The Case for the Chief Data Officer

Building the Data-driven Enterprise

Presented by Peter Aiken, Ph.D.

Peter Aiken, Ph.D.

- 30+ years of experience in data management
- Multiple international awards & recognition
- Founder, Data Blueprint (datablueprint.com)
- Associate Professor of IS, VCU (vcu.edu)
- (Past) President, DAMA Int. (dama.org)
- 9 books and dozens of articles
- Experienced w/ 500+ data management practices in 20 countries
- Multi-year immersions with organizations as diverse as the US DoD, Nokia, Deutsche Bank, Wells Fargo, Walmart, and the Commonwealth of Virginia
The Case for the Chief Data Officer
Recasting the C-Suite to Leverage your most Valuable Asset

- As-Is/Cause for Concern
  - Disclaimer/Bad Data Decisions Spiral
  - Data Management Practices Hierarchy Structure
  - Lack of Architecture/Engineering Capabilities/Costs
  - Self Assessment/Root cause analysis
- To-Be/Necessary (but insufficient) CDO Prerequisites
  1. Dedicated solely to data asset leveraging
  2. Unconstrained by an IT project mindset
  3. Reporting directly to the business
• Have accomplished astounding technological feats
• Have developed excellent organizational skill sets
• Have delivered phenomenal business value

IT Project Failure Rates (moving average)

Failed | Challenged | Succeeded
---|---|---
1994 | 16% | Failed
1993 | 16% | Challenged
1998 | 15% | Failed
2000 | 10% | Challenged
2002 | 14% | Failed
2004 | 14% | Challenged
2009 | 14% | Failed

Source: Standish Chaos Reports as reported at: http://www.galorath.com/wp/software-project-failure-costs-billions-better-estimation-planning-can-help.php
Bad Data Decisions Spiral

46% of companies report they made an inaccurate business decision based on bad or outdated data.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.

C-level decision makers are not data knowledgable

Most CIOs are not data knowledgable

Bad data decisions

Poor treatment of organizational data assets

Poor quality data

Poor organizational outcomes
The Case for the Chief Data Officer
Recasting the C-Suite to Leverage your most Valuable Asset

- As-Is/Cause for Concern
  - Disclaimer/Bad Data Decisions Spiral
  - Data Management Practices Hierarchy Structure
  - Lack of Architecture/Engineering Capabilities/Costs
  - Self Assessment/Root cause analysis
- To-Be/Necessary (but insufficient) CDO Prerequisites
  1. Dedicated solely to data asset leveraging
  2. Unconstrained by an IT project mindset
  3. Reporting directly to the business
You can accomplish Advanced Data Practices without becoming proficient in the Basic Data Management Practices however this will:

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

Data Management Practices Hierarchy

**Advanced Data Practices**
- MDM
- Mining
- Big Data
- Analytics
- Warehousing
- SOA

**Basic Data Management Practices**
- Data Stewardship
- Data Development
- Organizational Data Integration
- Data Program Management
- Data Support Operations
Data Management is an Integrated System of Five Practice Areas

Five Integrated DM Practices

- **Manage data coherently.**
  - Data Program Coordination

- **Share data across boundaries.**
  - Organizational Data Integration

- **Assign responsibilities for data.**
  - Data Stewardship

- **Engineer data delivery systems.**
  - Data Development

- **Maintain data availability.**
  - Data Support Operations
Weakest Link Results Reporting Results

- Understand five organizational data management practice areas
  - Rate each area per capability maturity model
- Understand the "weakest link" nature of the results reporting
  - Engineered components can only be as strong as their weakest component
  - Low scores seem harsh but are realistic
  - (and on the upside) easily improvable
  - A single “1” degrades the entire practice area – as shown with "stewardship"
- DMPA results are granulated for each practice area providing improvement process guidance

The Case for the Chief Data Officer
Recasting the C-Suite to Leverage your most Valuable Asset

- As-Is/Cause for Concern
  - Disclaimer/Bad Data Decisions Spiral
  - Data Management Practices Hierarchy Structure
  - Lack of Architecture/Engineering Capabilities/Costs
  - Self Assessment/Root cause analysis
- To-Be/Necessary (but insufficient) CDO Prerequisites
  1. Dedicated solely to data asset leveraging
  2. Unconstrained by an IT project mindset
  3. Reporting directly to the business
He who doesn’t lay his foundations before hand, may by great abilities do so afterward ...

... although with great trouble to the architect and danger to the building.

You cannot architect after implementation!
What is this?

• It is tall
• It has a clutch
• It was built in 1942
• It is still in regular use!

Architecture and Engineering

• Architecture enables complex "things" to be built
• Engineering ensures a disciplined approach to development
• Permits organizations to better manage their sole non-depleteable, non-degrading, durable, strategic asset - data
  – within the organization, and
  – with organizational data exchange partners
• Leverage
  – Obtained by implementation of data-centric technologies, processes, and human skill sets
  – Increased by elimination of data ROT (redundant, obsolete, or trivial)
    • The bigger the organization, the greater potential leverage exists
• Treating data more asset-like simultaneously
  1. lowers organizational IT costs and
  2. increases organizational knowledge worker productivity

---

**A Model Specifying Relationships Among Important Terms**

1. Each FACT combines with one or more MEANINGS.
2. Each specific FACT and MEANING combination is referred to as a DATUM.
3. An INFORMATION is one or more DATA that are returned in response to a specific REQUEST.
4. INFORMATION REUSE is enabled when one FACT is combined with more than one MEANING.
5. INTELLIGENCE is INFORMATION associated with its USES.

[Built on definition by Dan Appleton 1983]
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

The Case for the Chief Data Officer

- As-Is/Cause for Concern
  - Disclaimer/Bad Data Decisions Spiral
  - Data Management Practices Hierarchy Structure
  - Lack of Architecture/Engineering Capabilities/Costs
  - **Self Assessment/Root cause analysis**
- To-Be/Necessary (but insufficient) CDO Prerequisites
  1. Dedicated solely to data asset leveraging
  2. Unconstrained by an IT project mindset
  3. Reporting directly to the business
Organizations Surveyed

- Results from more than 500 organizations
- 32% government
- Appropriate public company representation
- Enough data to demonstrate European organization DM practices are generally more mature

Data Management Capability Maturity Model Levels

We have a process for improving our DM capabilities

- We manage our DM processes so that the whole organization can follow our standard DM guidance
- We have experience that we have standardized so that all in the organization can follow it
- We have DM experience and have the ability to implement disciplined processes
- Our DM practices are ad hoc and dependent upon "heroes" and heroic efforts

Optimizing
(5)

- Managed
(4)
- Defined
(3)
- Repeatable
(2)
- Initial
(1)

One concept for process improvement, others include:
- Norton Stage Theory
- TQM
- TQdM
- TDQM
- ISO 9000

and focus on understanding current processes and determining where to make improvements.
### Data Management Practice Areas

<table>
<thead>
<tr>
<th>Practice Area</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data program coordination</td>
<td>DM is practiced as a coherent and coordinated set of activities</td>
</tr>
<tr>
<td>Organizational data integration</td>
<td>Delivery of data is support of organizational objectives – the currency of DM</td>
</tr>
<tr>
<td>Data stewardship</td>
<td>Designating specific individuals caretakers for certain data</td>
</tr>
<tr>
<td>Data development</td>
<td>Efficient delivery of data via appropriate channels</td>
</tr>
<tr>
<td>Data support</td>
<td>Ensuring reliable access to data</td>
</tr>
</tbody>
</table>

### Capability Maturity Model Levels

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – Initial</td>
<td>Our DM practices are ad hoc and dependent upon &quot;heroes&quot; and heroic efforts</td>
</tr>
<tr>
<td>2 - Repeatable</td>
<td>We have DM experience and have the ability to implement disciplined processes</td>
</tr>
<tr>
<td>3 - Documented</td>
<td>We have standardized DM practices so that all in the organization can perform it with uniform quality</td>
</tr>
<tr>
<td>4 - Managed</td>
<td>We manage our DM processes so that the whole organization can follow our standard DM guidance</td>
</tr>
<tr>
<td>5 - Optimizing</td>
<td>We have a process for improving our DM capabilities</td>
</tr>
</tbody>
</table>

### Data Management Practices Measurement (DMPA)

<table>
<thead>
<tr>
<th>Practice Area</th>
<th>Initial (I)</th>
<th>Repeatable (II)</th>
<th>Documented (III)</th>
<th>Managed (IV)</th>
<th>Optimizing (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Program Coordination</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organizational Data Integration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Stewardship</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Support Operations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- CMU's Software Engineering Institute (SEI) Collaboration
- Results from hundreds organizations in various industries including:
  - Public Companies
  - State Government Agencies
  - Federal Government
  - International Organizations
- Defined industry standard
- Steps toward defining data management "state of the practice"
Data Management Practices Assessment

- Data Program Coordination
- Organizational Data Integration
- Data Stewardship
- Data Development
- Data Support Operations

High Marks for IFC's Audit

- Leadership & Guidance
- Asset Creation
- Metadata Management
- Quality Assurance
- Change Management
- Data Quality
Largely Ineffective Investments

- Approximately, 10% percent of organizations achieve parity and (potential positive returns) on their DM investments
- Only 30% of DM investments achieve tangible returns at all
- Seventy percent of organizations have very small or no tangible return on their DM investments

Comparison of DM Maturity 2007-2012

- Data Program Coordination
- Organizational Data Integration
- Data Stewardship
- Data Development
- Data Support Operations

2007 Maturity Levels
2012 Maturity Levels
Conclusion must be?

1. CIOs are unaware of the strategic nature of data; or

2. CIOs are not concerned about how data management is accomplished in their organizations; or

3. CIOs think data management is being adequately accomplished in their organizations.

% of DM organizations labeled "successful"

• In 25 years:
  – "Successful" DM organizations fell from 43% to 15%
  – "Unsuccessful" increased from 5% to 21%.
CIOs are distancing themselves from DM

<table>
<thead>
<tr>
<th>Percentage Reporting Directly to the CIO</th>
<th>1981</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>74%</td>
<td></td>
<td>43%</td>
</tr>
</tbody>
</table>

| Percentage reporting 3 or more levels below the CIO | 26% | 57% |

Copyright 2013 by Data Blueprint

Why ETL and Data Migration Projects Fail

- Assessed 1200 migration projects!
  - Surveyed only experienced migration specialists who have done at least four migration projects
- The median project cost is over 10 times the amount planned!
  - Biggest Challenges: Bad Data; Missing Data; Duplicate Data
- The survey did not consider projects that were cancelled largely due to data migration difficulties
- "… problems are encountered rather than discovered"
Root Cause Analysis of IT Challenges

Poor Data Asset Management

5% Sales Increase Versus Data Volume

- Sales
- Data Volume

Driver: Information Overload!

- Increasing demand
- Increasing amount
- Heightened public and corporate sensitivity to security, privacy, and compliance
- Data dependent IT initiatives (BI, SOA, Analytics, Big Data ...)
- New data emphasis
- Leveraging data for competitive advantage and profitability

A likely state of your data

Take control of your data growth.

Redundancy

Multiple Data Sources

Inconsistent Data Quality

IT are data owners

Very Silo’d or conflicting data sources

Multiple changes to source system

Inconsistent data definitions of common terms

Lots of Data... Minimum Information

Difficult to report and mine against
A likely state of your data management efforts

Various Maturity Frameworks

Data is the new oil!
Data is the new (s)oil!
Data is the new bacon!

Data is a lubricant!
Data becomes a fuel!

Adapted from John Ladley
**Enron**

- August 2001 Enron stock falls to $42/share from $90/share
- Dynergy brings several $ billion in an attempted rescue
- Enron spends entire amount in 1 week
  - Any person can write a check at Enron for
  - Any amount of money for
  - Any purchase at
  - Any time
- Enron goes back to Dynergy for more $
- Dynergy: *What happened to the several $ billion I gave you last week?*
- Enron:

---

**CFO Necessary Prerequisites/Qualifications**

- CPA
- CMA
- Masters of Accountancy
- Other recognized degrees/certifications
- These are necessary but insufficient prerequisites/qualifications
What do we teach knowledge workers about data?

What percentage of the deal with it daily?

100%

What do we teach IT professionals about data?

- 1 course
  - How to build a new database
  - 80% if IT expenses are used to improve existing IT assets
- What impressions do IT professionals get from this education?
  - Data is a technical skill that is used to develop new databases
- This is not the best way to educate IT and business professionals - every organization's
  - Sole, non-depletable, non-degrading, durable, strategic asset
Organizations aren't ...

- The ultimate authority on organizational informational assets
- Able to devote the required time/attention to the management of organizational informational assets
- Possessed of the requisite expertise to manage organizational informational assets
- Situated to achieve success organizationally as long as they have a technology development perspective

---

The Case for the Chief Data Officer
Recasting the C-Suite to Leverage your most Valuable Asset

- As-Is/Cause for Concern
  - Disclaimer/Bad Data Decisions Spiral
  - Data Management Practices Hierarchy Structure
  - Lack of Architecture/Engineering Capabilities/Costs
  - Self Assessment/Root cause analysis
- To-Be/Necessary (but insufficient) CDO Prerequisites
  1. Dedicated solely to data asset leveraging
  2. Unconstrained by an IT project mindset
  3. Reporting directly to the business
CIO Infrastructure Focus

Top Five CIO Concerns 2005-2011

<table>
<thead>
<tr>
<th>Survey/Monitoring</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grant Thornton-State CIO Survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Ameritech ITAA Annual CIO survey</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CIO Magazine-State of the CIO</td>
<td></td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>UK CIO Survey</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gartner Annual Priorities</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Informationweek Global CIO top 10 issues</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accenture CIO Survey</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KPMG &amp; Harvey Nash</td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NASCIO Survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>Robert Half Technology</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robert Half Technology</td>
<td>3</td>
<td>6</td>
<td>9</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
</tbody>
</table>
The Top Job

- Finance
- Operations
- Sale/Marketing
- HR
- Risk
- Information
- Technology/CIO
  - Align IT initiatives with business goals
  - Improving IT operations performance
  - Cultivating the IT/business partnership
  - Cost control/expense management
  - Implementing new systems
  - Leading change efforts
  - Driving business innovation
  - Redesigning business processes
  - Developing and refining business strategy
  - Negotiating with IT vendors
  - Managing IT crises
  - Developing market strategies & technologies
  - Security management
  - Studying trends to identify opportunities

Where does data go?
Chief Electrification Officer – responsible for electrical generating and distribution systems. The title was used mainly in developed countries from the 1880s to 1940s during the electrification of industry, but is still used in some developing countries.

It is difficult to remain data-focused

- Question: What is the hardest part of doing analysis?
- Answer: Not doing design!
- Question: What is the hardest part of a CDO's job?
- Answer: Remaining data focused!
- Anything the CDO does that is not ...
  - ... applying organizational data assets to the implementation of organizational strategy ...
  - ... is a similar distraction
The Case for the Chief Data Officer
Recasting the C-Suite to Leverage your most Valuable Asset

- As-Is/Cause for Concern
  - Disclaimer/Bad Data Decisions Spiral
  - Data Management Practices Hierarchy Structure
  - Lack of Architecture/Engineering Capabilities/Costs
  - Self Assessment/Root cause analysis

- To-Be/Necessary (but insufficient) CDO Prerequisites
  1. Dedicated solely to data asset leveraging
  2. Unconstrained by an IT project mindset
  3. Reporting directly to the business

Application-Centric Development

- In support of strategy, organizations develop specific goals/objectives
- The goals/objectives drive the development of specific systems/applications
- Development of systems/applications leads to network/infrastructure requirements
- Data/information are typically considered after the systems/applications and network/infrastructure have been articulated
- 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
Data-Centric Development

• In support of strategy, the organization develops specific goals/objectives
• The goals/objectives drive the development of specific data/information assets with an eye to organization-wide usage
• Network/infrastructure components are developed to support organization-wide use of data
• Development of systems/applications is derived from the data/network architecture
• 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

Data-Centric Perspective

• Measure success differently
• Same project
• Same process
• Different measures for success
  – Asking if the data is correct
  – Valuing data more than valuing "on time and within budget";
  – Valuing correct data more than correct processes
  – Auditing data rather than project documents
• Articulation by Linda Bevolo
This represents a gradual shift from application to data-centric

"Waterfall" model creates new data siloes
Evolving Data is Different than Creating New Systems

Common Organizational Data (and corresponding data needs requirements)

Future State

Future State (Version +1)

Evolve

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

Create

New Organizational Capabilities

Data is not a Project

• Durable asset
  – An asset that has a usable life more than one year

• Reasonable project deliverables
  – 90 day increments
  – Data evolution is measured in years

• Data
  – Evolves - it is not created
  – Significantly more stable

• Readymade data architectural components
  – Prerequisite to agile development

• Only alternative is to create additional data siloes!
The Case for the Chief Data Officer
Recasting the C-Suite to Leverage your most Valuable Asset

• As-Is/Cause for Concern
  - Disclaimer/Bad Data Decisions Spiral
  - Data Management Practices Hierarchy Structure
  - Lack of Architecture/Engineering Capabilities/Costs
  - Self Assessment/Root cause analysis

• To-Be/Necessary (but insufficient) CDO Prerequisites
  1. Dedicated solely to data asset leveraging
  2. Unconstrained by an IT project mindset
  3. Reporting directly to the business

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?"
There is enough work to justify the function
There is not much talent
The CDO provides significant input to the Top Information Technology Job

CDO Reporting Particulars

1. Report outside of IT and the current CIO altogether;

2. Report to the same organizational structure that the CFO and other "top" jobs report into; and

3. Focus on activities that are outside of (and more importantly) upstream from any system development lifecycle activities (SDLC).
**New division of labor**

- **Reporting to IT**
  - Data Development
  - Database Operations Management
- **Shared with the business**
  - Metadata Management
  - Data Security Management
- **Reporting to Business**
  - Data Architecture Management
  - Reference & Master Data Management
  - Data Warehousing & BI Management
  - Document & Content Management
  - Data Quality Management
  - Data Governance

---

**Ishikawa Fishbone Diagrams**

- **Why is infant mortality so high?**
  - Malnourished mothers
- **Why are mothers malnourished?**
  - Substandard biology educations in high school
- **Why do biology programs substandard?**
  - Poor education of high school biology teachers
- **Why do we have poor biology teacher education?**
  - Biology profession unaware of consequences

- **Why are so many organizational technology experiences so poor?**
  - Misunderstanding of data's role in IT
- **Why do so few understand data's role in IT?**
  - Little, if any, focus on enterprise-wide data use in the educational system
- **Why is the educational system not addressing this gap?**
  - Lack of recognition by the system
- **Why has the system not yet been made aware of this deficiency?**
  - Lack of understanding at the C-level of these issues
- **Why do they not understand?**
  - Little, if any, focus on enterprise-wide data use in the educational system

---

**Asking "why" repeatedly!**
The Case for the Chief Data Officer
Recasting the C-Suite to Leverage your most Valuable Asset

- As-Is/Cause for Concern
  - Disclaimer/Bad Data Decisions Spiral
  - Data Management Practices Hierarchy Structure
  - Lack of Architecture/Engineering Capabilities/Costs
  - Self Assessment/Root cause analysis

- To-Be/Necessary (but insufficient) CDO Prerequisites
  1. Dedicated solely to data asset leveraging
  2. Unconstrained by an IT project mindset
  3. Reporting directly to the business