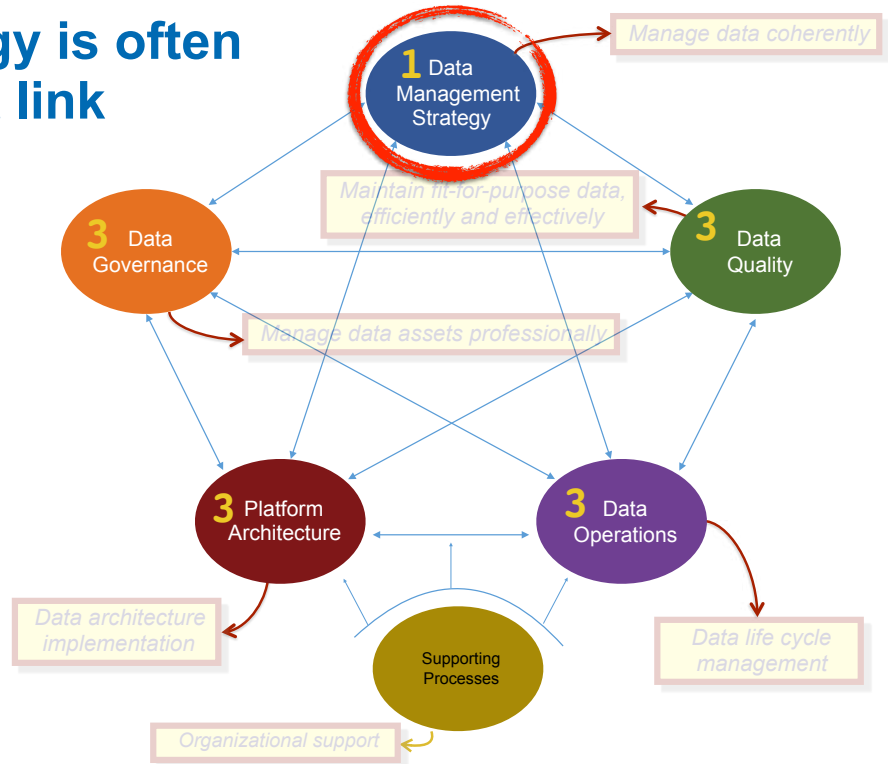
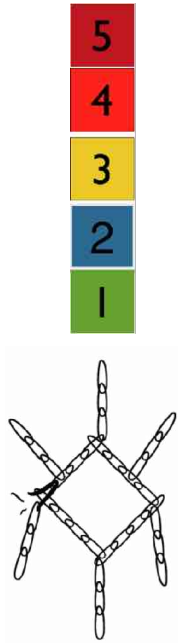
DMMSM Structure of 5 Integrated DM Practice Areas



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 54

Data Strategy is often the weakest link



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 55

~~Data Strategy Metrics Methodology~~

- Two Pet Peeves
 - Use "method" noun
a cause of annoyance:
 - not "methodology" noun (plural methodologies)
a system of methods used in a particular area of study or activity:
 - and use "measure" noun
 - not "metric" noun
1 technical a system or standard of measurement.
- Data Strategy Measures
 - Effectiveness
 - Over time
 - Volume (length)
 - Should be shorter than the organizational strategy
 - Versions
 - Should be sequential (with score keeping)
 - Understanding
 - Common agreement can be measured



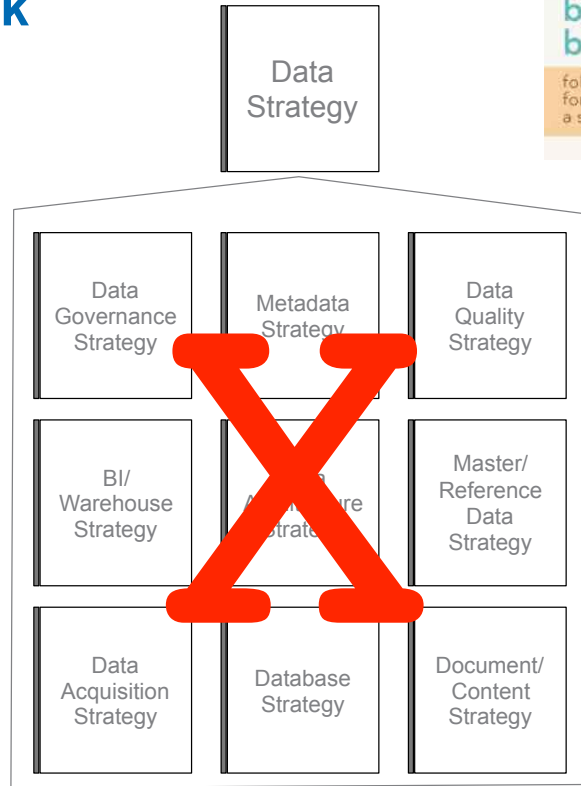
UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 55

By the Book

by the book

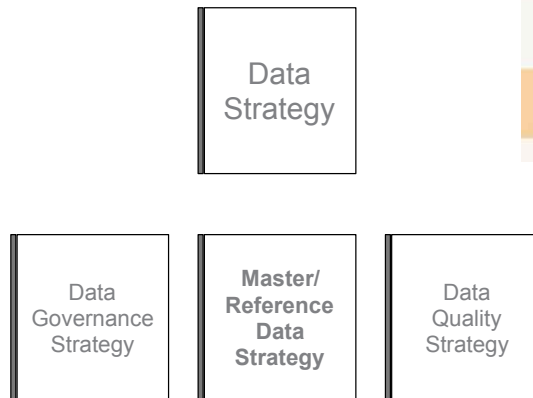
following all the rules for doing something in a strict way



Version 1

by the book

following all the rules for doing something in a strict way



Version 2

Data Strategy

by the book

following all the rules for doing something in a strict way



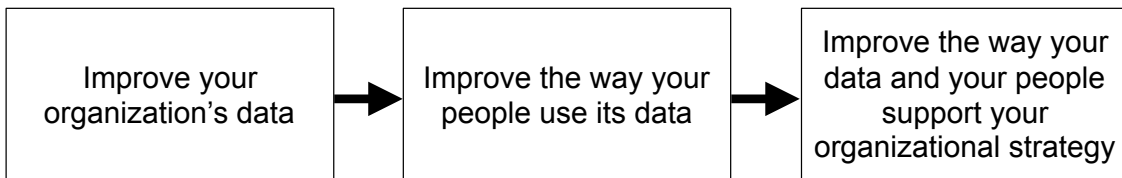
Data Governance Strategy

BI/ Warehouse Strategy

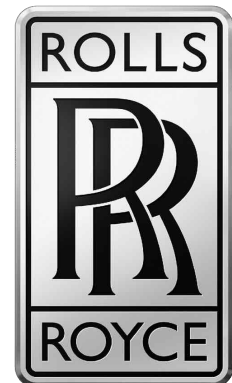
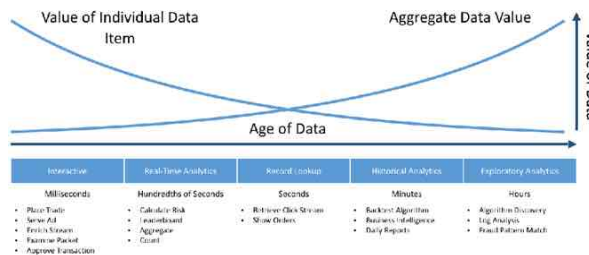
Data Acquisition Strategy



Reasons for a Data Strategy



- Because data points to where valuable things are located
- Because data has intrinsic value by itself
- Because data has inherent combinatorial value
- Valuing Data
 - Use data to measure change
 - Use data to manage change
 - Use data to motivate change
- Creating a competitive advantage with data



What did Rolls Royce Learn from *Formula 1*

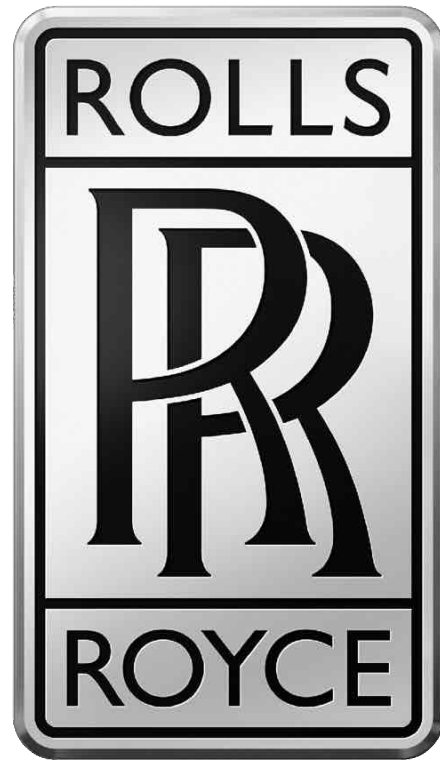
- Old model
 - Sell jet engines
- New model
 - Sell hours of thrust power
 - Power-by-the-hour
 - No payment for down time
 - Wing to wing
 - When was it invented?

Rolls-Royce celebrates 50th anniversary of Power-by-the-Hour

Tuesday, 30 October 2012

Rolls-Royce, the global power systems company, today celebrates the 50th anniversary of "Power-by-the-hour", its pioneering approach to engine maintenance management that forms the basis of the company's market-leading CorporateCare® services.

"Power-by-the-hour", a Rolls-Royce trademark, was invented in 1962 to support the Viper engine on the de Havilland Hawk. Subsequently, it was used on a fixed-wing aircraft engine and was the first engine service to be offered on a fixed-price, per-flight-hour basis. This aligned the interests of the manufacturer and operator, who only paid for engines that performed well.



The focus of data strategy should be sequenced



Exorcising the Seven Deadly Data Sins

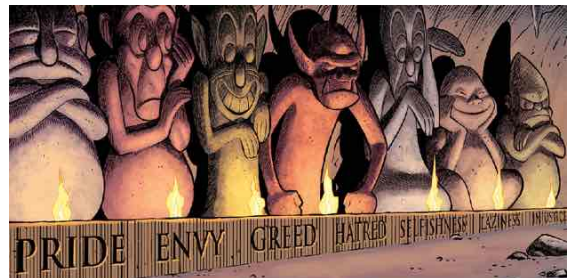
5 Failing to Adequately Manage Expectations



6 Not Sequencing Data Strategy Implementation

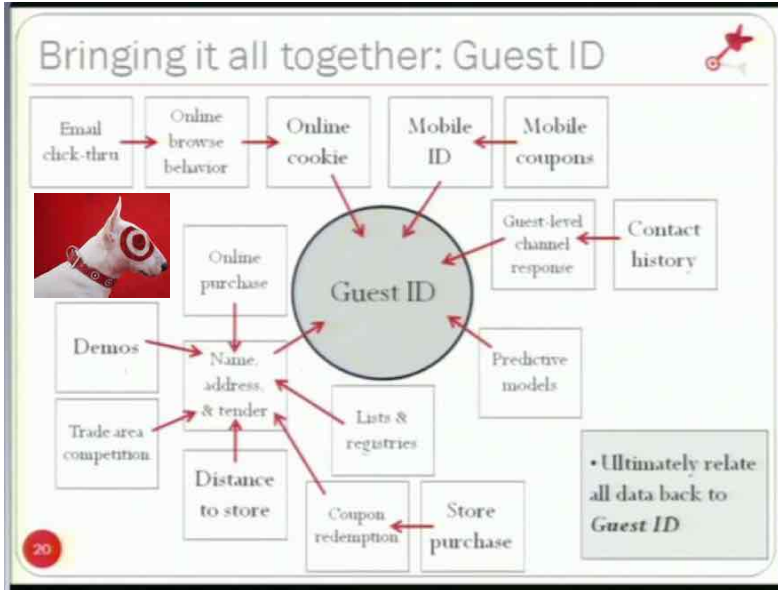


7 Failing To Address Cultural And Change Management Challenges



Target Isn't Just Predicting Pregnancies

<http://mportal.perfomedia.com/node/1372> and <http://www.predictusa.com/policies/target-calls-predict-beers-pregnancy-inside-story/> http://mportal.perfomedia.com/m/page/0/playlist_01f373



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 65

Target Corporations Database Contents



- Your **age**
- Marital **status**
- Part of town you live in
- How long it takes you to drive to work
- Estimated **salary**
- If you have recently **moved**
- Credit cards carried in your wallet
- What **websites** you visit
- Your ethnicity
- Your job history
- The magazines you read
- If you've ever declared **bankruptcy** or got divorced
- The year you bought (or **lost**) your house
- Where you went to college
- What kinds of topics you talk about **online**
- Whether you prefer certain brands of coffee, paper towels, cereal or applesauce
- Your political leanings, reading habits, charitable giving and
- The number of cars you own
- ...



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 66

Managing Assets?



ASHLEY MADISON®
Life is Short. Have an Affair.®



TARGET



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 67

Managing Data with Guidance?

- Federal employees
 - 44 users from whitehouse.gov
- Thousands of military and government e-mails
 - Canadian citizens
 - One-fifth of Quebec



TARGET



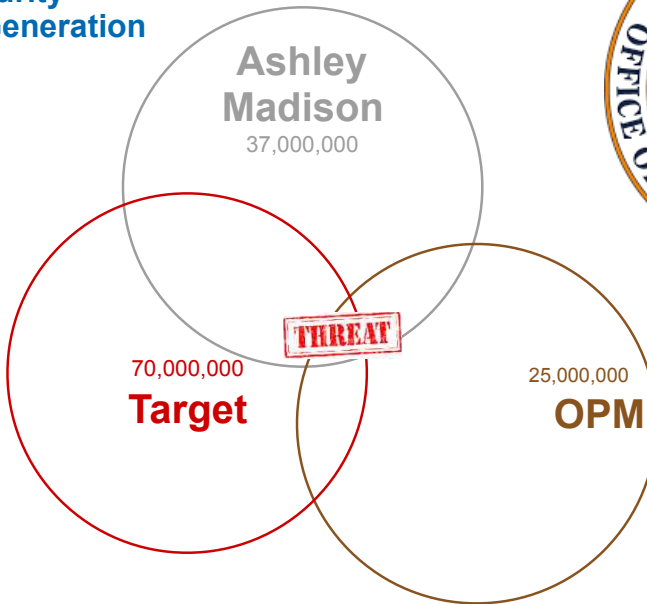
ASHLEY MADISON®
Life is Short. Have an Affair.®



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 68

How the Government Jeopardized
Our National Security
for More than a Generation



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 69

Committee on Oversight and Government Reform
U.S. House of Representatives
114th Congress



How the Government Jeopardized Our National
Security for More than a Generation

Majority Staff Report

Hon. Jason Chaffetz, Chairman
Committee on Oversight and Government Reform

Hon. Mark Meadows, Chairman
Subcommittee on Government Operations

Hon. Will Hurd, Chairman
Subcommittee on Information Technology

September 7, 2016

www.oversight.house.gov

<https://oversight.house.gov/report/opm-data-breach-government-jeopardized-national-security-generation/>



UNLOCKING BUSINESS VALUE

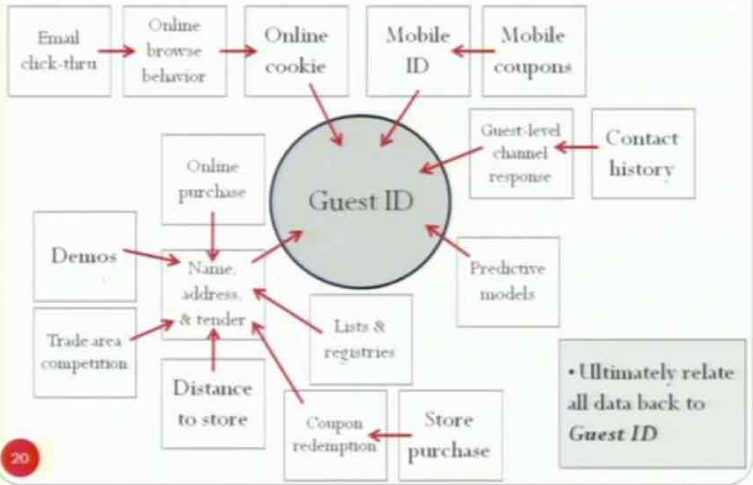
Copyright 2018 by Data Blueprint Slide # 70

Target Isn't Just Predicting Pregnancies



HUGGIES JEANS

Bringing it all together: Guest ID



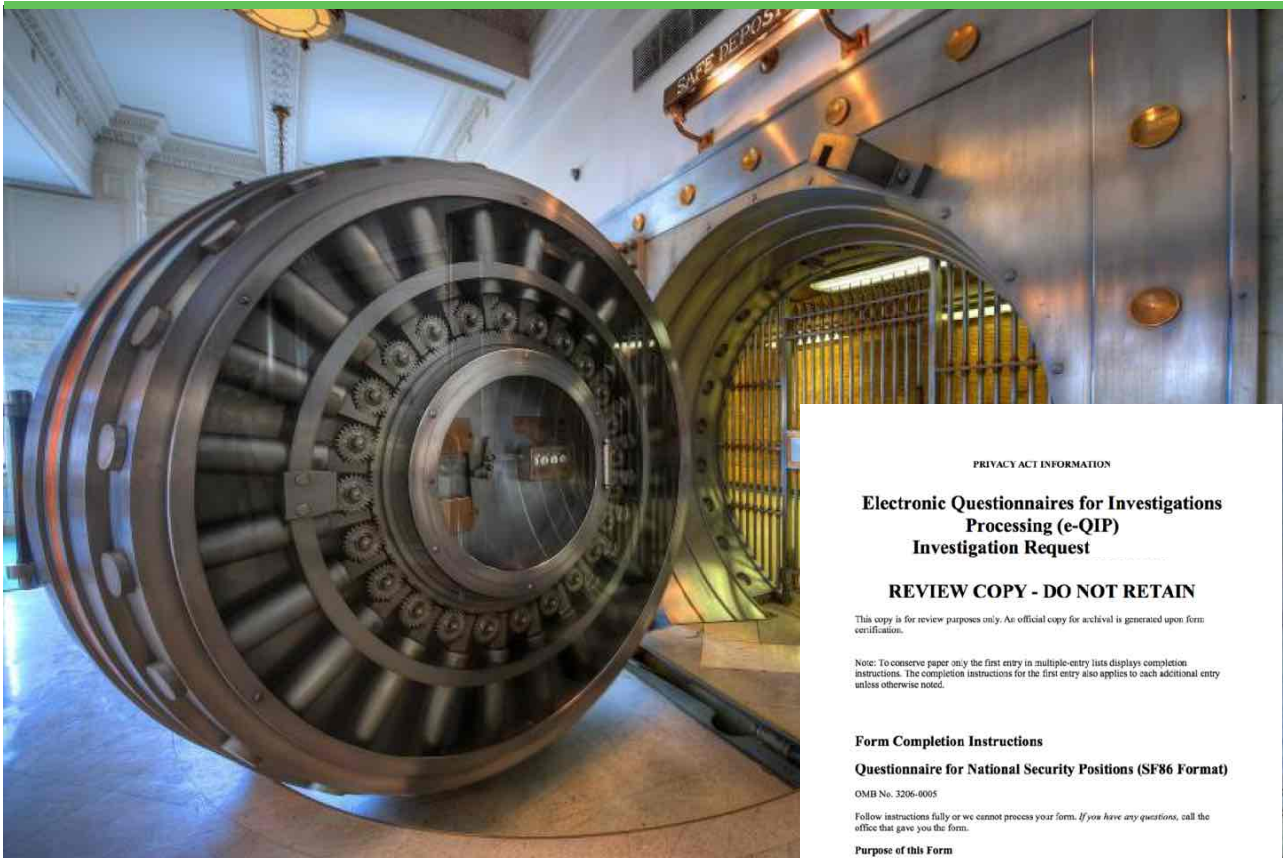
20

<http://rportal.performedia.com/node/1373>
<http://www.predictiveanalyticsworld.com/patimes/target-really-predict-teens-pregnancy-inside-story/>
http://rportal.performedia.com/rm/paw10/gallery_01#1373



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 71



PRIVACY ACT INFORMATION

Electronic Questionnaires for Investigations Processing (e-QIP) Investigation Request

REVIEW COPY - DO NOT RETAIN

This copy is for review purposes only. An official copy for archival is generated upon form verification.

Note: To conserve paper only the first entry in multiple-entry lists displays completion instructions. The completion instructions for the first entry also applies to each additional entry unless otherwise noted.

Form Completion Instructions

Questionnaire for National Security Positions (SF86 Format)

OMB No. 3206-0005

Follow instructions fully or we cannot process your form. If you have any questions, call the office that gave you the form.

Purpose of this Form



UNLOCKING BUSINESS VALUE

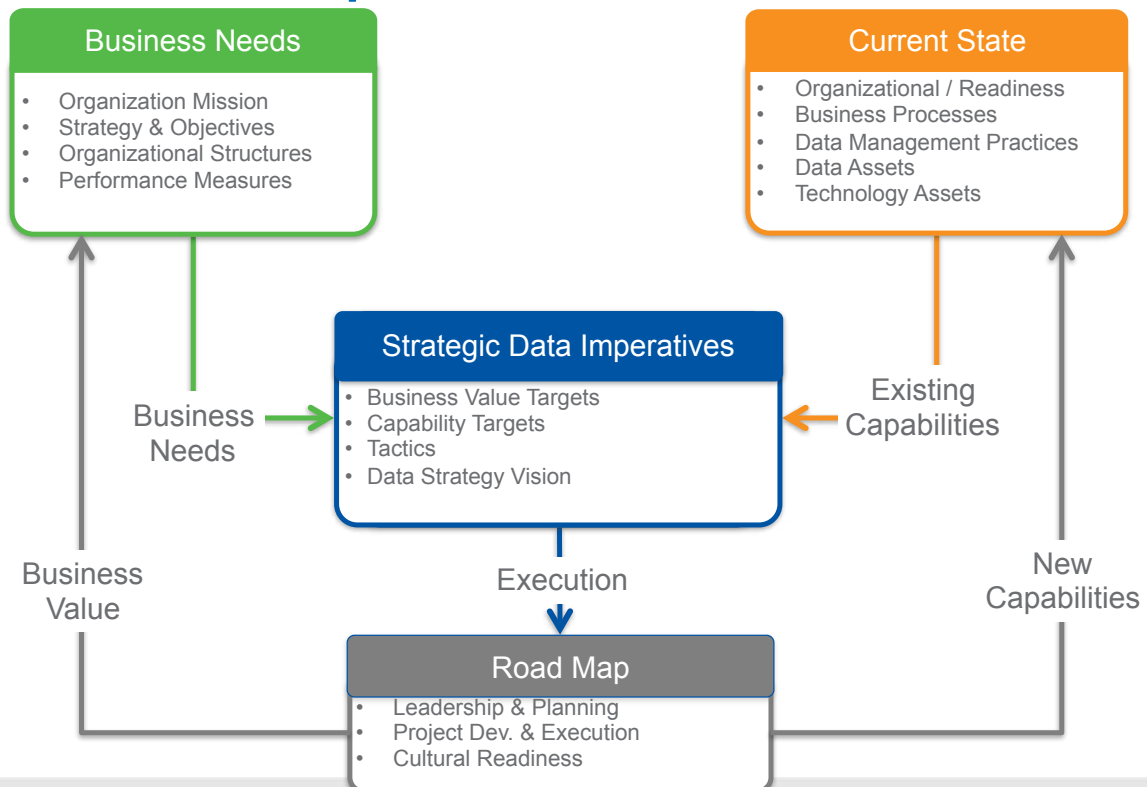
Copyright 2018 by Data Blueprint Slide # 72

Target Corporations Database Contents

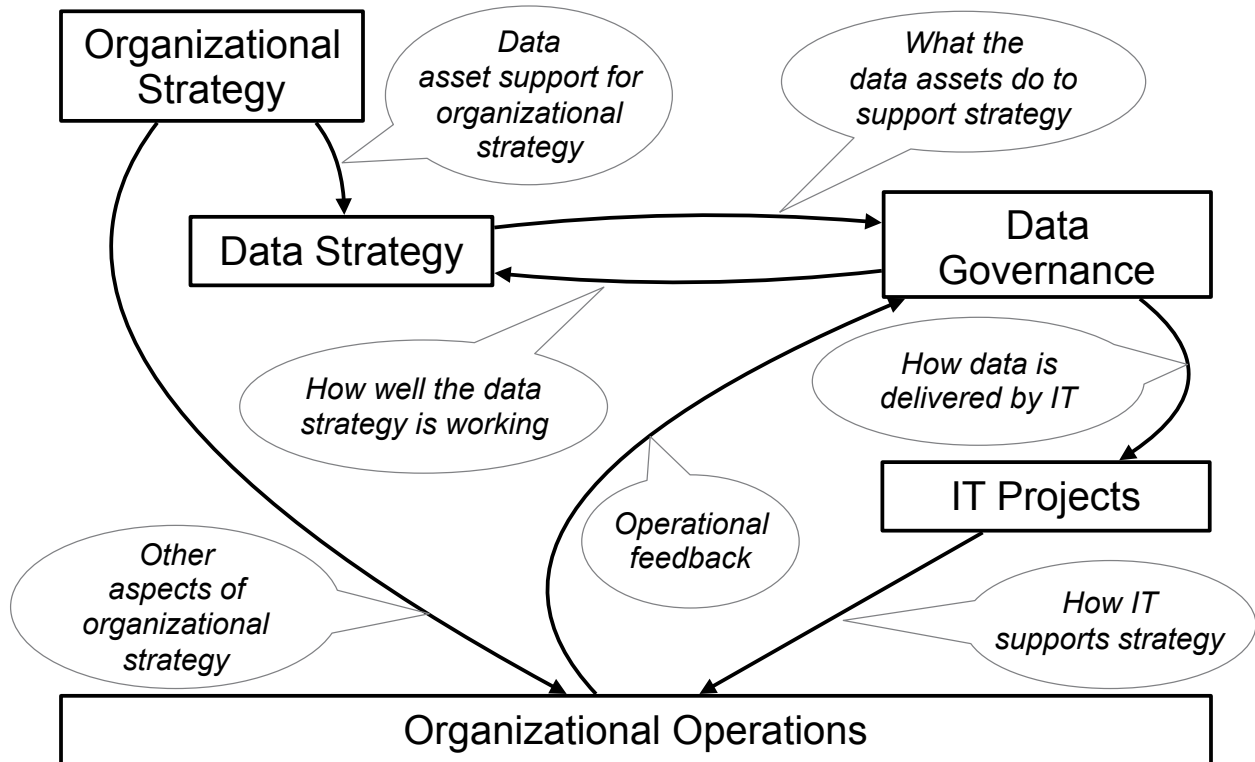
- Your age
- Marital status
- Part of town you live in
- How long it takes you to drive to work
- Estimated salary
- If you have recently moved
- Credit cards carried in your wallet
- What websites you visit
- Your ethnicity
- Your job history
- The magazines you read
- If you've ever declared bankruptcy or got divorced
- The year you bought (or lost) your house
- Where you went to college
- What kinds of topics you talk about online
- Whether you prefer certain brands of coffee, paper towels, cereal or applesauce
- Your political leanings, reading habits, charitable giving and
- The number of cars you own



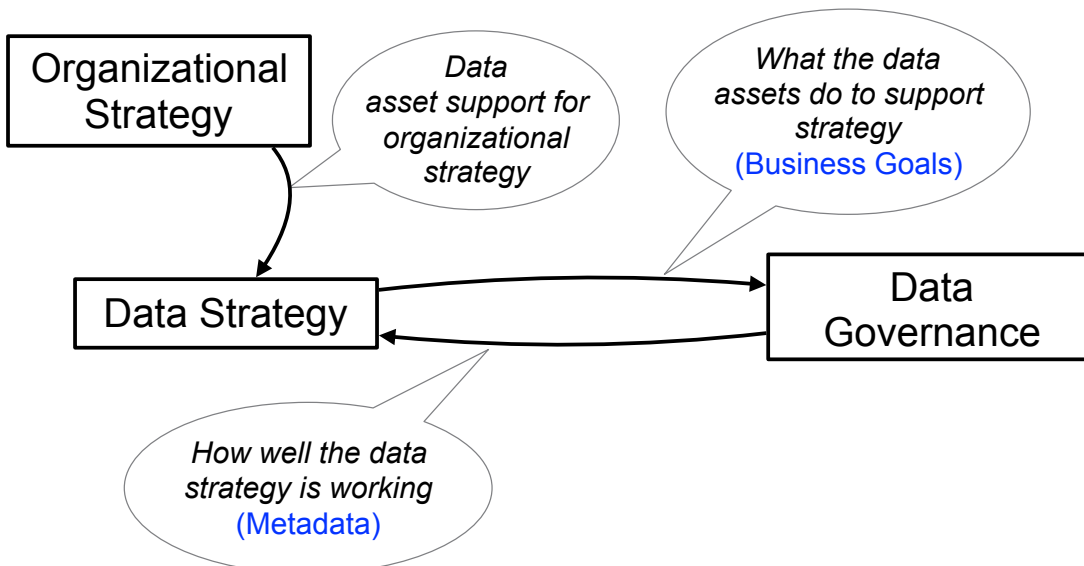
Data Implementation Framework



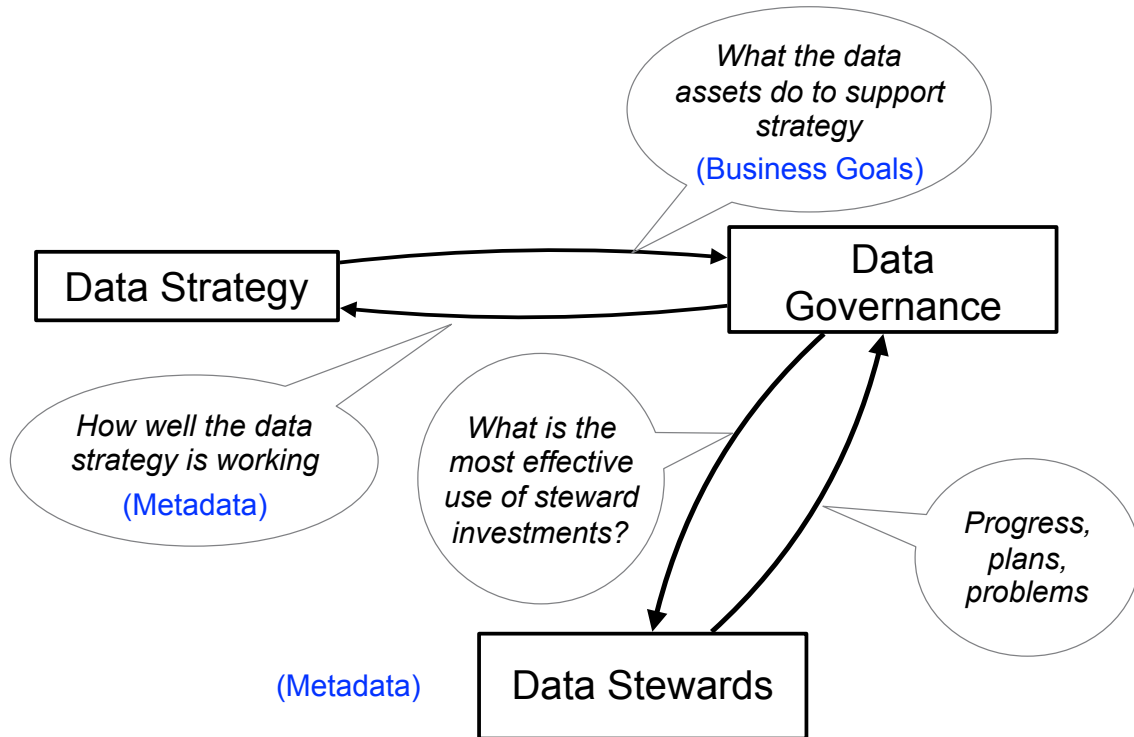
Data Strategy and Data Governance in Context



Data Strategy & Data Governance



Data Governance & Data Stewards



Data Management Program Expenses

- 5 Data Professionals
- \$100,000 each/annually
- When will you be done?
- *"It's okay my CIO gave me 5 years!"*



ESTIMATING

These are all questions we ask ourselves when we, as developers, are presented with a new problem and asked to provide an estimate as to how long it will take AND when it will be ready.



Virginia Governor's Data Interns Program



DATA

Virginia Internship Program Pits Grad Students Against Gov Data

For a second year, fresh sets of eyes and cutting-edge data analytics skills are the tools grad students will bring Virginia through the state's data internship program.

BY COLIN WOOD / AUGUST 25, 2015



Virginia Commonwealth University

FLICKR/ANDREW BAIN

Commonwealth of Virginia
 Office of Governor Terry McAuliffe

For Immediate Release
 July 23, 2015

Office of the Governor
 Contact: Brian Coy
 Email: Brian.Coy@governor.virginia.gov

Governor McAuliffe Announces 2015-16 Data Internships

~ Virginia Commonwealth University graduate student teams to explore the use of data to improve government efficiency ~

RICHMOND – Governor Terry McAuliffe today announced that Virginia state government and the Virginia Commonwealth University School of Business will again work together on data re-engineering internships to explore the use of data to improve the effectiveness and efficiency of state government.

In the 2014-2015 school year, the data internship program's first, 45 graduate students and more than 20 state agencies participated. Those internships have resulted in tangible dollar savings and improved agency processes. Student/agency teams have worked on successful projects, such as

Virginia Secretary of Technology Karen Jackson and CIO of the Commonwealth Nelson Moe are leading the effort on behalf of the state. Students who want to apply for internships should contact Peter Aiken (peter.aiken@vcu.edu) for additional information.

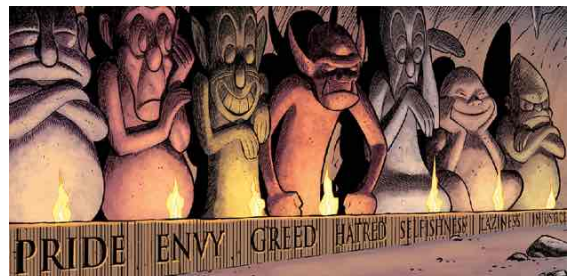


UNLOCKING BUSINESS VALUE

Copyright 2015 by Data Blueprint Slide # 79

Exorcising the Seven Deadly Data Sins

- 4 Not Aligning The Data Program with IT Projects
- 5 Failing to Adequately Manage Expectations
- 6 Not Sequencing Data Strategy Implementation
- 7 Failing To Address Cultural And Change Management Challenges



UNLOCKING BUSINESS VALUE

Copyright 2017 by Data Blueprint Slide # 80

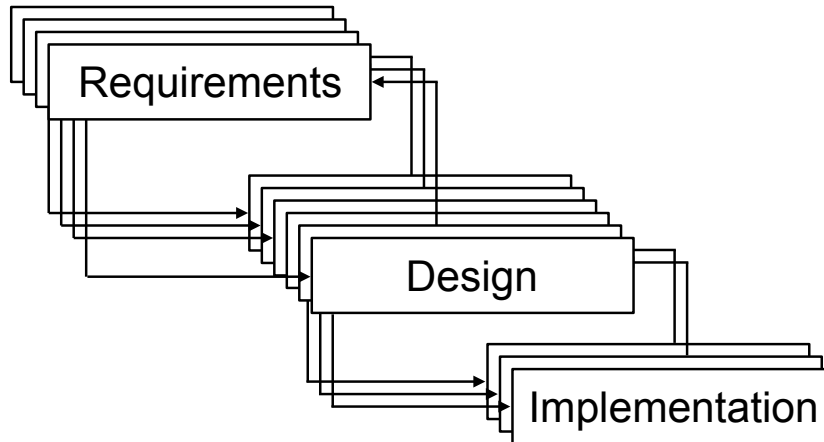
Differences between Programs and Projects

- Programs are Ongoing, Projects End
 - Managing a program involves long term strategic planning and continuous process improvement is not required of a project
- Programs are Tied to the Financial Calendar
 - Program managers are often responsible for delivering results tied to the organization's financial calendar
- Program Management is Governance Intensive
 - Programs are governed by a senior board that provides direction, oversight, and control while projects tend to be less governance-intensive
- Programs Have Greater Scope of Financial Management
 - Projects typically have a straight-forward budget and project financial management is focused on spending to budget while program planning, management and control is significantly more complex
- Program Change Management is an Executive Leadership Capability
 - Projects employ a formal change management process while at the program level, change management requires executive leadership skills and program change is driven more by an organization's strategy and is subject to market conditions and changing business goals



Adapted from http://top.idownloadnew.com/program_vs_project/ and <http://management.simplicable.com/management/new/program-management-vs-project-management>

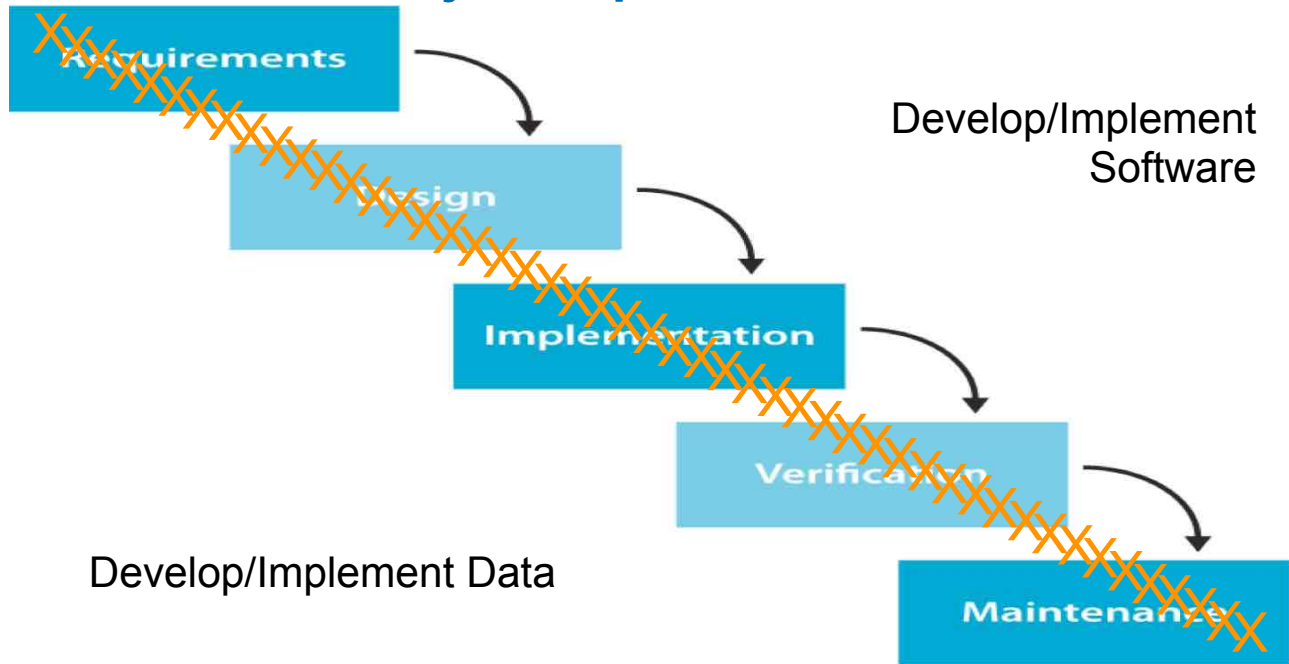
Systems Development (as described by Winston Royce in 1959)



Verification

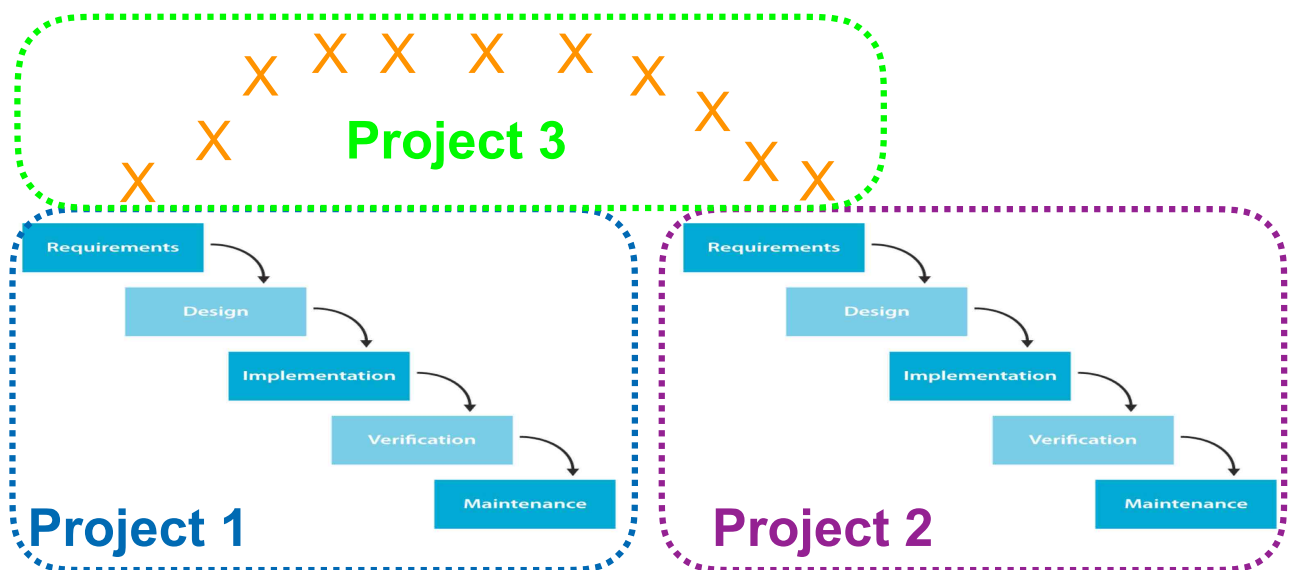
Maintenance

Project Implementation



This approach can only work when no sharing of data occurs!

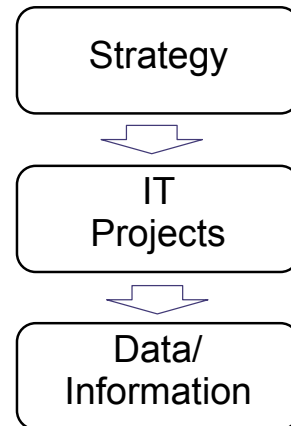
Projects Are Silos



Shared data structures require programmatic development and evaluation

IT Project or Application-Centric Development

- In support of strategy, organizations implement IT projects
- Data/information are typically considered within the scope of IT projects
- Problems with this approach:
 - Ensures data is formed to the applications and not around the organizational-wide information requirements
 - Process are narrowly formed around applications
 - Very little data reuse is possible



Original articulation from Doug Bagley @ Walmart

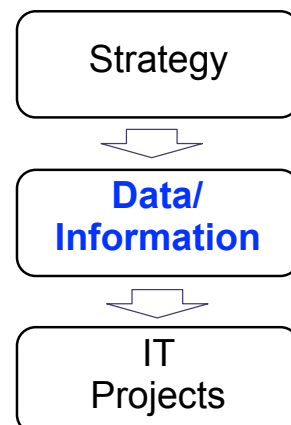


UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 85

Data-Centric Development

- In support of strategy, the organization develops specific, shared data-based goals/objectives
- These organizational data goals/objectives drive the development of specific IT projects with an eye to organization-wide usage
- Advantages of this approach:
 - Data/information assets are developed from an organization-wide perspective
 - Systems support organizational data needs and compliment organizational process flows
 - Maximum data/information reuse



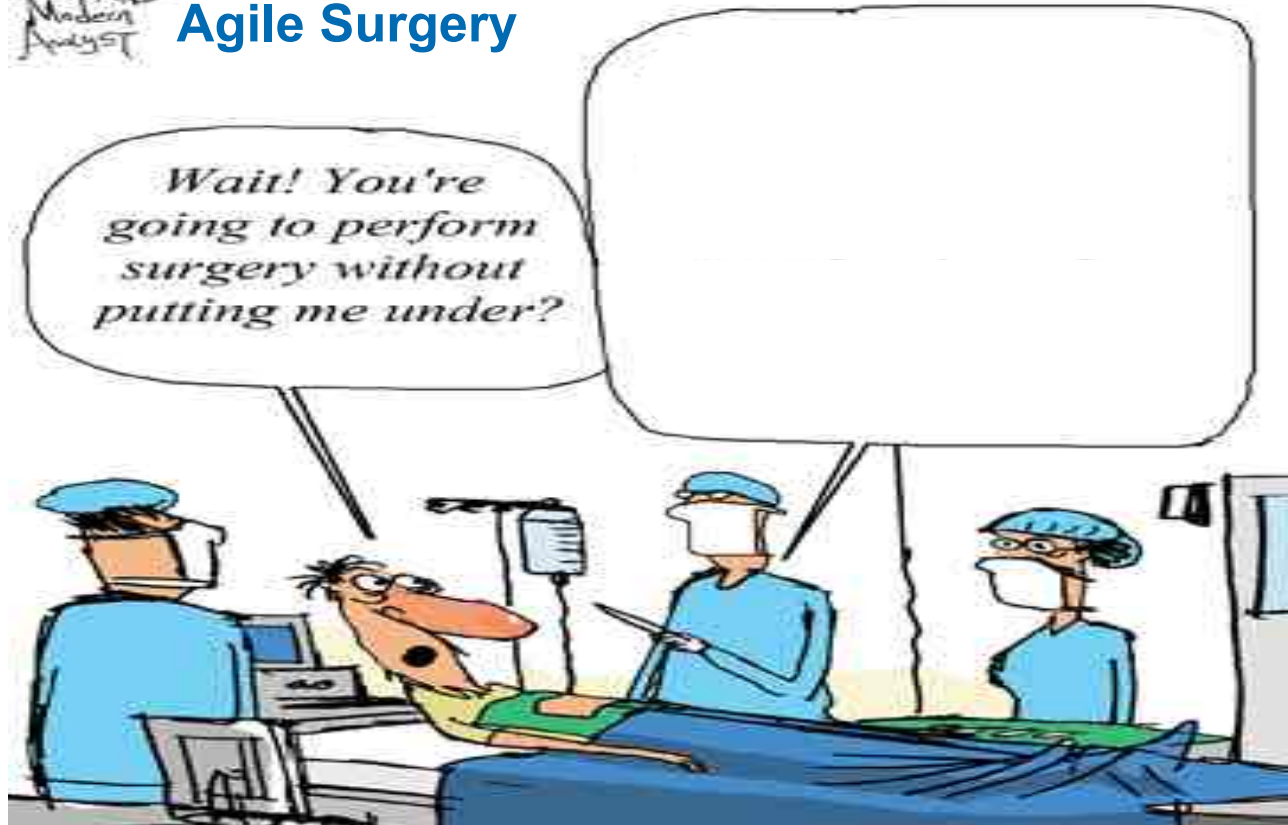
Original articulation from Doug Bagley @ Walmart








UNLOCKING BUSINESS VALUE

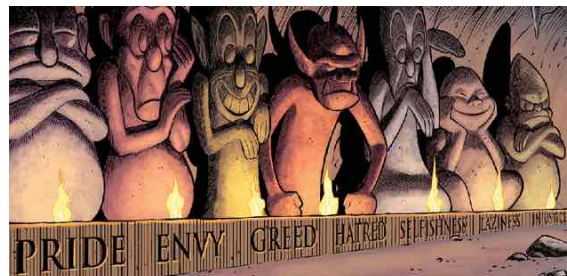
Copyright 2018 by Data Blueprint Slide # 86

Agile Surgery



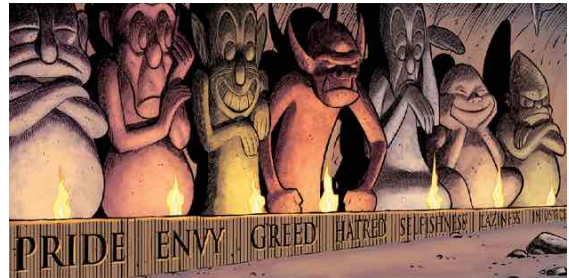
Exorcising the Seven Deadly Data Sins

-  3 Not implementing a Robust, Programmatic Means of Developing Shared Data
-  4 Not Aligning The Data Program with IT Projects
-  5 Failing to Adequately Manage Expectations
-  6 Not Sequencing Data Strategy Implementation
-  7 Failing To Address Cultural And Change Management Challenges



Exorcising the Seven Deadly Data Sins

- 2 Lacking Qualified Data Leadership
- 3 Not implementing a Robust, Programmatic Means of Developing Shared Data
- 4 Not Aligning The Data Program with IT Projects
- 5 Failing to Adequately Manage Expectations
- 6 Not Sequencing Data Strategy Implementation
- 7 Failing To Address Cultural And Change Management Challenges



What do we teach knowledge workers about data?

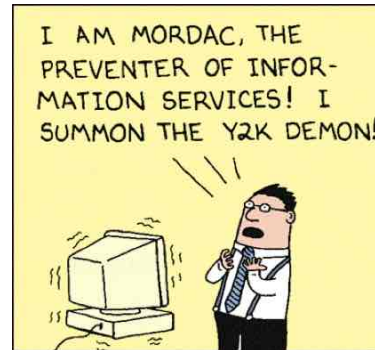
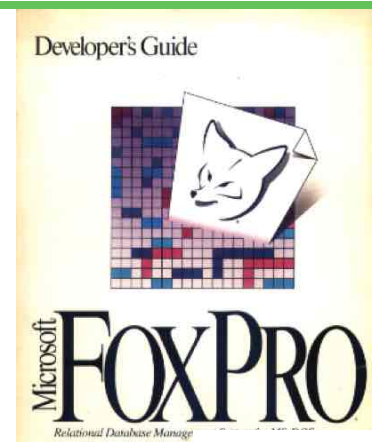


What percentage of the deal with it daily?

100%

Why should a knowledge worker

- with a PhD in Chemical Engineering
- have to know whether this product was Y2K compliant?



UNLOCKING BUSINESS VALUE

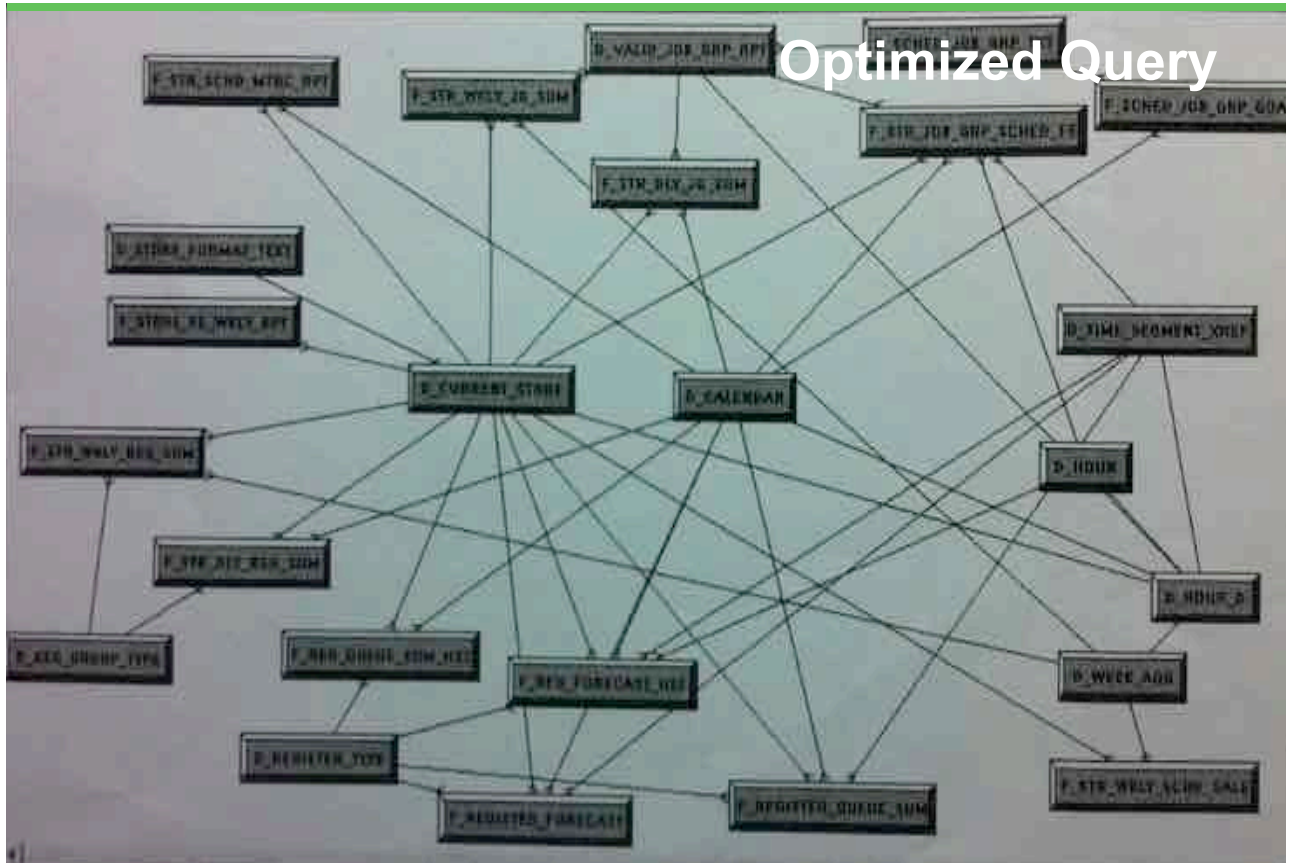
Copyright 2018 by Data Blueprint Side # 91



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Side # 92

Optimized Query



Repeat **100s**, thousands, millions of times ...

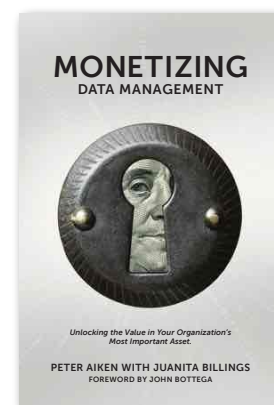




Data is a hidden IT Expense

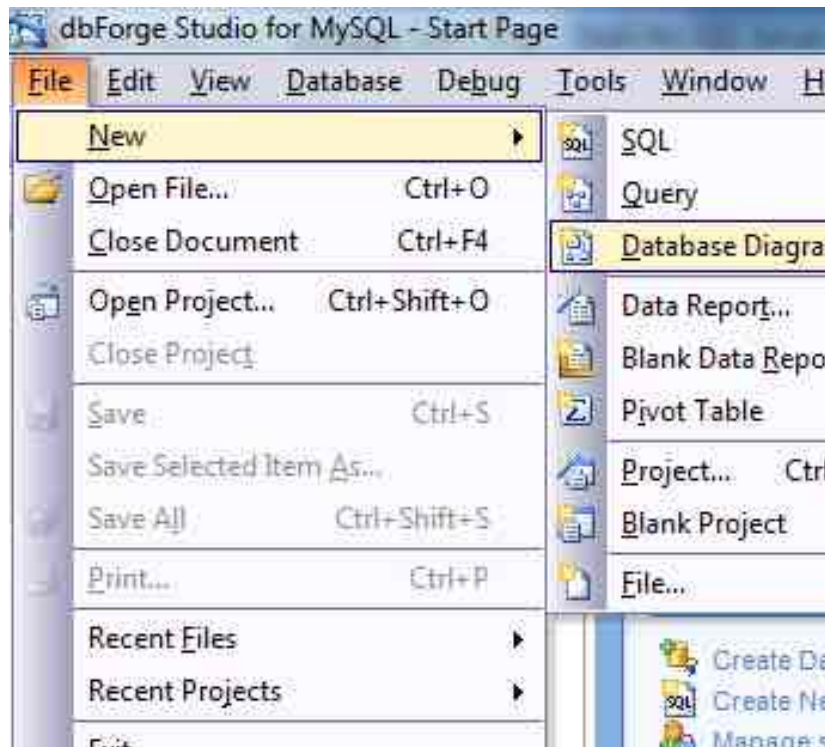
- Organizations spend between 20 - 40% of their IT budget evolving their data - including:
 - Data **migration**
 - Changing the location from one place to another
 - Data **conversion**
 - Changing data into another form, state, or product
 - Data **improving**
 - Inspecting and manipulating, or re-keying data to prepare it for subsequent use

– Source: *John Zachman*



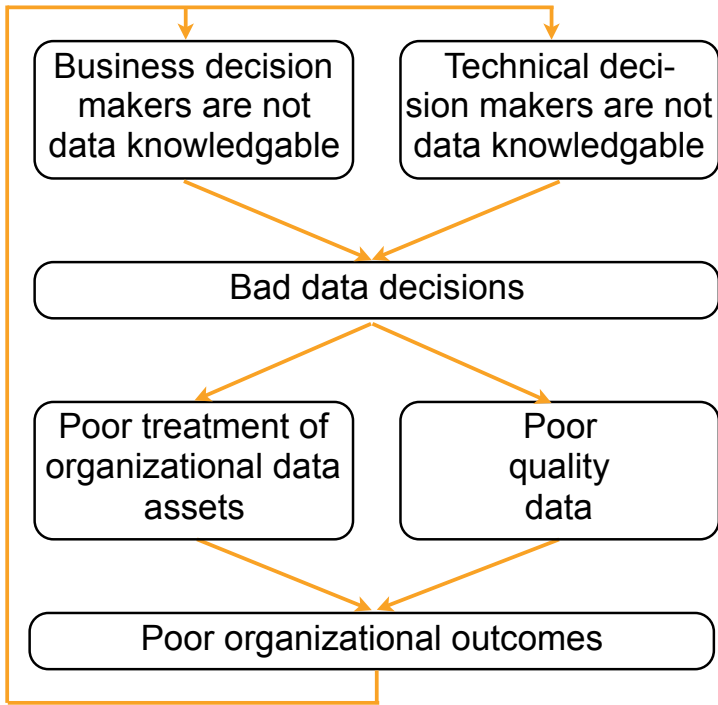
What do we teach IT professionals about data?

- 1 course
 - How to build a new database
- What impressions do IT professionals get from this education?
 - Data is a technical skill that is needed when developing new databases



If the only tool you know is a hammer you tend to see every problem as a nail (slightly reworded from Abraham Maslow)

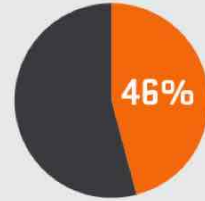
Bad Data Decisions Spiral



NEWS FLASH!

46% of companies report they made an inaccurate business decision based on bad or outdated data. Bad data leads to bad business decisions. Companies need to be careful that their data is sound – especially when dealing with investors.

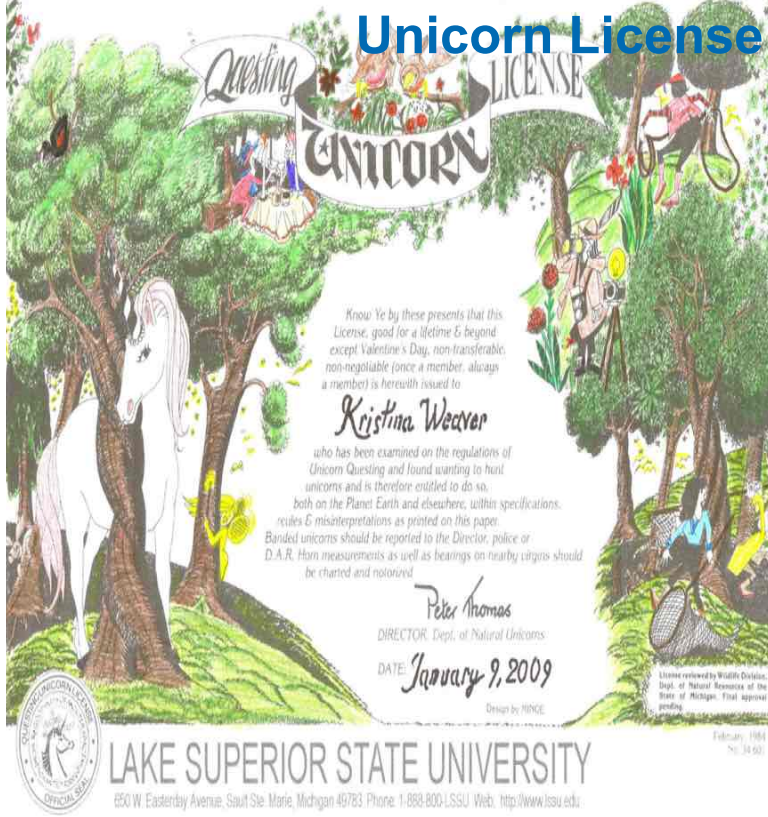
[Like](#) [Comment](#) [Share](#)



Hiring Panels Are Often Challenged to Help

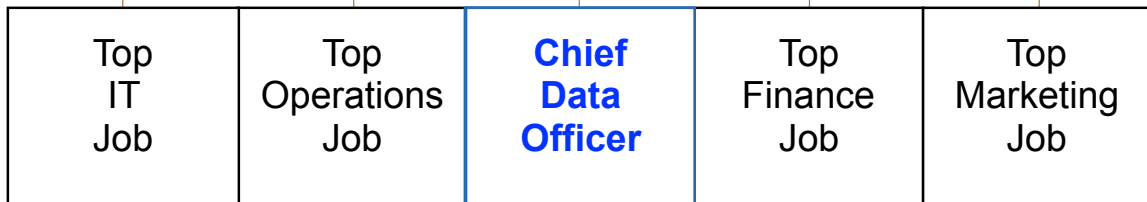


Unicorn License (There Are No Unicorns)



Top Data Job

Top Job



Data Governance Organization

- Dedicated solely to data asset leveraging
- Unconstrained by an IT project mindset
- Reporting to the business








• *90 Percent of Large Global Organizations Will Have Appointed Chief Data Officers By 2019*
 (Gartner website accessed January 26, 2016 <http://www.gartner.com/newsroom/id/3190117?>)

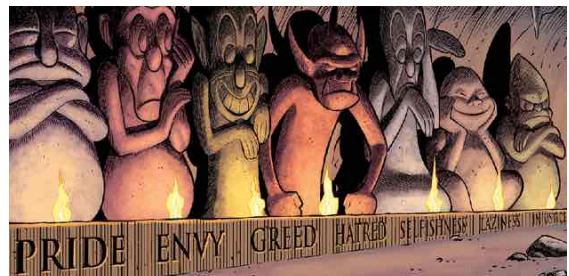




The Enterprise Data Executive Takes One for the Team

Exorcising the Seven Deadly Data Sins

-  Not Understanding Data-Centric Thinking
-  Lacking Qualified Data Leadership
-  Not implementing a Robust, Programmatic Means of Developing Shared Data
-  Not Aligning The Data Program with IT Projects
-  Failing to Adequately Manage Expectations
-  Not Sequencing Data Strategy Implementation
-  Failing To Address Cultural And Change Management Challenges



A Story from Michael Gorman

- An executive walks into an elevator
 - Michael has been trying to get an appointment
 - Trying to get an executive to validate a modeling concept
 - **Can a project have multiple Department owners?**
 - Executive: *I don't have time for technical issues*
 - Michael: *This isn't technical, it's policy*
 - Executive refuses to engage
- Mike (persistent) tries more several times but with the same result so he keeps looking for an answer
- Meets the DBA who is implementing the business rule
 - *Of course a project can have multiple Department owners*
 - Database: *Project, Department, Department Project*
- Next time Michael runs into the elevator executive
 - Michael: (says nothing)
 - Executive: *Thought you'd bother me about that 'technical issue?'*
 - Michael: *No, the DBA solved it. Projects have multiple Department owners.*
 - Executive: *No, that's wrong, that's not corporate policy*
 - Michael: *The dba is implementing so that Projects can have multiple Department owners. So, she's more important than you!*



Data is executed policy!



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 105



George Box
British Statistician
(1919-2013)

“All models are wrong, ...
... some are useful.”



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 106

We are uncovering better ways of developing IT systems by doing it and helping others do it. Through this work we have come to value:

Data programmes preceding software development

Stable data structures preceding stable code

Shared data preceding completed software

Data reuse preceding reusable code

That is, while there is value in the items on the right, we value the items on the left more.

Data programmes preceding software development

Common Organizational Data
(and corresponding data needs requirements)



Future State



Evolve

Data evolution is separate from, external to, and precedes system development life cycle activities!

Data management and software development must be separated and sequenced



Systems Development Activities

Build

New Organizational Capabilities



Data programmes preceding software development



theDataDoctrine.com



We are uncovering better ways of developing IT systems by doing it and helping others do it.
Through this work we have come to value:

Data programmes preceding software development

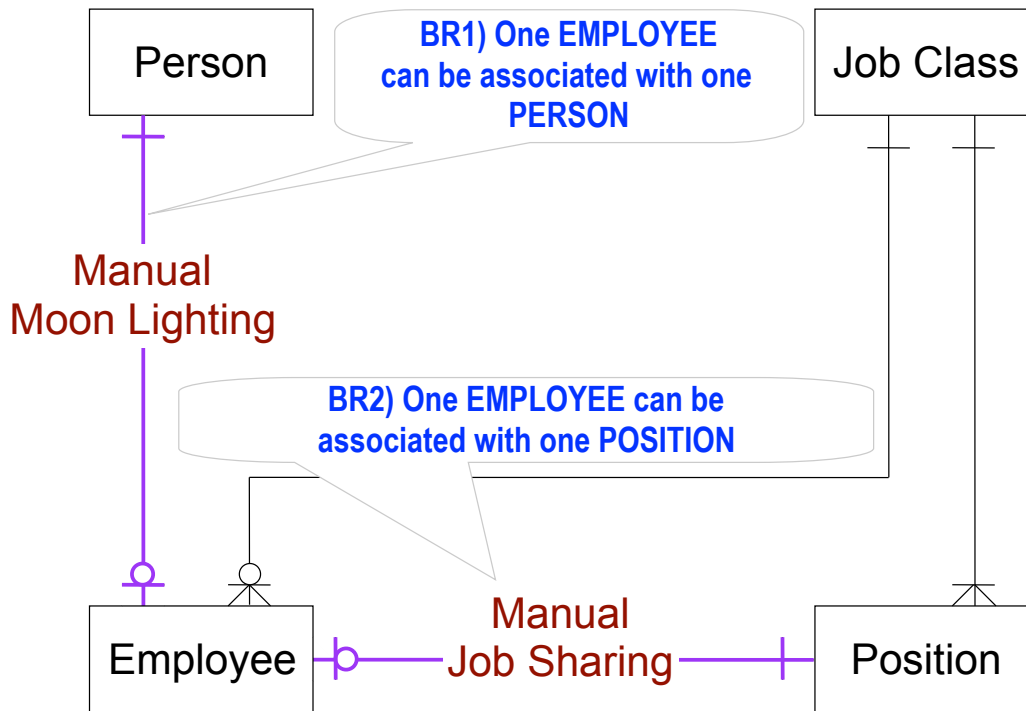
Stable data structures preceding stable code

Shared data preceding completed software

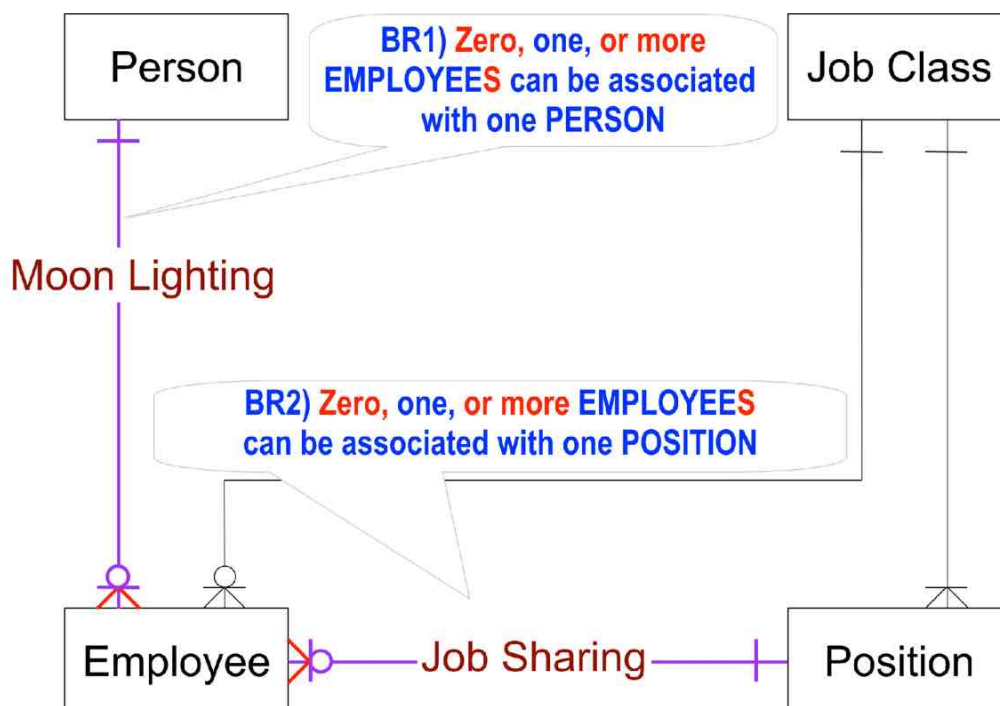
Data reuse preceding reusable code

That is, while there is value in the items on the right, we value the items on the left more.

Stable data structures preceding stable code



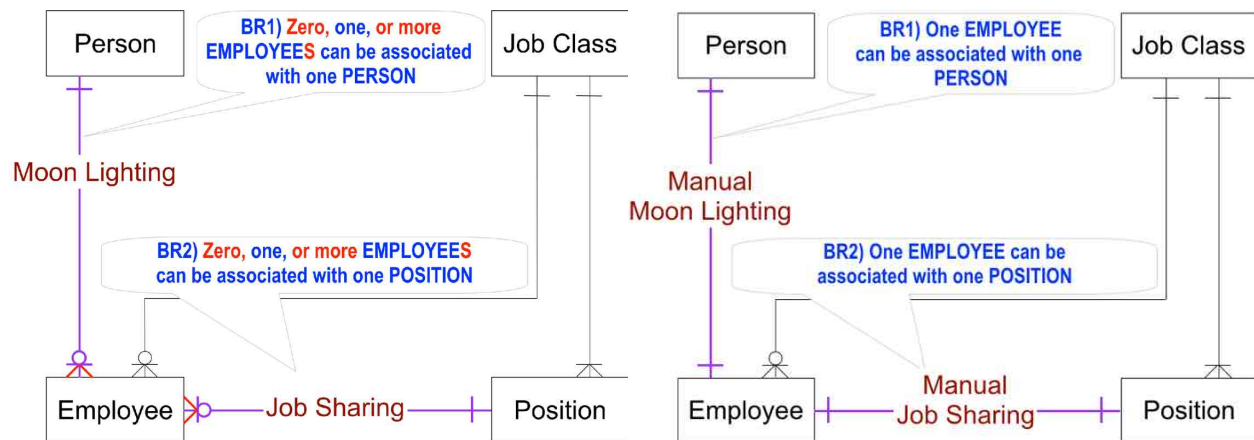
Stable data structures preceding stable code



Stable data structures preceding stable code

More flexible data structure

Less flexible data structure



(Requires 2 structural loops more than the more flexible data structure)

Data structures must be specified prior software development/acquisition

theDataDoctrine.com



We are uncovering better ways of developing IT systems by doing it and helping others do it. Through this work we have come to value:

Data programmes preceding software development

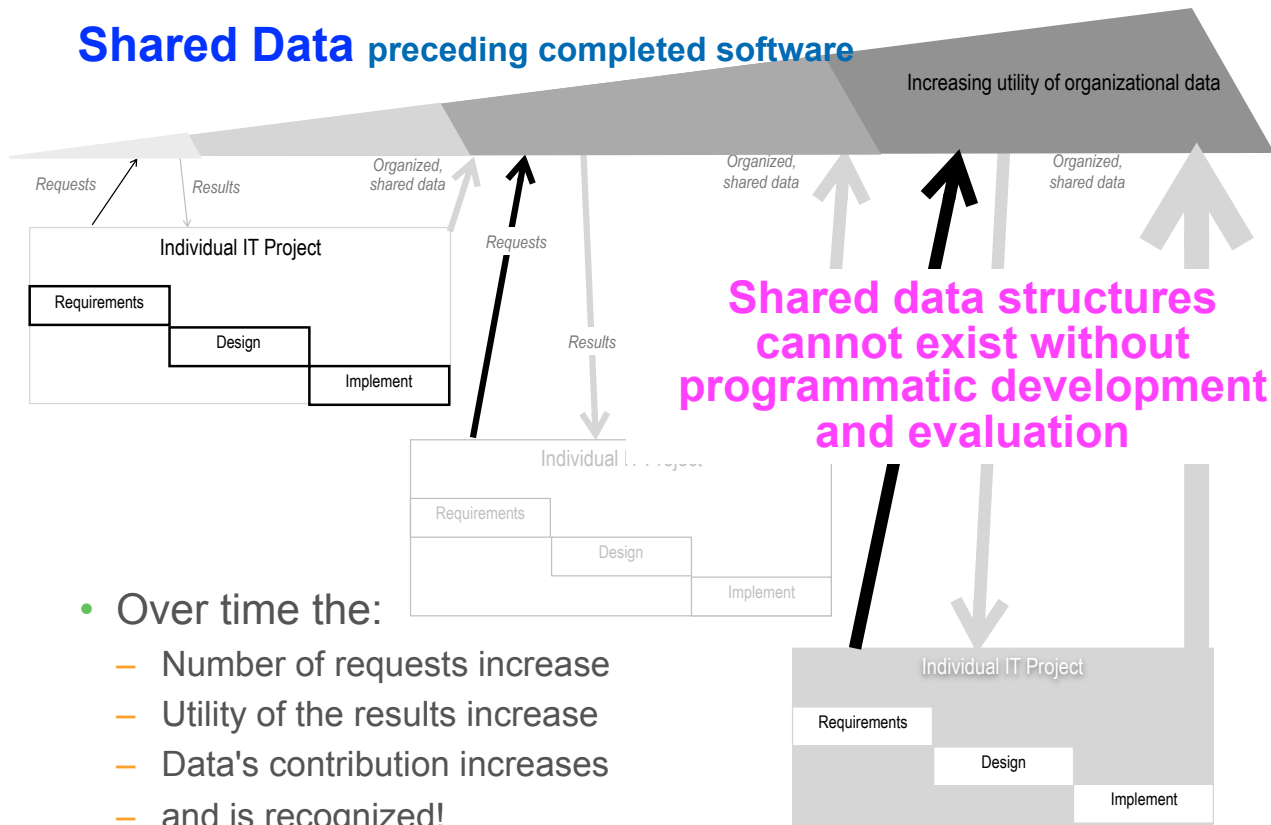
Stable data structures preceding stable code

Shared data preceding completed software

Data reuse preceding reusable code

That is, while there is value in the items on the right, we value the items on the left more.

Shared Data preceding completed software



theDataDoctrine.com



We are uncovering better ways of developing IT systems by doing it and helping others do it. Through this work we have come to value:

Data programmes preceding software development

Stable data structures preceding stable code

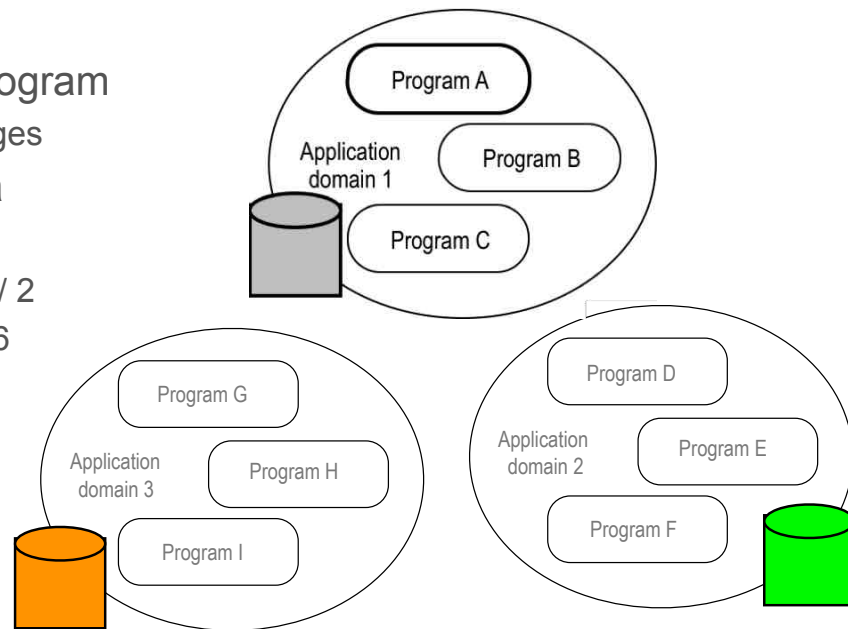
Shared data preceding completed software

Data reuse preceding reusable code

That is, while there is value in the items on the right, we value the items on the left more.

Data reuse preceding reusable code

- Reusable software has been valued more than reusable data
- Who makes decisions about the range and scope of common data usage?
- Change a program
 - 9 max changes
- Change data
 - Worst case
 - $(N * (N - 1)) / 2$
 - $(9 * 8) / 2 = 36$



theDataDoctrine.com



We are uncovering better ways of developing IT systems by doing it and helping others do it. Through this work we have come to value:

Data programmes preceding software development

Stable data structures preceding stable code

Shared data preceding completed software

Data reuse preceding reusable code

That is, while there is value in the items on the right, we value the items on the left more.

The Mission

These principles guide organizations attempting to better use data as an asset

What are the principles of Data-Centric Thinking?

Leaders can increase organizational effectiveness by focusing on data as a central, shared resource (or better still) as your sole, non-depletable, non-degrading, durable strategic asset. Data-centric thinking means practicing four data doctrine precepts.

Data-Centric Thinking Values: (In the spirit of the Agile Manifesto)

We are uncovering better ways of developing systems by doing it and helping others do it. Through this work we have come to value:

Data Programmes Preceding Software Projects
Stable Data Structures Preceding Stable Code
Shared Data Preceding Completed Software
Reusable Data Preceding Reusable Code

That is, while there is value in the items on the right, we value the items on the left more.

Agile Software Development isn't enough:

The Manifesto for Agile Software Development is an excellent step in the right direction, however more needs to be done to address the fundamental challenges facing IT. Systems are comprised of different components including people, processes, hardware, software, and data. Agile software development practices deliver better quality software products more rapidly. Better data products pervade and persist in all aspects of systems and increasingly benefit organizations.

The Need for the Data Doctrine:

In order for organizations to effectively incorporate data assets in support of organizational strategy, they need to establish a data management programme that is separate from, external to, and precedes software development projects!

- Data management and software development must be separated and sequenced
- Data structures must be stabilized before the software accessing them can be correctly constructed.
- Shared data structures require programmatic development and evaluation.
- Reusable data should be leveraged by reusable software.

Note: We are deliberately using the British spelling of the word programme to differentiate it from a software program - in this case used to denote an effort that is initiated and continues until the organization decides that it no longer needs to perform this kind of work, or the organization ceases to exist.

Look around and you can see the consequences of ignoring the foundational role that data plays in our organizations:

- Inadequate or nonexistent data education at all levels leads to knowledge workers under-appreciating the value of shared data assets. This, in turn, leads organizations to over rely on efforts such as software development.
- Lack of this data education leads organizations to omit data programmes and instead try to manage shared organizational data assets at the project level.
- Increase IT spending compensates for lack of data programmes. Organizations, consequently, spend resources on activities like integrating and cleaning up data and managing far more data than is necessary to manage strategically.

<http://www.thedatadoctrine.com>



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 119

Exorcising the Seven Deadly Data Sins

1

Not Understanding Data-Centric Thinking



2

Lacking Qualified Data Leadership



3

Not implementing a Robust, Programmatic Means of Developing Shared Data



4

Not Aligning The Data Program with IT Projects



5

Failing to Adequately Manage Expectations



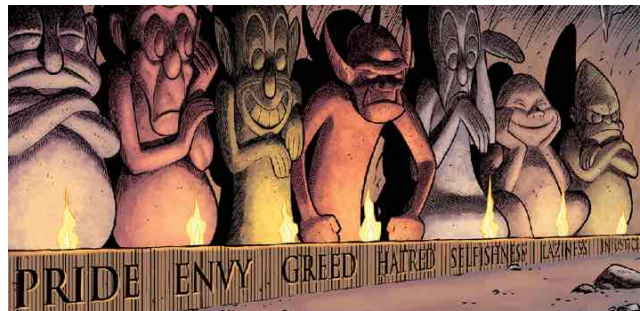
6

Not Sequencing Data Strategy Implementation



7

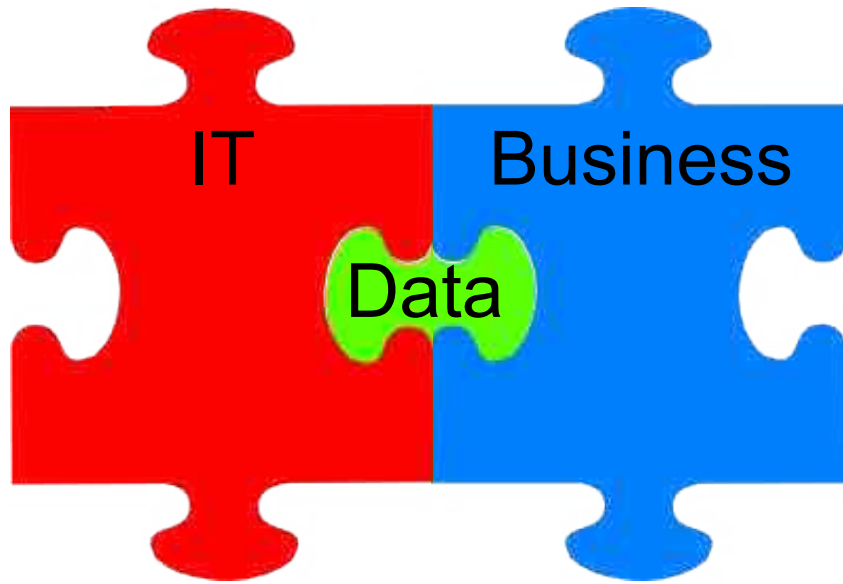
Failing To Address Cultural And Change Management Challenges



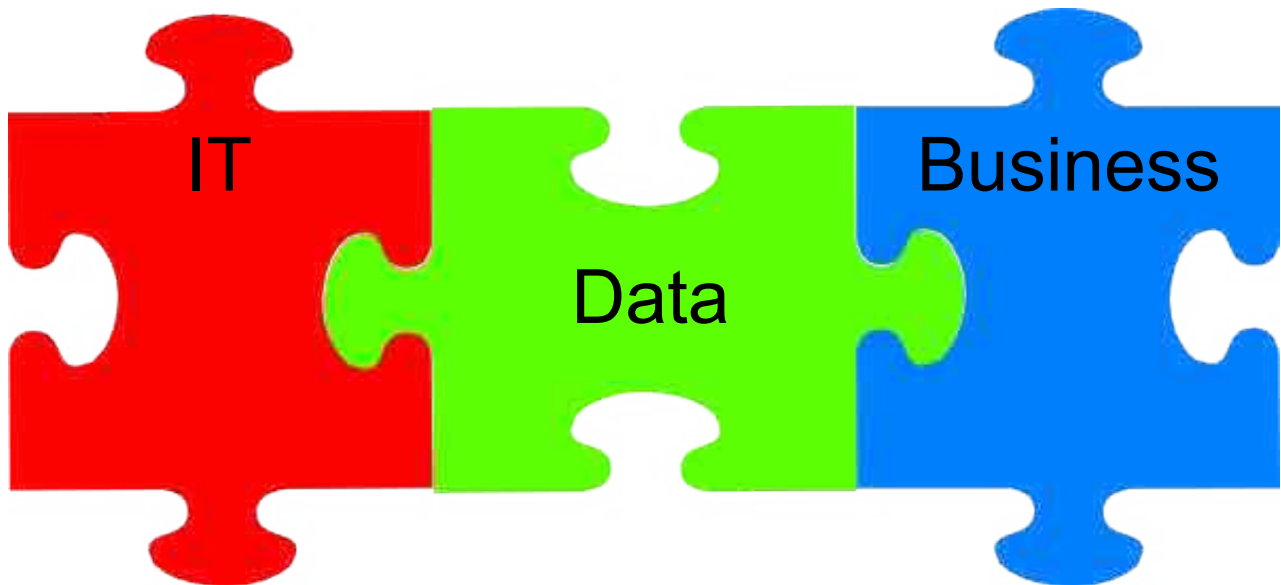
UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 120

As Is State of Data (as Perceived)



Desired To Be State of Data (as Understood)



YOUR DATA STRATEGY

Making It Concise, Actionable, and Understandable by Business and IT!

Peter H. Aiken • Todd Harbour
Foreword by Micheline Casey

Digital Literacy Series



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 123

Questions?



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 124

The Abbreviated State



An imaginary documentary about abbreviating all 50 States down to two letter codes made up by Gary Gulman for this 7-13-2016 stand-up routine on CONAN • <https://www.youtube.com/watch?v=dLECCmKnrys>



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 125



UNLOCKING BUSINESS VALUE



datablueprint.com

10124 W. Broad Street, Suite C
Glen Allen, Virginia 23060
804.521.4056

Copyright 2018 by Data Blueprint Slide # 125