



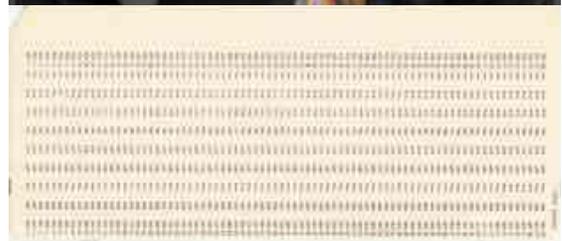
# Exorcising the Seven Deadly Data Sins

Peter Aiken, PhD

[datablueprint.com](http://datablueprint.com)

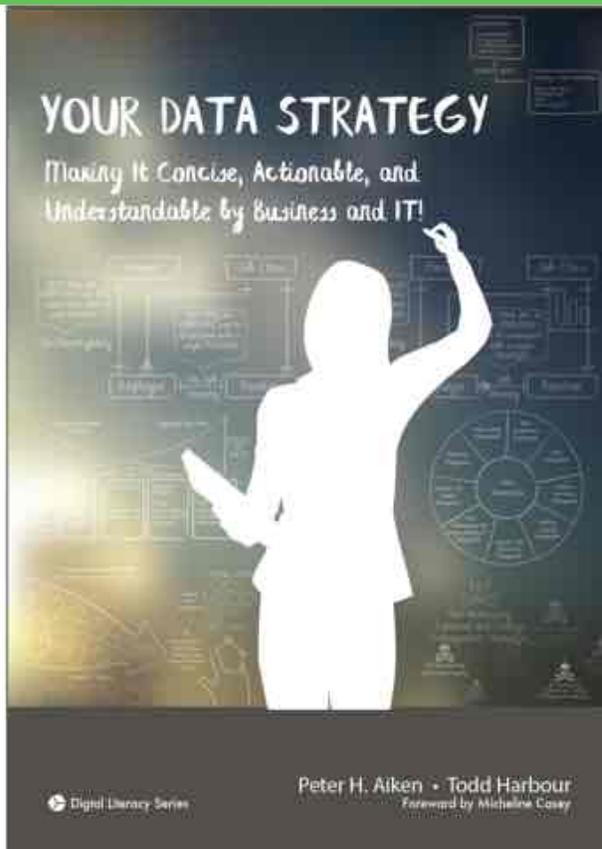
## Peter Aiken, Ph.D.

- 33+ years in data management
- Repeated international recognition
- Founder, Data Blueprint ([datablueprint.com](http://datablueprint.com))
- Associate Professor of IS ([vcu.edu](http://vcu.edu))
- DAMA International ([dama.org](http://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
  - ...



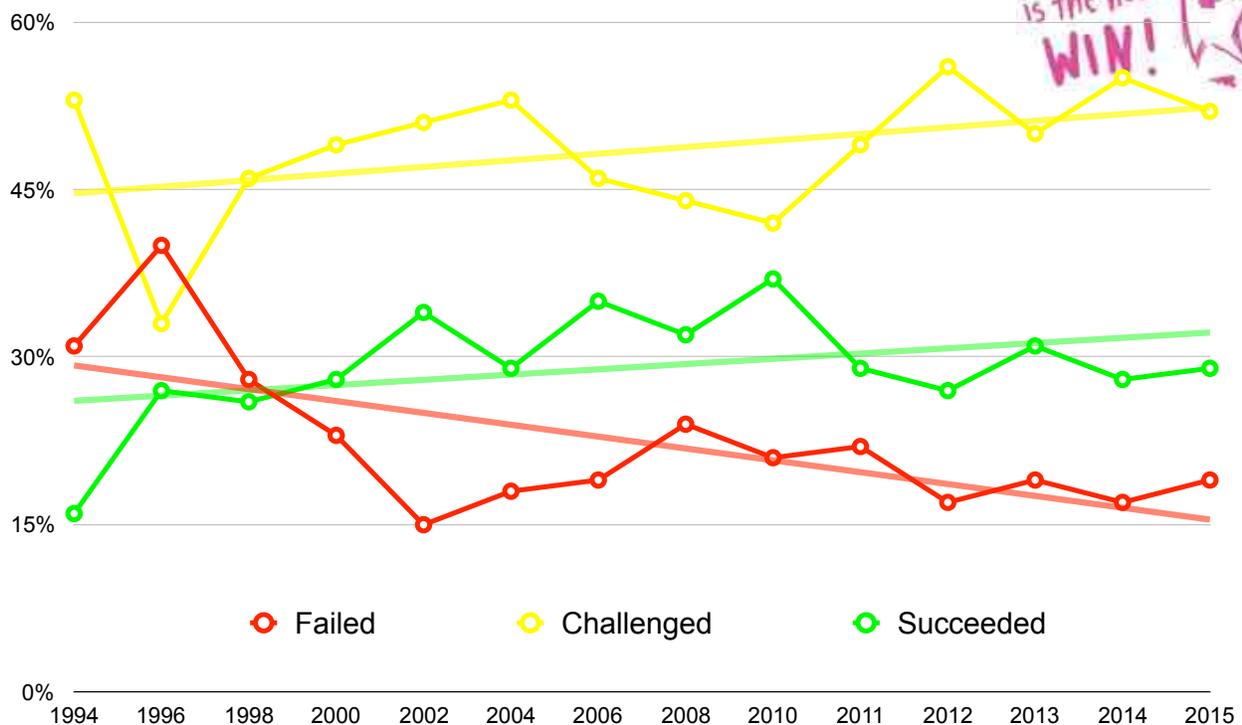
UNLOCKING BUSINESS VALUE

Excerpted from  
**Your Data Strategy**



Source: Standish Chaos Reports as reported at: <http://standishgroup.com>

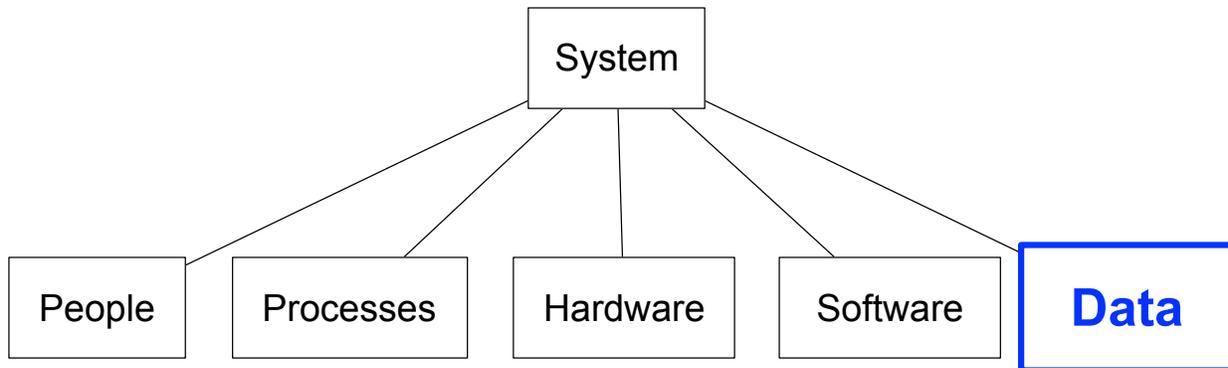
## IT Project Failure Rates (1994-2015)



# 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>



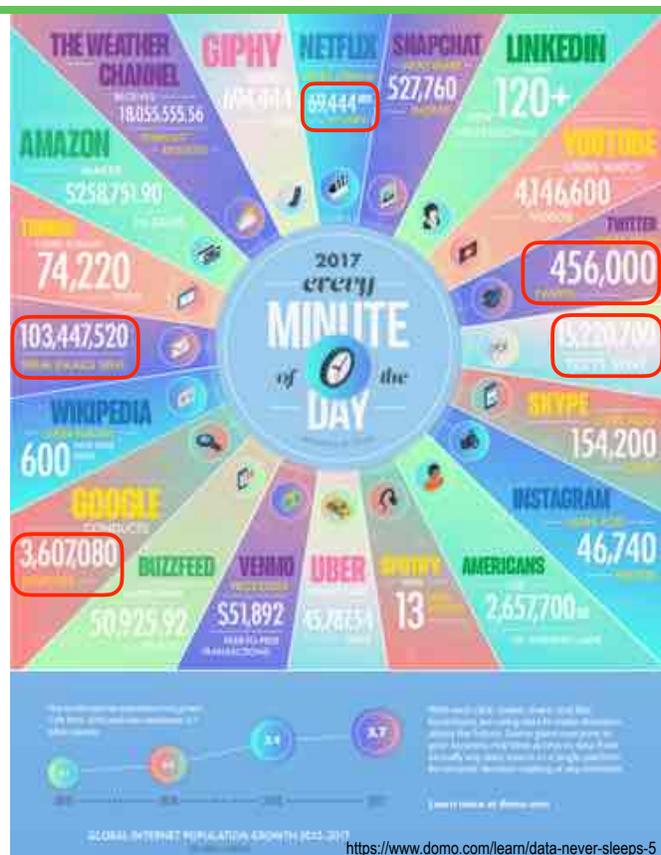
UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 5

## 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 # 6

# 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

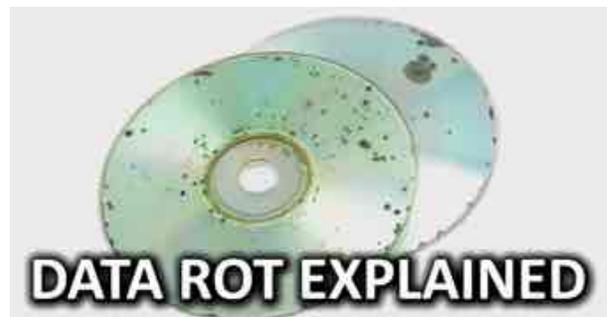
Copyright 2018 by Data Blueprint Slide # 7

## 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**



- **R**edundant
- **I**ncomplete
- **O**bsolute
- **T**rivial



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 8



As articulated by Micheline Casey



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 9

## Exorcising the Seven Deadly Data Sins



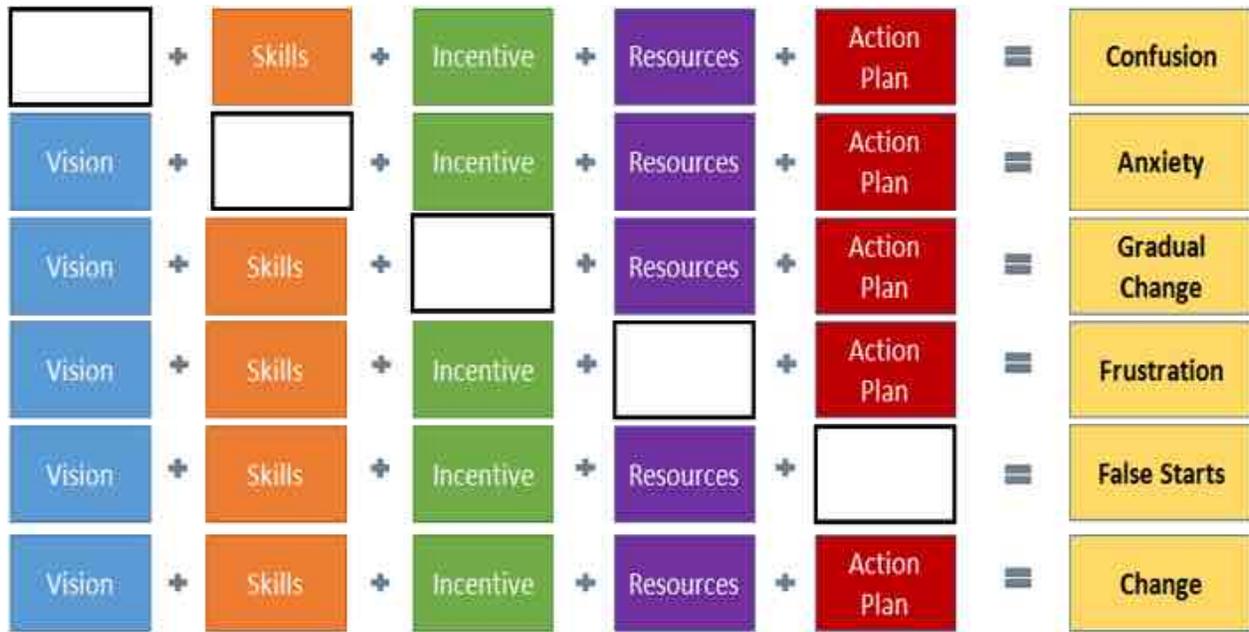
Failing To Address  
Cultural And Change  
Management Challenges



UNLOCKING BUSINESS VALUE

Copyright 2017 by Data Blueprint Slide # 10

## Diagnosing Organizational Readiness



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

## 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





# Exorcising the Seven Deadly Data Sins

**6** Not Sequencing Data Strategy Implementation

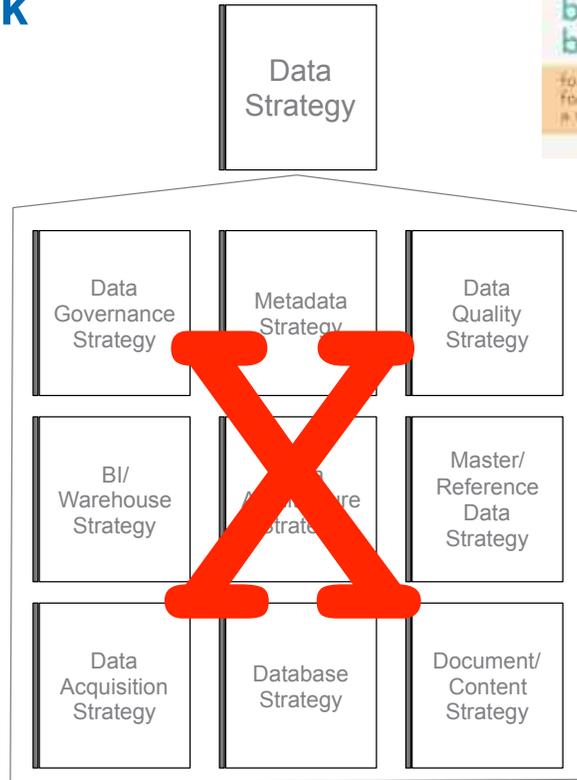
**7** Failing To Address Cultural And Change Management Challenges



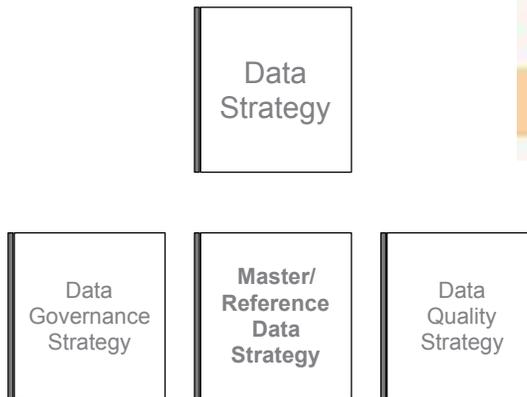
## The focus of data strategy should be sequenced



# By the Book



# Version 1



## 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



## ~~Data Strategy Metrics Methodology~~

- Two Pet Peeves **noun**
  - Use "method" **a cause of annoyance:**
  - not "methodology" **noun (plural methodologies)**  
and use "measure" **a system of methods used in a particular area of study or activity:**
  - 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



# Exorcising the Seven Deadly Data Sins

**5** Failing to Adequately Manage Expectations



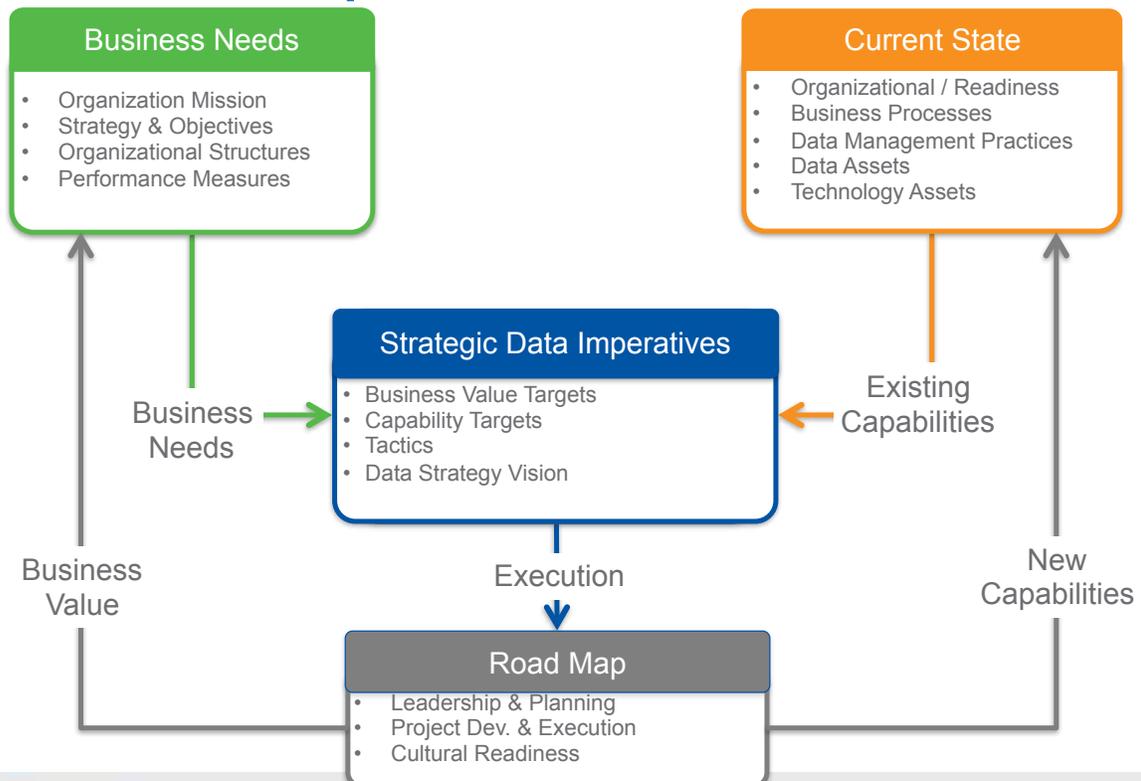
**6** Not Sequencing Data Strategy Implementation



**7** Failing To Address Cultural And Change Management Challenges



## Data Implementation Framework



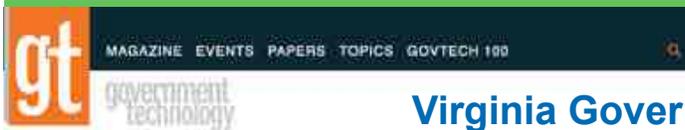
# Data Management Program Expenses

- 5 Data Managers
- \$100,000 Annually
- When will you be done?
- "It's okay my CIO gave me 5 years!"



UNLOCKING BUSINESS VALUE

Copyright 2015 by Data Blueprint Slide # 23



DATA

## Virginia Governor's Data Interns Program



Commonwealth of Virginia  
Office of Governor Terry McAuliffe

### 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 26, 2015



Virginia Commonwealth University

Flickr/Andrew Bain

For Immediate Release  
July 23, 2015

Office of the Governor  
Contact: Brian Coy  
Email: [Brian.Coy@governor.virginia.gov](mailto: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. These 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](mailto:peter.aiken@vcu.edu)) for additional information.



UNLOCKING BUSINESS VALUE

Copyright 2015 by Data Blueprint Slide # 24

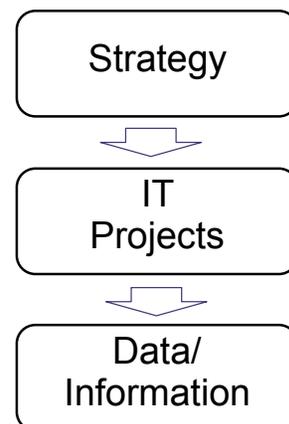
# Exorcising the Seven Deadly Data Sins

- 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



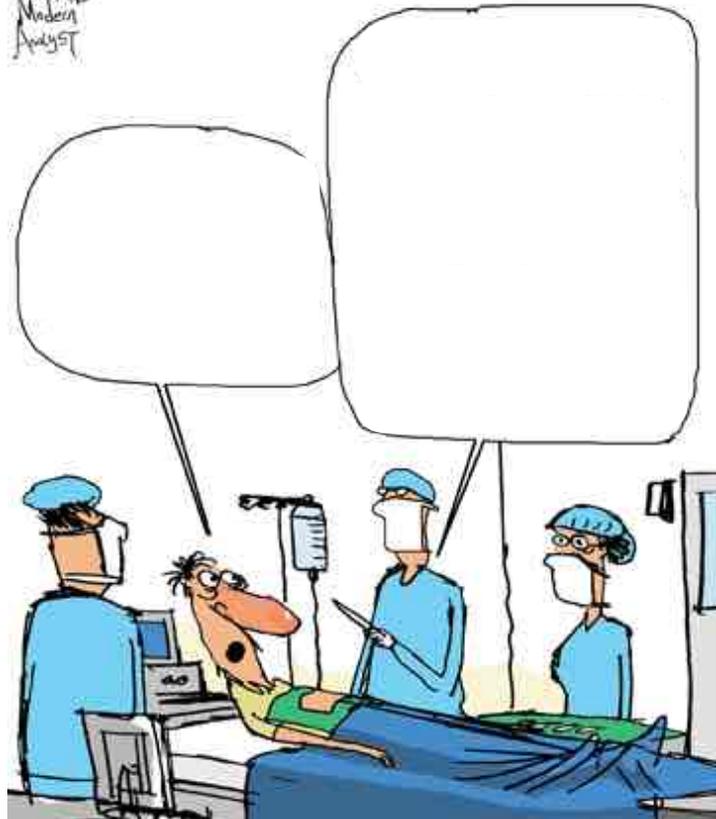
## IT Project or Application-Centric Development

- 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



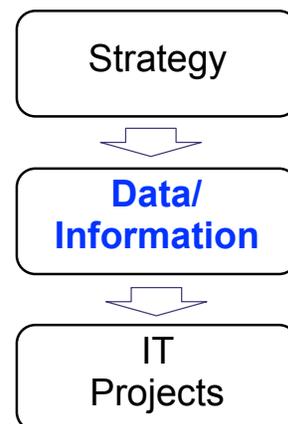
# Agile Surgery

© 1996  
Modern  
Analyst



## 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



# 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

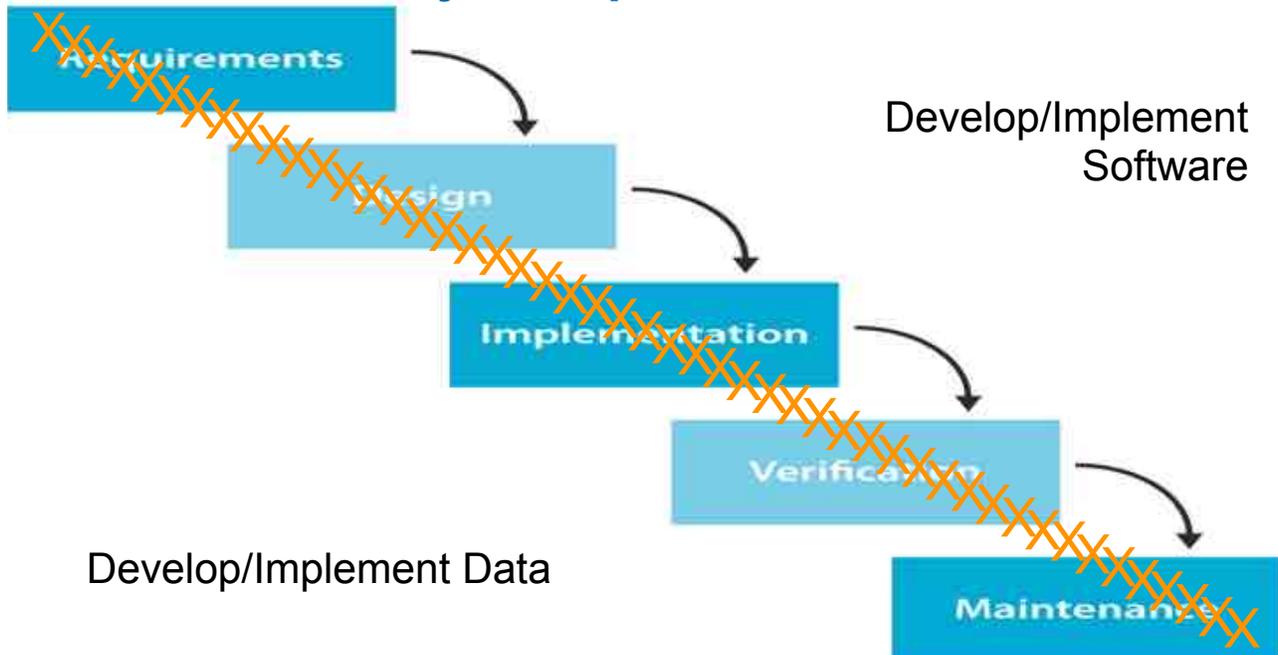
5 Failing to Adequately Manage Expectations

6 Not Sequencing Data Strategy Implementation

7 Failing To Address Cultural And Change Management Challenges

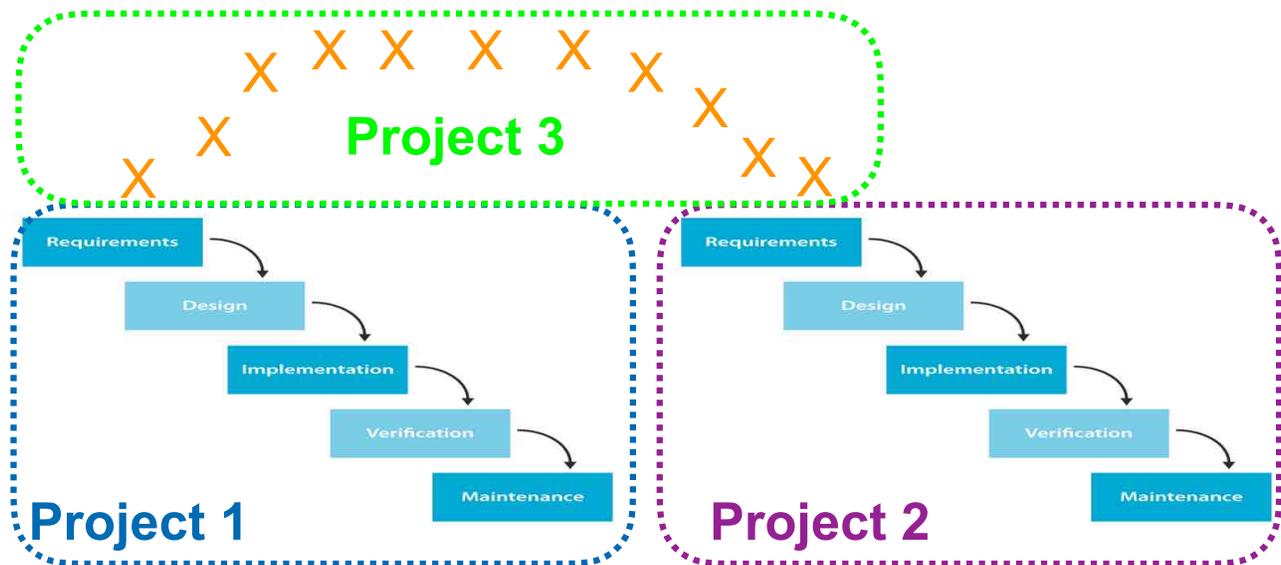


## Project Implementation



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

# Projects Are Silos



*Shared data structures require programmatic development and evaluation*

## 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



## 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?

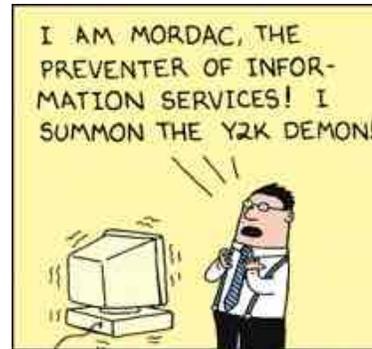
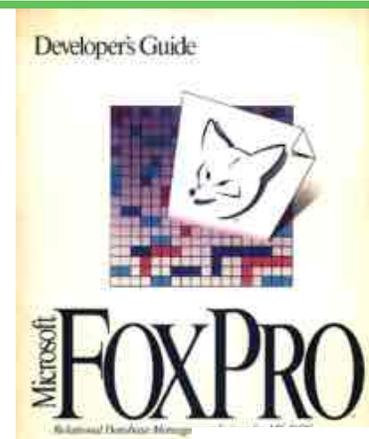


What percentage of the deal with it daily?

100%

# Why should a knowledge worker

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



UNLOCKING BUSINESS VALUE

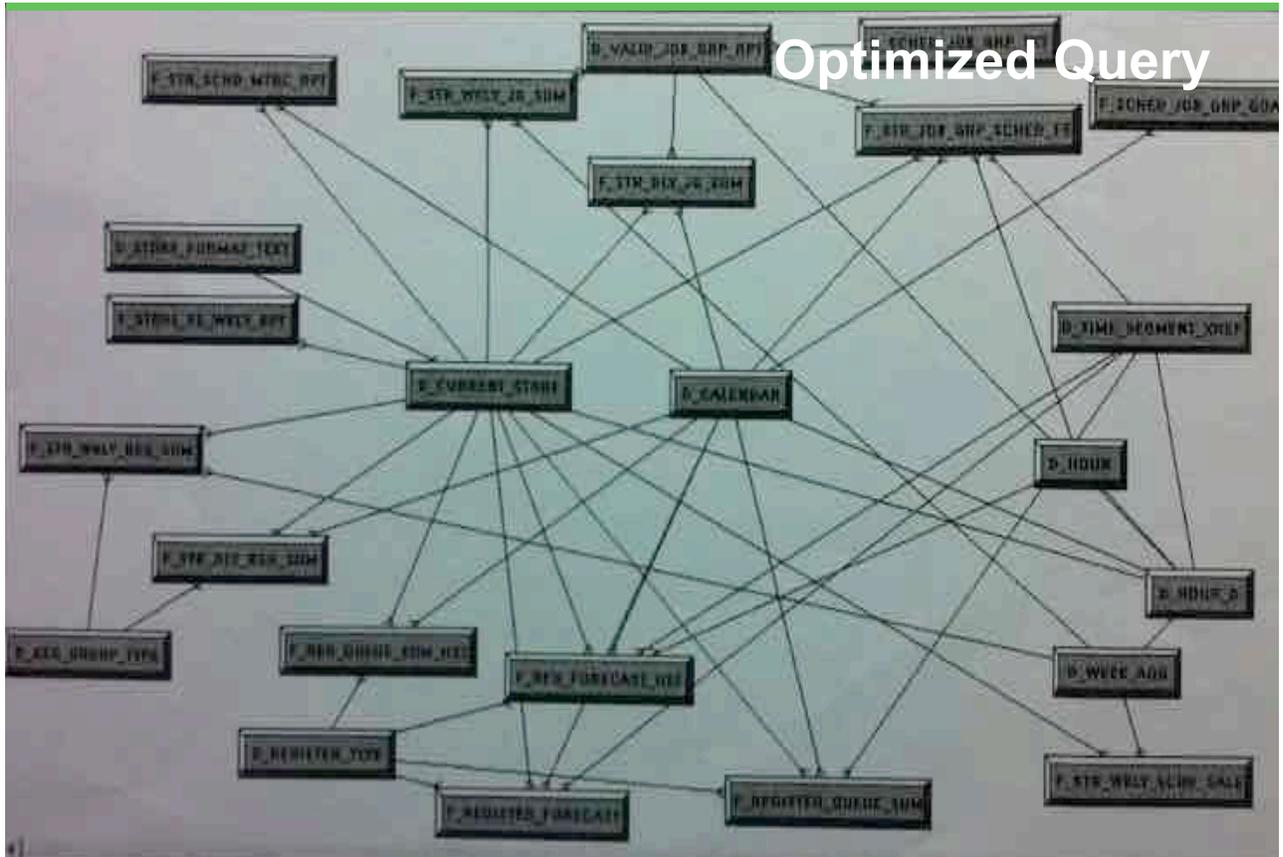
Copyright 2018 by Data Blueprint Side # 35



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Side # 36

# Optimized Query



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

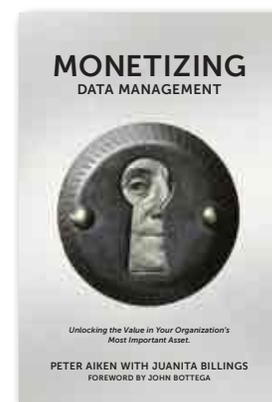




## Data is a hidden IT Expense

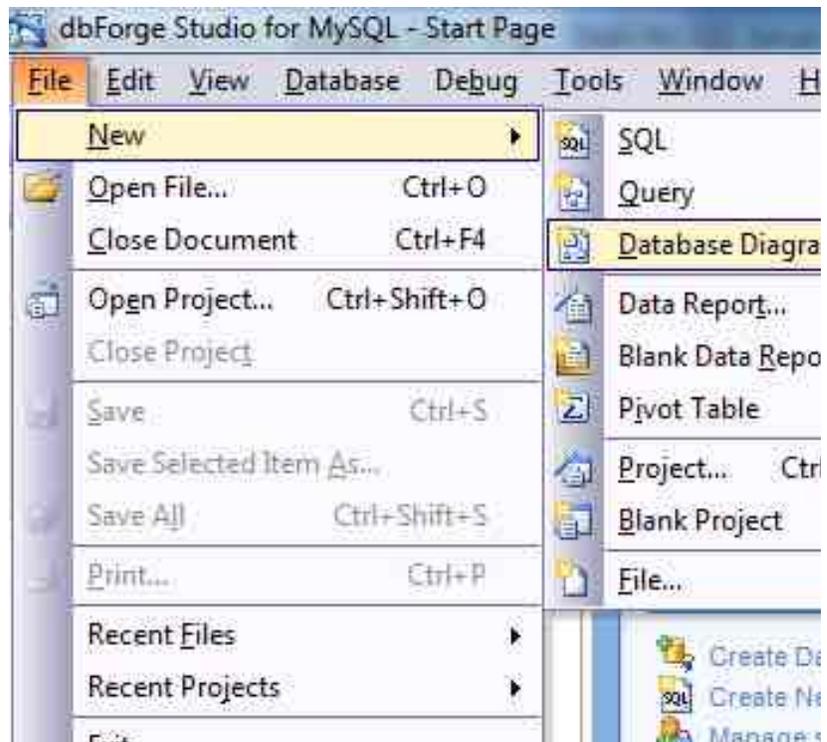
- Organizations spend between 20 - 40% of their IT budget evolving their data - including:
  - Data **migration**
    - Changing the location from one place to another
  - Data **conversion**
    - Changing data into another form, state, or product
  - Data **improving**
    - Inspecting and manipulating,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

— Source: *John Zachman*



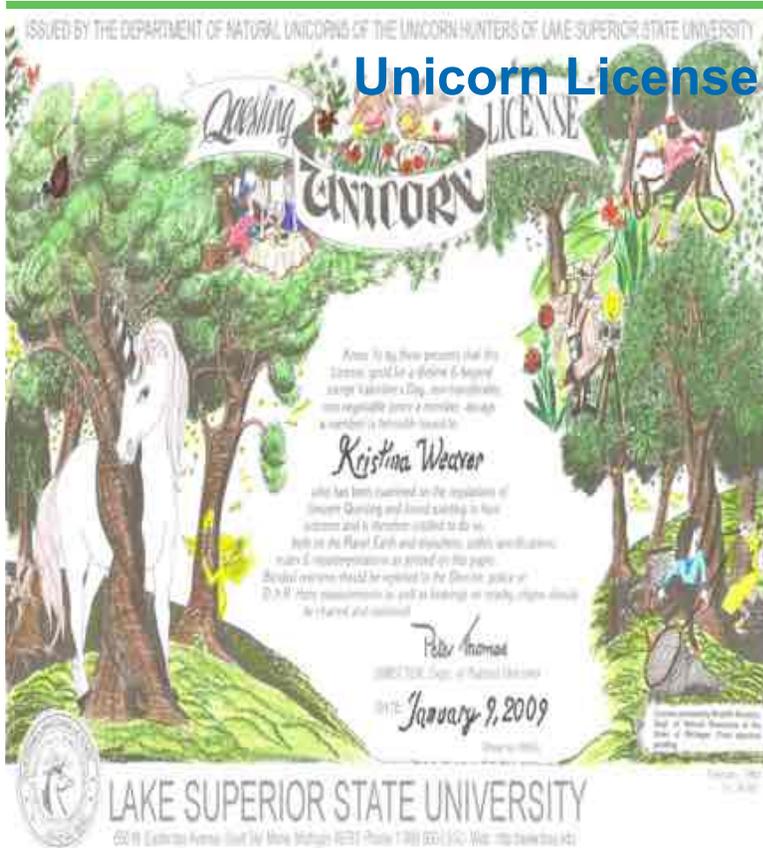
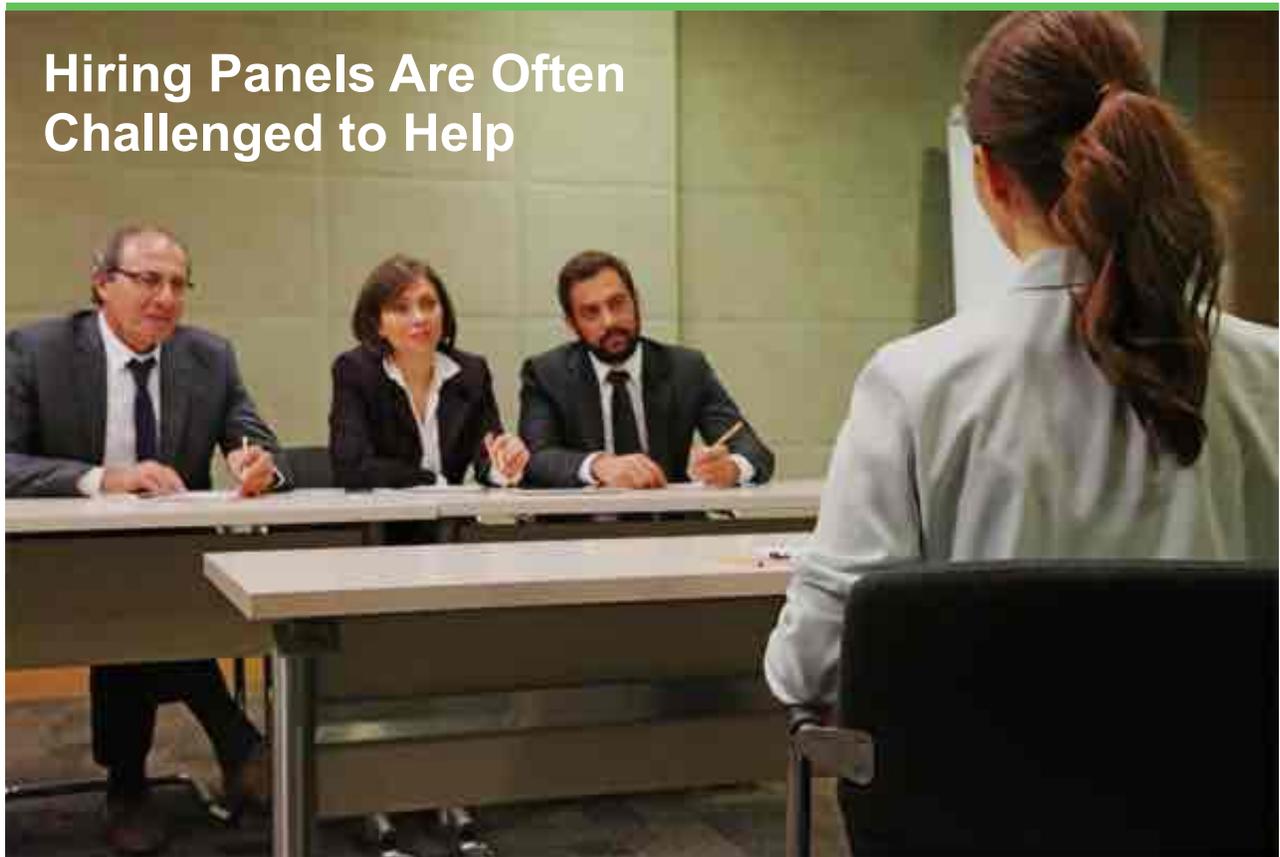
# What do we teach IT professionals about data?

- 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)

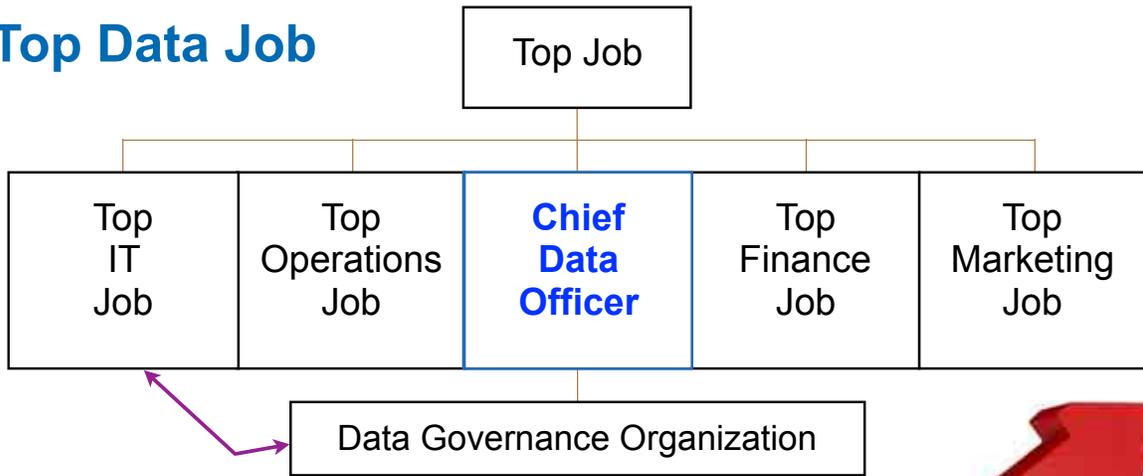
# Hiring Panels Are Often Challenged to Help



## Unicorn License (There Are No Unicorns)



# Top Data Job



- 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?>)



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 45

## The Enterprise Data Executive Takes One for the Team



UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 46

## 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



## George Box British Statistician (1919-2013)

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



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



# Data programmes preceding software development

Common Organizational Data  
(and corresponding data needs requirements)



**Data management and software development must be separated and sequenced**



Systems Development Activities

**Evolve**

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

Future State



(Version +1)

**Create**

New Organizational Capabilities



**data blueprint**

UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 51

## [theDataDoctrine.com](http://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

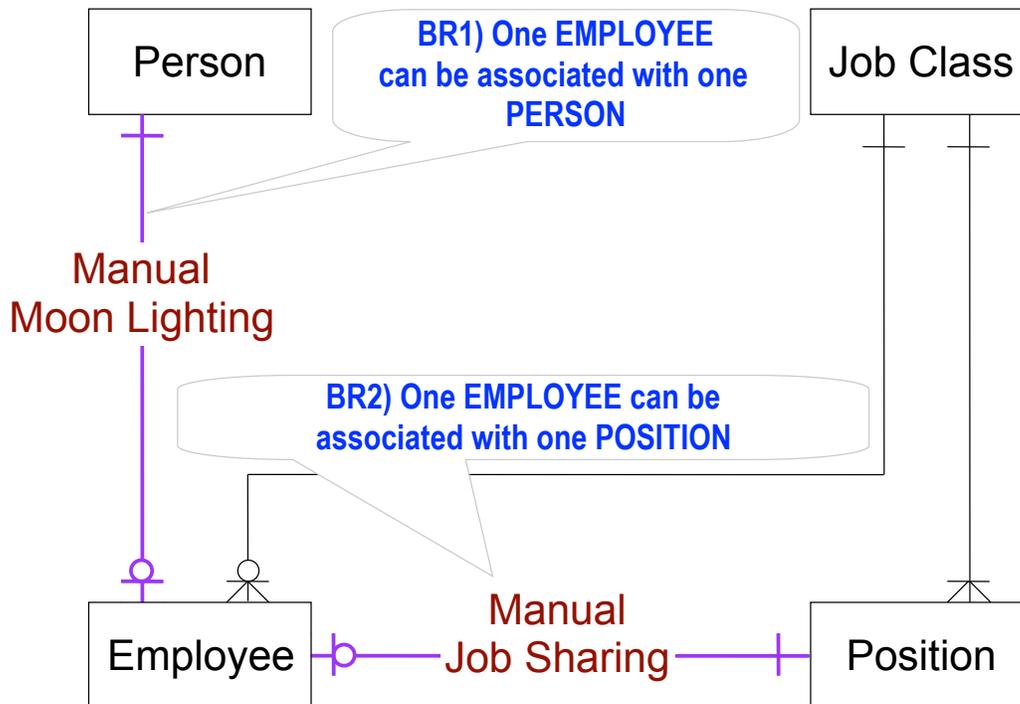
That is, while there is value in the items on the right, we value the items on the left more.

**data blueprint**

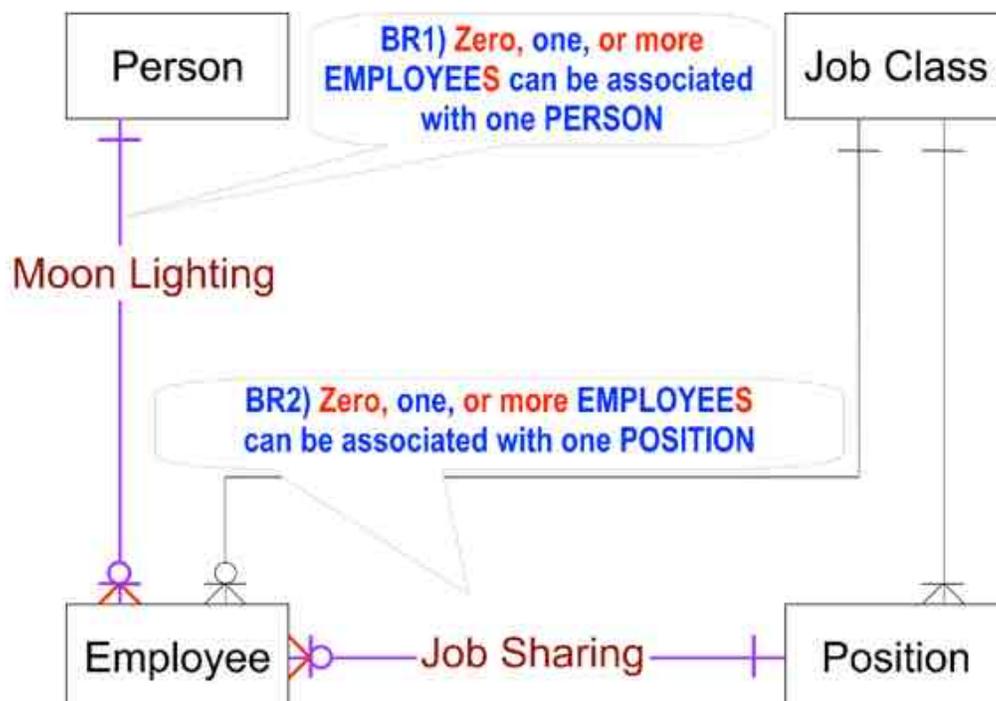
UNLOCKING BUSINESS VALUE

Copyright 2018 by Data Blueprint Slide # 52

## Stable data structures preceding stable code



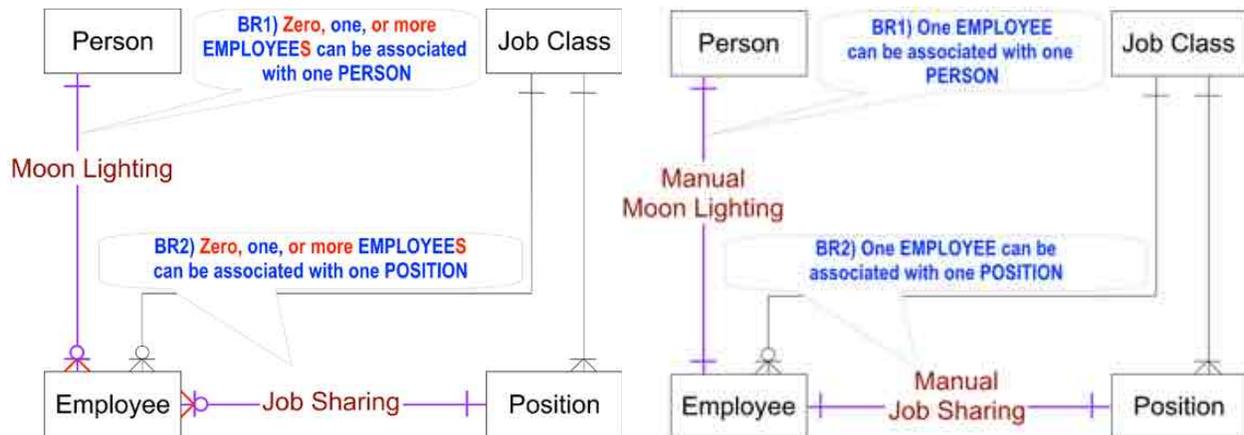
## Stable data structures preceding stable code



# 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](http://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

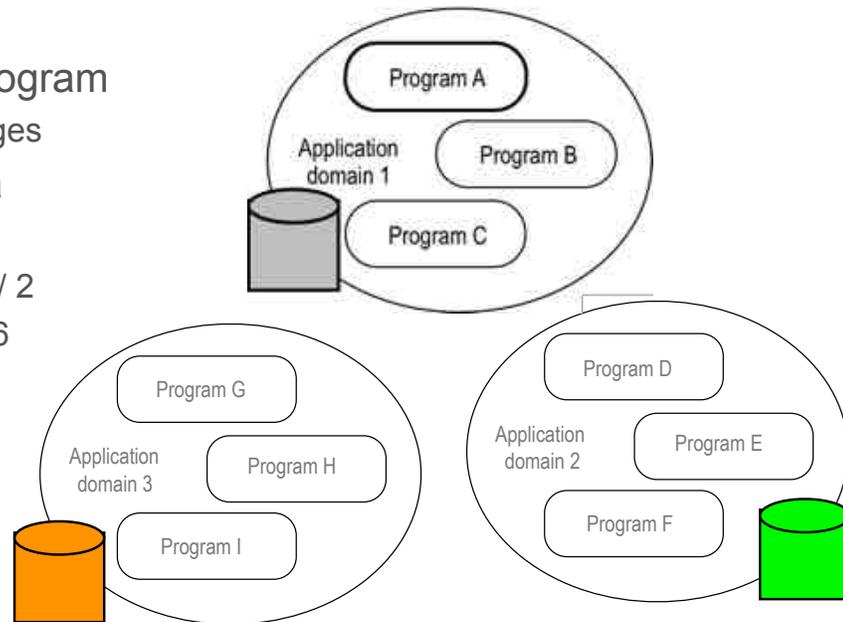
**Data reuse** preceding reusable code

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](http://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, there 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, however, deliver quality software products more rapidly. Better data products pervade and persist in all aspects of systems and increasingly diverse 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 governance and evaluation.
- Reusable data should be leveraged by reusable software.

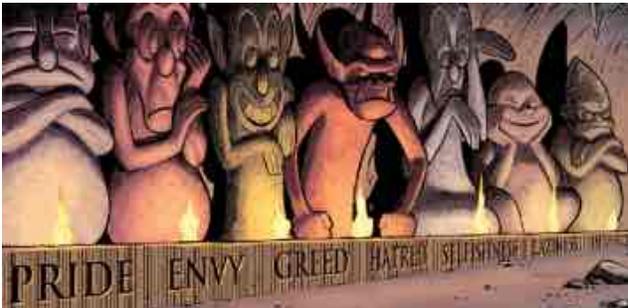
Look around and you can see the consequences of ignoring the foundational role that data plays in our organizations:

- 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.
- Increased 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>

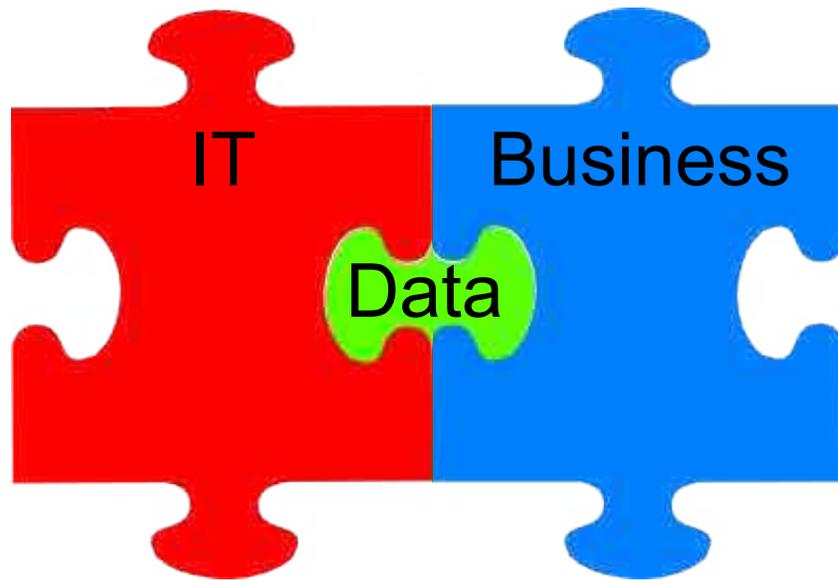
## Exorcising the Seven Deadly Data Sins

-  **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



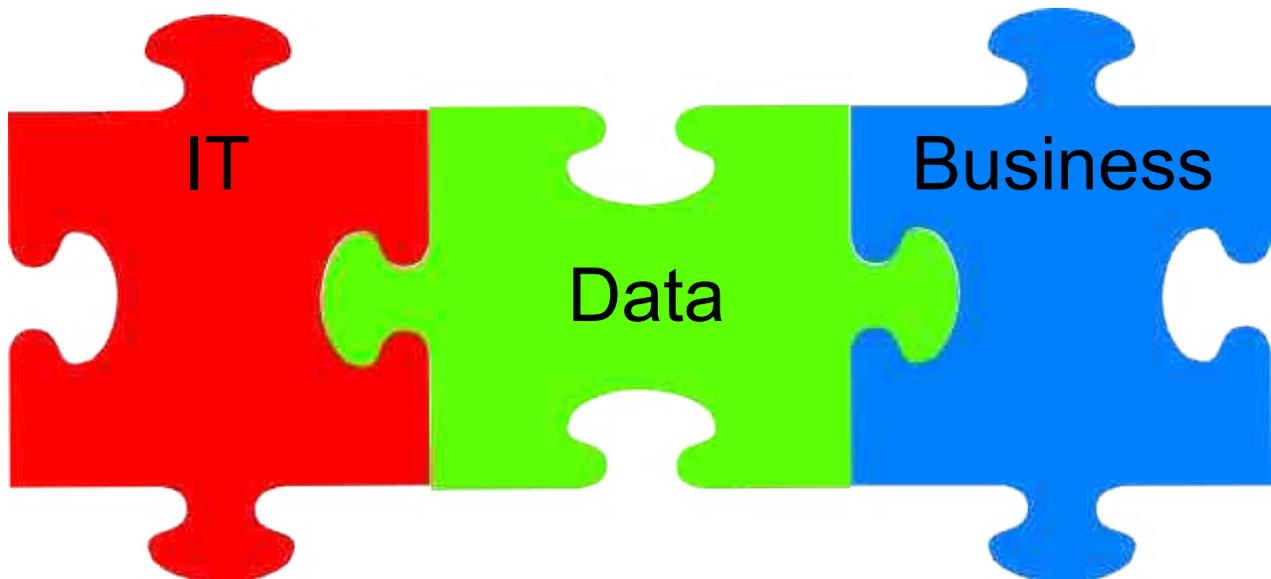
---

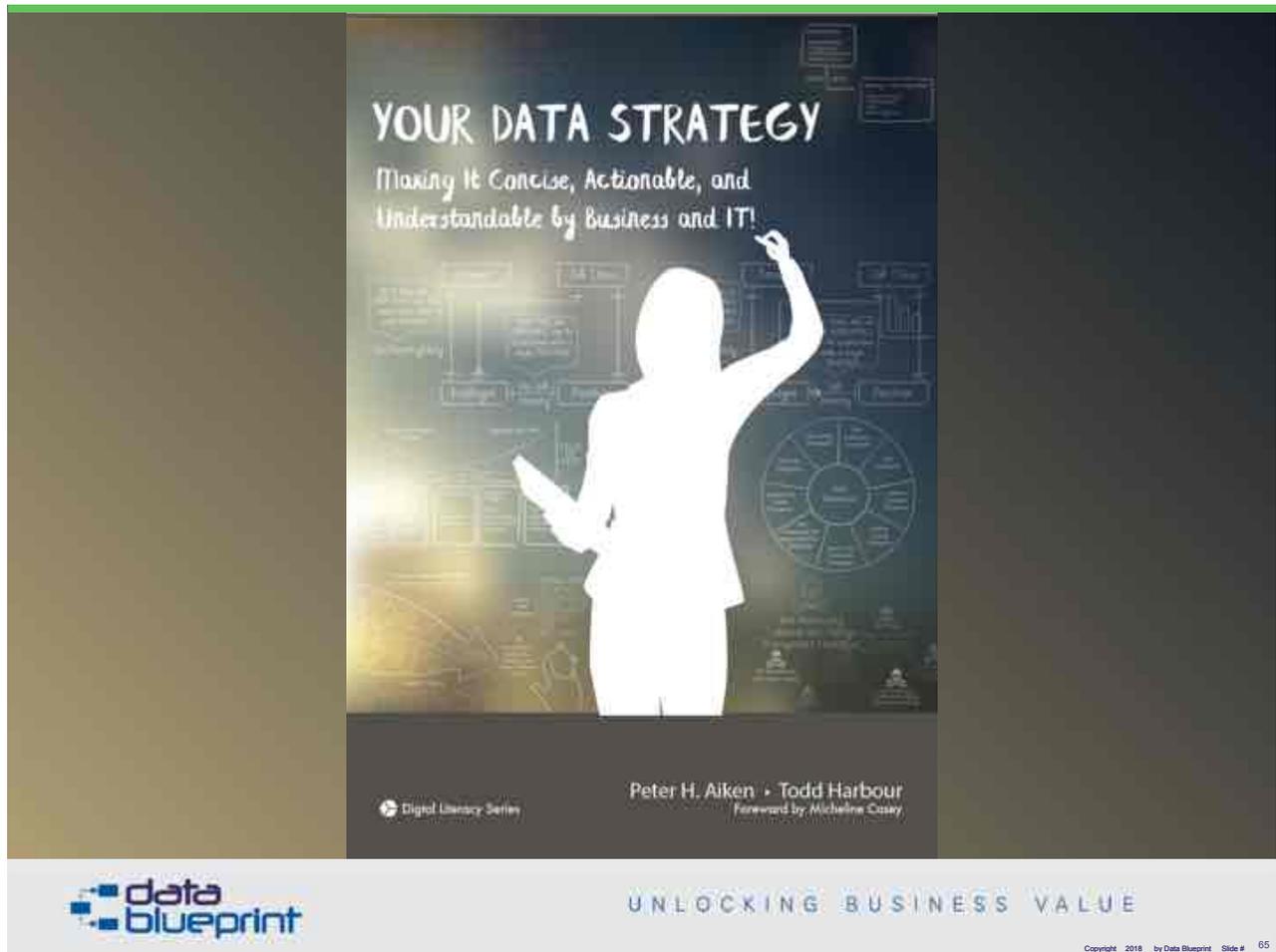
## As Is State of Data (as Perceived)



---

## Desired To Be State of Data (as Understood)





## Questions?



It's your turn!

Use the chat feature or Twitter (#dataed) to submit your questions to Peter now!





UNLOCKING BUSINESS VALUE



[datablueprint.com](http://datablueprint.com)

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