



Peter Aiken, Ph.D.

- I've been doing this a long time
- My work is recognized as useful
- Associate Professor of IS (vcu.edu)
- Institute for Defense Analyses (ida.org)
- DAMA International (dama.org)
- MIT CDO Society (iscdo.org)
- Anything Awesome (anythingawesome.com)
- Experienced w/ 500+ data management practices worldwide
- 12 books and dozens of articles
- Multi-year immersions
 - US DoD (DISA/Army/Marines/DLA)
 - Nokia
 - Deutsche Bank
 - Wells Fargo
 - Walmart
 - HUD ...

\$1,500,000,000.00 USD



Program overview

- Data's Confounding Characteristics

- Uneven understanding
- Has lead fractured views of data and to
- Increasing organizational data debt

1. Keeping DG practically focused on strategy

- This is a young profession and must
- Directly support organizational strategy by
- Improving data and its use in the short and long term

2. DG must exist at the same level as HR

- In order to achieve effectiveness,
- DG is central to DM (and central to digitization efforts)
- Must be de-coupled from IT strategy

3. Gradually add ingredients (practicing and getting better)

- Digital and data are dependent on high speed automation/data processing
- Employ a DG Frameworks to refine focus
- Plan to evolve (PDCA)

4. Data governance in action: Storytelling

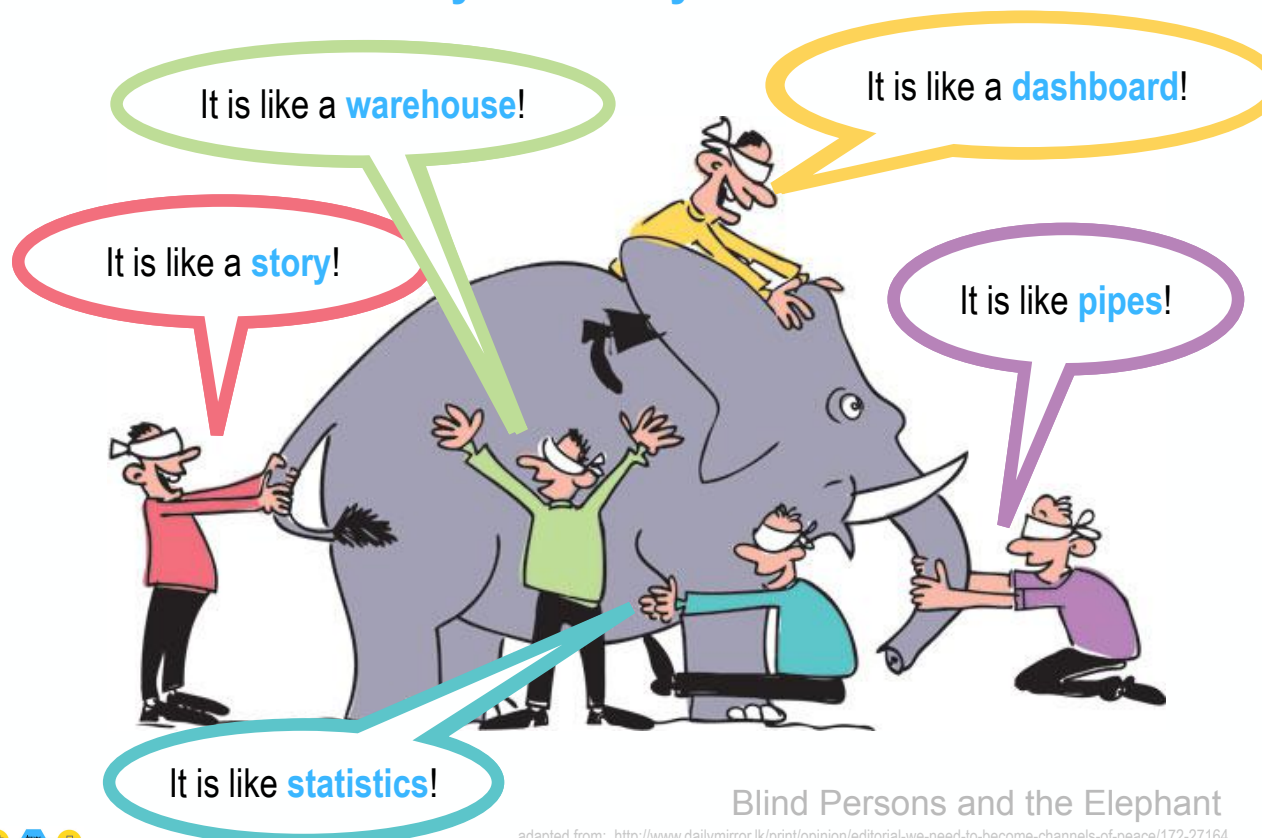
- Take Aways/References/Q&A



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 3

Data is not broadly or widely understood



Blind Persons and the Elephant

adapted from: <http://www.dailymirror.lk/print/opinion/editorial-we-need-to-become-channels-of-peace/172-27164>

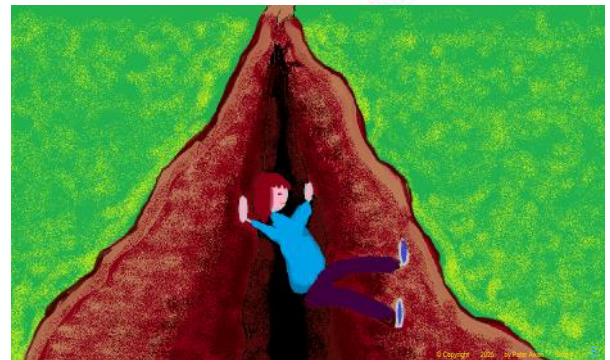


<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 4

Confusion as to data responsibility

- 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?"



The Princess on the Pea

by
Hans Christian
Andersen



Sleepless
(forever?)

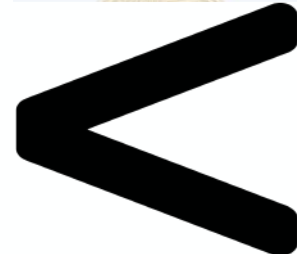


Doing a poor job with data governance

- Failure to understand the role of data governance re: proposed and existing software/services
 - Locks in imperfections for the life of the application
 - Restricts data investment benefits
 - Decreases organizational data leverage
- Accounts for 20-40% of IT budgets devoted to evolving
 - Data **migration** (Changing the data location)
 - Data **conversion** (Changing data form, state, or product)
 - Data **improving** (Inspecting and manipulating, or re-keying data to prepare it for subsequent use)
- Lack of data governance causes everything else to
 - Take longer
 - Cost more
 - Deliver less
 - Present greater risk (with thanks to Tom DeMarco)



<https://creativecommons.org/licenses/by-nc-sa/4.0/>



gemini.google.com



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Provide examples of how poor data governance causes everything to cost more, deliver less, take longer, and present greater risk



Let's illustrate how poor data governance can lead to increased costs, reduced value, extended timelines, and greater risks with concrete examples:

1. Increased Costs:

- **Duplicate Data Entry:** Without standardized data entry processes and a single source of truth, employees may enter the same information multiple times across different systems. This wastes time and resources, increases the chance of errors, and inflates storage costs. Example: A non-profit having to mail the same fundraising appeal to the same donor multiple times because their CRM has duplicate records.
 - **Data Rework:** Poor data quality (inaccurate, incomplete, inconsistent data) necessitates costly data cleansing and correction efforts. Employees spend valuable time fixing data instead of using it for analysis or decision-making. Example: A marketing team launching a campaign with incorrect customer addresses, resulting in wasted postage and lost potential revenue.
 - **System Integration Issues:** When systems are built on poorly governed data with differing standards, integrating them becomes complex and expensive. Data mapping and transformation require significant resources and can introduce further errors. Example: A hospital merging two departments and finding that patient records are incompatible due to differing data formats, leading to costly and time-consuming data migration efforts.
 - **Missed Opportunities for Discounts/Savings:** Inaccurate or incomplete data can prevent organizations from qualifying for volume discounts, taking advantage of early payment incentives, or identifying cost-saving opportunities. Example: A manufacturing company missing out on a bulk discount on raw materials because their inventory management system has inaccurate stock levels.
- ## 2. Reduced Value/Deliver Less:
- **Inaccurate Reporting and Analytics:** Decisions based on flawed data lead to poor outcomes. Inaccurate sales forecasts, flawed risk assessments, and misguided marketing campaigns can all result from poor data governance. Example: A retailer stocking up on the wrong products due to inaccurate sales data, leading to lost revenue and increased storage costs.
 - **Ineffective Marketing:** Targeting the wrong audience with marketing messages wastes resources and damages brand reputation. Poor data quality can lead to inaccurate customer segmentation and ineffective campaigns. Example: A software company sending email promotions for a product upgrade to customers who already have the latest version.
 - **Lost Revenue Opportunities:** Poor data governance can lead to missed opportunities for upselling, cross-selling, and customer retention. Example: A bank failing to offer a suitable loan product to a customer because their financial data is incomplete or inaccurate.

- **Decreased Customer Satisfaction:** Inaccurate or inconsistent data can lead to poor customer service experiences, such as incorrect billing, delayed responses, and personalized offers that miss the mark. Example: A customer service representative unable to access a customer's complete order history, leading to frustration and a negative customer experience.
- ## 3. Take Longer:
- **Data Silos and Redundancy:** When data is scattered across different systems and departments, it takes longer to find the information needed for decision-making or operational tasks. Example: A research team spending weeks gathering data from various sources for a project, when a centralized data repository could have provided the information much faster.
 - **Lengthy Data Validation Processes:** Poor data quality necessitates time-consuming data validation and verification processes. Example: An accounting team spending days reconciling financial data from different systems due to inconsistencies and errors.
 - **Delayed Project Completion:** Data-dependent projects can be significantly delayed due to data quality issues, data access challenges, or system integration problems. Example: A new product launch being delayed because the product information database is incomplete and inaccurate.
- ## 4. Greater Risk:
- **Compliance Violations:** Poor data governance can lead to non-compliance with data privacy regulations (GDPR, CCPA, HIPAA, etc.), resulting in hefty fines and legal repercussions. Example: A healthcare provider accidentally disclosing patient medical information due to inadequate data security measures.
 - **Security Breaches:** Lack of proper data security controls increases the risk of data breaches and cyberattacks, which can lead to financial losses, reputational damage, and legal liabilities. Example: A financial institution experiencing a data breach that exposes customer account information, leading to financial losses and reputational harm.
 - **Reputational Damage:** Data quality issues and security breaches can damage an organization's reputation and erode customer trust. Example: A social media company facing public backlash due to a data privacy scandal.
 - **Strategic Missteps:** Decisions based on flawed data can lead to costly strategic mistakes, such as entering the wrong markets, investing in the wrong products, or pursuing ineffective business strategies. Example: A retail company expanding into a new geographic area based on inaccurate market research data, leading to financial losses and business closure.

By addressing data governance proactively, organizations can avoid these costly consequences and unlock the true potential of their data assets.



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 6 9

Provide examples of how poor data governance causes everything at Walmart to cost more, deliver less, take longer, and present greater risk



It's great you're thinking about the impact of poor data governance at a company like Walmart! Given its scale, the effects of poor data governance can be massive. Here are some examples of how it can cause increased costs, reduced delivery, longer timelines, and greater risk:

INCREASED COSTS:

Inventory Management Inefficiencies:

- Poor data governance can lead to inaccurate inventory data. This means Walmart might over-order products that aren't selling well, leading to storage costs, waste from expired goods, and markdowns to clear inventory.
- Conversely, inaccurate data can lead to under-ordering popular items, resulting in lost sales and dissatisfied customers.

Supply Chain Disruptions:

- Walmart's supply chain is incredibly complex. If data on supplier performance, shipping times, or demand forecasting is inaccurate or inconsistent, it can lead to delays, higher transportation costs (e.g., needing to expedite shipments), and increased labor costs to resolve issues.

Marketing Campaign Ineffectiveness:

- Poor customer data quality (e.g., duplicate records, incorrect contact information) can result in wasted marketing spend.
- For example, sending the same promotional email multiple times to the same customer or sending offers for products they don't want reduces the return on investment (ROI) of marketing campaigns.

Financial Reporting Errors:

- Inaccurate sales data, expense data, or other financial information can lead to errors in financial reports.
- This can result in increased costs for audits, corrections, and potential penalties.

REDUCED DELIVERY:

E-commerce Fulfillment Issues:

- Inaccurate inventory data directly impacts e-commerce fulfillment. If the system shows an item is in stock when it's not, it leads to order cancellations, delayed shipments, and customer dissatisfaction.
- Incorrect address data can also cause delivery failures and increased shipping costs for redelivery.

In-Store Availability Problems:

- Poor data governance can disrupt the flow of goods to stores. If the data on product demand or store inventory is inaccurate, shelves may not be stocked with the right products at the right time, leading to lost sales and frustrated customers.

LONGER TIMELINES:

Product Development Delays:

- Walmart develops and sources a vast number of products. If data related to product specifications, supplier information, or quality control is poorly managed, it can lead to delays in the product development lifecycle.

Decision-Making Bottlenecks:

- When data is inconsistent or unreliable, it takes longer for managers to make informed decisions.
- They may need to spend extra time verifying data, resolving discrepancies, and seeking approvals, which slows down business processes.

IT Project Delays:

- Poor data governance can significantly impact IT projects. For example, if data migration is required for a system upgrade, data quality issues can cause significant delays, cost overruns, and project failures.

GREATER RISK:

Compliance Violations:

- Walmart handles a lot of sensitive data (customer data, financial data, etc.). Poor data governance increases the risk of data breaches, privacy violations, and non-compliance with regulations like GDPR or CCPA. This can lead to hefty fines and reputational damage.

Operational Risks:

- Inaccurate data can lead to poor operational decisions, such as stocking unsafe products or failing to identify and address supply chain vulnerabilities.

Reputational Damage:

- Data breaches, inaccurate product information, or supply chain issues caused by poor data governance can severely damage Walmart's reputation and erode customer trust.

Financial Risks:

- All of the above issues (increased costs, reduced delivery, longer timelines) ultimately translate into financial risks for Walmart, impacting profitability, shareholder value, and long-term sustainability.

In conclusion, effective data governance is absolutely critical for Walmart to operate efficiently, effectively, and responsibly. Poor data governance creates a ripple effect throughout the organization, leading to a wide range of negative consequences.



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 6 10

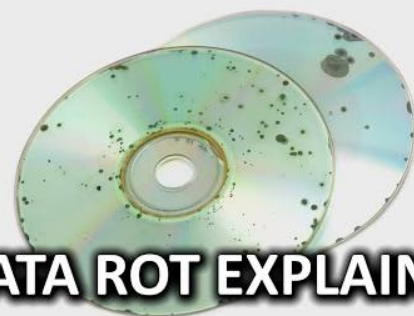


<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Organizing the Wheat Separated from the Chaff



Is well organized data worth more?



DATA ROT EXPLAINED



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Pre-Information Age Metadata



- Examples of information architecture achievements that happened well before the information age:

- Page numbering
- Alphabetical order
- Table of contents
- Indexes
- Lexicons
- Maps
- Diagrams



"While we can arrange things with the intent to communicate certain information, we can't actually make information. Our users do that for us."

Example from: *How to make sense of any mess* by Abby Covert (2014) ISBN: 1500615994



DATA ROT EXPLAINED



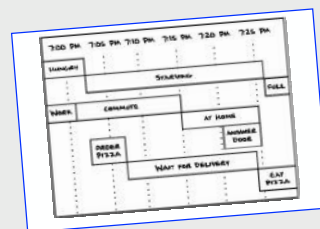
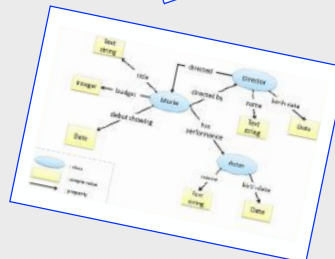
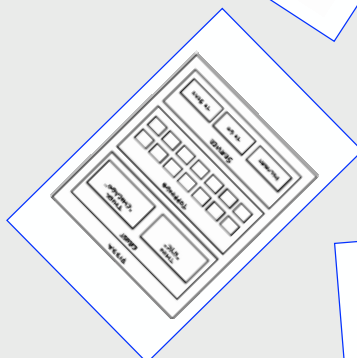
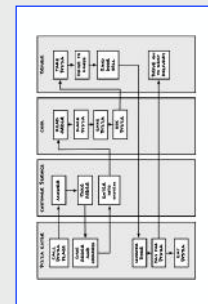
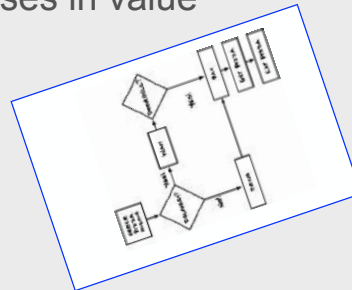
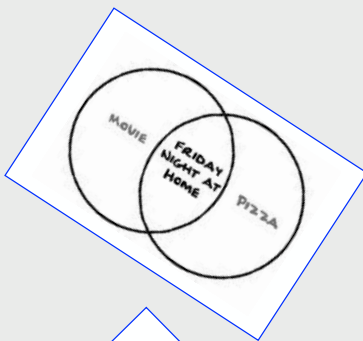
<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 13

Remove the structure and things fall apart rapidly



- Better organized data increases in value



DATA ROT EXPLAINED



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 14

Organizing the Wheat Separated from the Chaff



- Better organized data increases in value
- Poor data management practices are costing organizations money/time/effort
- 80% of organizational data is **ROT**
 - **R**edundant
 - **O**bsolute
 - **T**rivial
- The question is which data to eliminate?
 - Most enterprise data is never analyzed



Who is best qualified to accomplish this?

How do i go about identifying data rot?



Data rot, also known as data decay, is a significant concern for any organization that relies on data for its operations and decision-making. It refers to the degradation or obsolescence of data over time, making it inaccurate, unreliable, or unusable. Identifying data rot is crucial for maintaining data quality and ensuring that data remains a valuable asset. Here's how you can go about it:

1. Establish Data Quality Metrics and Monitoring

Define Key Metrics:

- Start by defining key data quality metrics relevant to your business needs. These might include:
 - Accuracy: Does the data correctly reflect reality?
 - Completeness: Are all required data fields populated?
 - Consistency: Is the data consistent across different systems and sources?
 - Timeliness: Is the data up-to-date and relevant for current use?
 - Validity: Does the data conform to defined formats and rules?

Implement Monitoring:

- Set up systems and processes to regularly monitor these metrics.
- This could involve:
 - Automated Data Quality Checks: Implement rules and checks within your databases or data processing pipelines to automatically identify data quality issues.
 - Data Profiling: Use data profiling tools to analyze the characteristics of your data and identify anomalies or patterns that might indicate data rot.
 - Regular Audits: Conduct periodic audits of your data to assess its quality and identify any areas of concern.

2. Look for Specific Signs and Symptoms

Inconsistencies and Discrepancies:

- Look for data that contradicts itself or is inconsistent across different systems or reports.
- For example, the same customer might have different addresses or contact information in different databases.

Missing or Incomplete Data:

- Identify records with missing or incomplete data fields.
- This could indicate that data is not being captured correctly or that it has been lost or corrupted over time.

Outdated Information:

- Check for data that is no longer current or relevant.
- For example, customer contact information, product pricing, or supplier details might become outdated.

Data Entry Errors:

- Look for typos, incorrect formatting, or other data entry errors.
- These errors can accumulate over time and make data unreliable.

Data Drift:

- Data drift refers to changes in the statistical properties of data over time.
- This can occur when data is not updated to reflect changes in the real world or when data collection methods change.
- Data drift can lead to inaccurate analysis and predictions.

Broken Integrations:

- If data integrations between systems are not working correctly, it can lead to data inconsistencies and loss.
- Monitor data flows between systems to ensure that data is being transferred accurately and reliably.

3. Implement Data Governance Practices

Data Governance Policies:

- Establish clear data governance policies and procedures to define data ownership, data quality standards, and data management responsibilities.

Data Stewardship:

- Assign data stewards to be responsible for the quality and integrity of specific data domains.
- Data stewards can play a key role in identifying and addressing data rot.

Data Lifecycle Management:

- Implement data lifecycle management practices to ensure that data is properly managed throughout its lifecycle, from creation to deletion or archiving.
- This includes processes for data validation, data cleansing, and data maintenance.

4. Leverage Technology and Tools

Data Quality Tools:

- Use data quality tools to automate data quality checks, data profiling, and data cleansing.
- These tools can help you identify and correct data rot more efficiently.

Data Monitoring Tools:

- Implement data monitoring tools to track data quality metrics and alert you to potential issues.

Data Catalog:

- A data catalog can help you understand the lineage and usage of your data, making it easier to identify and address data rot.

5. Foster a Data-Driven Culture

Data Awareness:

- Promote a data-driven culture within your organization, where employees understand the importance of data quality and are empowered to report data issues.

Training and Education:

- Provide training and education to employees on data quality principles and best practices.
- By implementing these strategies, you can proactively identify and address data rot, ensuring that your data remains a reliable and valuable asset for your organization.



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2025 by Peter Allen Slide 6 17

How do i go about identifying data rot at Walmart?



Identifying data rot at a massive organization like Walmart requires a multi-faceted approach. Here's how you can go about it, combining proactive monitoring, reactive detection, and a focus on key business areas:

I. Proactive Monitoring and Assessment

Data Quality Audits:

- Perform regular audits of critical data domains (customer, product, sales, etc.) to assess accuracy, completeness, consistency, and timeliness.
- Use data profiling tools to analyze data patterns, identify anomalies, and detect potential data quality issues.
- Establish key performance indicators (KPIs) for data quality and track them over time.

Data Validation Rules:

- Implement automated data validation rules at various points in the data lifecycle (e.g., data entry, data processing, data storage).
- These rules should check for adherence to data standards, business rules, and referential integrity.

Examples:

- Ensuring that product IDs follow a specific format.
- Verifying that customer addresses are complete and valid.
- Checking that sales transactions are associated with existing products and customers.

Metadata Management:

- Maintain a comprehensive and up-to-date metadata repository that documents the definition, source, and usage of data elements.
- This helps to understand the context of data and identify potential inconsistencies or redundancies.
- Actively monitor metadata for changes and ensure it aligns with the actual data.

Data Lineage Tracking:

- Implement data lineage tools to track the flow of data from its source to its destination.
- This helps to identify where data transformations or errors might occur and pinpoint the source of data rot.

System Monitoring:

- Monitor the performance and health of systems that store and process data.
- Identify any system errors, bottlenecks, or failures that could lead to data corruption or loss.

II. Reactive Detection and Analysis

User Reports and Feedback:

- Establish channels for business users to report data quality issues or concerns.
- Encourage users to provide specific examples of data errors or inconsistencies they encounter.
- Track and analyze user feedback to identify recurring data rot patterns.

Business Process Monitoring:

- Monitor key business processes that rely on data, such as:
 - Order fulfillment
 - Inventory management
 - Customer service
 - Financial reporting

- Identify any process inefficiencies, errors, or anomalies that could be attributed to data rot.

Anomaly Detection in Reporting:

- Analyze reports and dashboards for unexpected trends, outliers, or inconsistencies.
- Investigate any anomalies to determine if they are caused by data errors.
- Use statistical analysis or machine learning techniques to automate anomaly detection.

Reconciliation Processes:

- Regularly reconcile data between different systems or data sources to identify discrepancies.

For example:

- Reconciling inventory data between the warehouse management system and the sales system.
- Reconciling financial data between different accounting systems.

III. Focus on Key Business Areas at Walmart

Product Data:

- Focus on identifying inaccuracies in product descriptions, pricing, availability, and specifications.

- This is critical for e-commerce accuracy, supply chain optimization, and customer satisfaction.

Customer Data:

- Monitor for duplicate customer records, incorrect contact information, and outdated preferences.
- This impacts marketing effectiveness, customer service, and loyalty programs.

Sales and Transaction Data:

- Analyze sales data for errors in pricing, discounts, and tax calculations.

- Monitor transaction data for fraudulent activity or inconsistencies.

Supply Chain Data:

- Focus on data related to supplier performance, shipping information, and inventory levels.
- This is crucial for efficient logistics, timely deliveries, and cost control.

Store Data:

- Monitor data related to store inventory, staffing, and customer traffic.
- This helps optimize store operations and improve the customer experience.

IV. Technology and Tools

Data Quality Tools:

- Use specialized data quality tools for data profiling, data cleansing, data matching, and data monitoring.

Master Data Management (MDM) Systems:

- Implement MDM systems to centralize and manage critical master data domains, ensuring accuracy and consistency.

Data Governance Platforms:

- Use data governance platforms to define data policies, establish data ownership, and track data quality metrics.

Business Intelligence (BI) and Analytics Tools:

- Leverage BI and analytics tools to monitor data trends, detect anomalies, and identify potential data rot issues.

By combining these proactive, reactive, and focused approaches, Walmart can effectively identify and address data rot, ensuring the integrity and reliability of its data assets.



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2025 by Peter Allen Slide 6 18

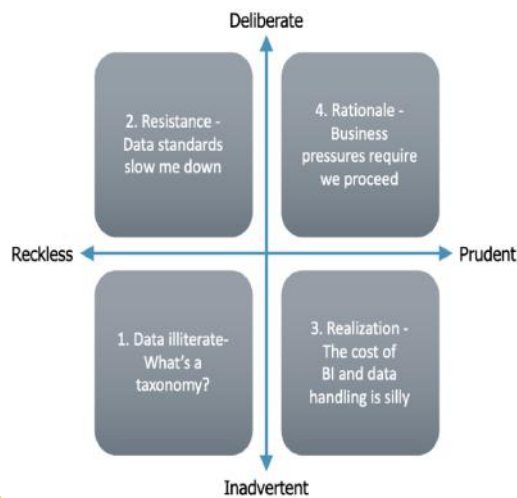
You must address data debt proactively



Data debt:

- Slows progress
- Decreases quality
- Increases costs
- Presents greater risks

- Data debt
 - The time and effort it will take to return your shared data to a governed state from its (likely) current state of ungoverned
- Getting back to zero
 - Involves undoing existing stuff
 - Likely new skills are required



<https://uk.nttdataservices.com/en/blog/2020/february/how-to-get-rid-of-your-data-debt>



<https://johnladley.com/a-bit-more-on-data-debt/>

<https://www.merkleinc.com/blog/are-you-buried-alive-data-debt>

© Copyright 2020 by Peter Allen Slide 6 19



It is not as easy to visualize the cost of Data Debt or that it depletes organizational resources:

- Slowing progress
- Decreasing quality
- Increasing costs
- Presenting greater risks



<https://johnladley.com/a-bit-more-on-data-debt/>

Where should I look in my organization for data debt?



Data debt, like financial debt, accumulates over time and can significantly hinder an organization's performance. It's often hidden, making it crucial to know where to look. Here's a breakdown of areas to investigate within your organization:

1. Data Silos and Fragmented Systems:

- **Departmental Databases:** Each department may have its own database, leading to inconsistent data definitions, redundant data storage, and difficulties in sharing information across the organization.
- **Legacy Systems:** Older systems, often poorly documented and difficult to maintain, can harbor significant data quality issues and inconsistencies.
- **Spreadsheets as Databases:** Over-reliance on spreadsheets for critical data management can create data quality problems, version control issues, and security vulnerabilities.
- **Shadow IT:** Unofficial systems and data stores created by employees outside of IT's control can introduce significant data quality and security risks.

2. Data Quality Issues:

- **Incomplete Data:** Missing values in critical data fields can limit the usefulness of the data and lead to inaccurate analysis.
- **Inaccurate Data:** Incorrect or outdated information can result in flawed decision-making and operational inefficiencies.
- **Inconsistent Data:** Different data formats, definitions, or units of measurement across systems can make it difficult to integrate and analyze data.
- **Duplicate Data:** Redundant data storage wastes resources and increases the risk of inconsistencies.
- **Outdated Data:** Data that is no longer current can lead to incorrect insights and missed opportunities.

3. Data Governance Gaps:

- **Lack of Data Ownership:** When no one is responsible for data quality and management, it can lead to neglect and data degradation.
- **Missing Data Standards:** Without clear data standards and definitions, it's difficult to ensure consistency and interoperability across systems.

- **Weak Data Security:** Inadequate security measures can increase the risk of data breaches and compliance violations.

- **Insufficient Metadata:** Missing or incomplete metadata (data about data) can make it difficult to understand the meaning and context of data, limiting its usefulness.

4. Business Processes and Workflows:

- **Manual Data Entry:** Manual data entry is prone to errors and can be a significant source of data debt.
- **Lack of Data Validation:** Without data validation checks, errors can easily propagate through systems and processes.
- **Inefficient Data Flows:** Complex or poorly designed data flows can create bottlenecks and increase the risk of data errors.

5. Technology and Infrastructure:

- **Outdated Technology:** Older technologies may not be able to handle the volume and complexity of data, leading to performance issues and data quality problems.
- **Lack of Integration:** Poorly integrated systems can make it difficult to share data and create a unified view of information.
- **Limited Scalability:** Systems that cannot scale to meet growing data volumes can become a bottleneck and hinder business growth.

Specific Questions to Ask:

- "Where do we have the most manual data entry?" This is often a hotbed of data debt.
- "Which reports do we trust the least?" This points to potential data quality issues.
- "How long does it take to get a single, accurate view of [key business metric]?" A long time suggests data silos and integration challenges.
- "Where are our biggest data-related compliance risks?" This highlights areas where data governance is weak.
- "What systems are we afraid to touch?" Legacy systems often hide significant data debt.

By systematically investigating these areas, you can uncover the hidden data debt within your organization and begin to develop a plan to address it. Remember that addressing data debt is an ongoing process, not a one-time fix. It requires a commitment to data governance, data quality, and continuous improvement.



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 21

Welcome to VCU's Policy Program

This page houses both the policy library which contains the most current version of VCU policies and resources for developing and revising policies at VCU. This resource is provided as a service to members of the VCU community and the general public. If you encounter any issues accessing pages on this site please contact the Policy Program at policy@vcu.edu or by calling 804-828-2336.

Please note: some policies may reference other VCU policies. Policies referenced are accessible by clicking the green buttons at the bottom of the screen when any given policy is open.

Please use the search box below to search policies by keyword.



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Search for documents



Found 162 record(s)

VCU Policy

Share
Settings

Sources

+ Add source

Select all sources

Academic Engagement With...
Academic Rights & Responsi...
Accessibility and Reasonabl...
Additional Sources of Acade...
Alcohol and Other Drugs.pdf
Animals on University Prope...
Annual Assessment of Facul...
Assessing Student Learning ...
Assessing Student Learning ...
Awarding Honorary Degrees...
Campus Expression and Spa...
Changing the Designation o...
Commencement Participatio...
Commencement Participatio...

Chat

VCU Policy

50 sources

These documents detail Virginia Commonwealth University's policies and procedures, covering diverse aspects of university operations. They address student academic engagement and conduct, faculty roles and responsibilities, accessibility and accommodations for disabilities, credit transfer and honorary awards, campus expression and space usage, computer and network resource use, corporate research agreements, policy creation and maintenance, criminal conviction investigations, data and information governance, debt management, deceased student refunds, reporting misconduct and retaliation protection, consensual employee-student relationships, financial aid adjustments, fraud identification and reporting, hazing prevention and discipline, the university's honor system, international travel, the learning management system, graduate handbook guidelines, military-affiliated student policies, nondiscrimination and affirmative action, parental notification for dependent students' mental health treatment, PhD program guidelines, and textbook adoption and sales, as well as visa sponsorship. Many policies emphasize compliance with federal and state laws. The documents provide definitions, specifics, procedures, and frequently asked questions to clarify expectations and ensure adherence to regulations.

Save to note

Add note

Audio Overview

Briefing doc

Studio

Audio Overview

Click to load the conversation.
Load

Notes

+ Add note

Study guide
Briefing doc
FAQ
Timeline

VCU Policy Summary
Okay, here's a detailed briefing document summarizing the main themes and important...

VCU Policy Timeline and Stakeholders
Okay, here is a detailed timeline and cast of characters based on the provided sources:...

VCU Student FAQs
FAQs How can students earn academic credit for prior learning or experiences outside of...

VCU Policy Summary
Okay, here's a detailed briefing document summarizing the main themes and important...

New note
Academic Policies and Resources at VCU: FAQ 1. What are the different ways I can earn...

New note
VCU Policy Handbook Study Guide Quiz Instructions: Answer the following short-...

New note
Virginia Commonwealth University Policy Guide Table of Contents Additional Sources of...

New note
Virginia Commonwealth University (VCU) Policy Briefing Document This document provides a...

Keep the proper focus

- Wrong question:
 - How should we govern all this data?
- Right question:
 - Should we include this data item within the scope of our current data governance practices?
- Regardless of the decision, document why!

VCU Policy

© Copyright 2020 by Peter Allen Slide 4 24

Program Overview

- Data's Confounding Characteristics

- Uneven understanding
- Has lead fractured views of data and to
- Increasing organizational data debt



1. Keeping DG practically focused on strategy

- This is a young profession and must
- Directly support organizational strategy by
- Improving data and its use in the short and long term

2. DG must exist at the same level as HR

- In order to achieve effectiveness,
- DG is central to DM (and central to digitization efforts)
- Must be de-coupled from IT strategy

3. Gradually add ingredients (practicing and getting better)

- Digital and data are dependent on high speed automation/data processing
- Employ a DG Frameworks to refine focus
- Plan to evolve (PDCA)



4. Data governance in action: Storytelling

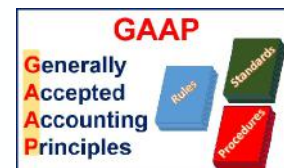
- Take Aways/References/Q&A



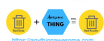
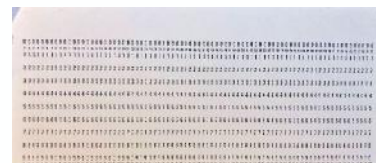
How old is your profession?



- 8,000+ years
- formalize practices
- GAAP



**Augusta Ada King
Countess of Lovelace
(1815-52)**



Current approaches are not and have not been working

What We Learned From Top Execs About Their Big Data And AI Initiatives

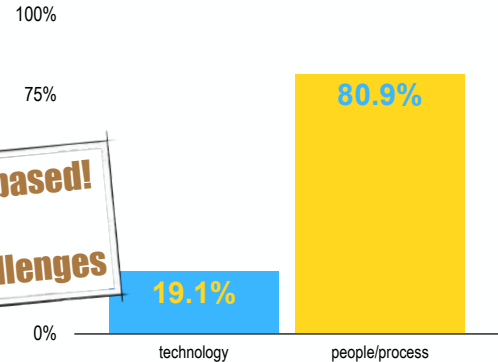


Randy Bean Contributor
CIO Network Contributor Group @
Enterprise & Cloud

Culture still eats strategy for breakfast

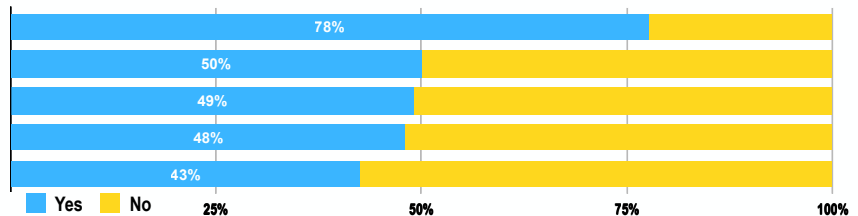
This aphorism is attributed to legendary management consultant Peter Drucker, and it certainly appears to hold true for data transformation efforts. Leading companies continue to identify culture – people, process, organization, change management – as the biggest impediment to becoming data-driven organizations – 92.2%. Few companies – only 30.0% – have even developed a well-articulated data strategy that culture could eat for breakfast. There is plenty of work to be done.

80% of data challenges are people/process based!
&
DG is the only resource to address these challenges



2018

Driving Innovation with Data
Competing on data and analytics
Managing data as a business asset
Created a data-driven organization
Forged a data culture



2023 → 2024 → 2025

Source: Big Data and AI Executive Survey by Randy Bean and Thomas Davenport: <https://www.randybeandata.com>

© Copyright 2025 by Peter Allen Slide 6 27

Corporate Governance

- "Corporate governance - which can be defined narrowly as the relationship of a company to its shareholders or, more broadly, as its relationship to society....", Financial Times, 1997.
- "Corporate governance is about promoting corporate fairness, transparency and accountability" James Wolfensohn, World Bank, President Financial Times, June 1999.
- "Corporate governance deals with the ways in which suppliers of finance to corporations assure themselves of getting a return on their investment", The Journal of Finance, Shleifer and Vishny, 1997.



<https://randybeandata.com>

28

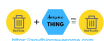
IT Governance

- "Putting structure around how organizations **align IT strategy with business strategy**, ensuring that companies stay on track to achieve their strategies and goals, and implementing good ways to measure IT's performance.
- It makes sure that all stakeholders' interests are taken into account and that processes provide **measurable results**.
- Framework should answer some key questions, such as how the IT department is functioning overall, what **key metrics** management needs and what **return** IT is giving back to the business from the investment it's making." CIO Magazine (May 2007)
- IT Governance Institute, 5 foci:
 - Strategic Alignment
 - Value Delivery
 - Resource Management
 - Risk Management
 - Performance Measures



7 Data Governance Definitions

- The formal orchestration of people, process, and technology to enable an organization to leverage data as an enterprise asset – **The MDM Institute**
- A convergence of data quality, data management, business process management, and risk management surrounding the handling of data in an organization – **Wikipedia**
- A system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods – **Data Governance Institute**
- The execution and enforcement of authority over the management of data assets and the performance of data functions – **KiK Consulting**
- A quality control discipline for assessing, managing, using, improving, monitoring, maintaining, and protecting organizational information – **IBM Data Governance Council**
- Data governance is the formulation of policy to optimize, secure, and leverage information as an enterprise asset by aligning the objectives of multiple functions – **Sunil Soares**
- The exercise of authority and control over the management of data assets – **DM BoK**



What is Data Governance?

Managing Data with Guidance

Go Ask
Anyone!™

*Would
you
want
your
sole,
non-
depletable,
non-
degrading,
durable,
strategic
asset
managed
without
guidance?*



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

31

Data Governance is

Managing Data **Decisions** with Guidance

Go Ask
Anyone!™

*Would
you
want
your
sole,
non-
depletable,
non-
degrading,
durable,
strategic
asset
managed
without
guidance?*



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

32

Bad Data Decisions for Big Decisions

<https://managedservicesjournal.com/articles/bad-data-makes-bad-decisions-58-of-leaders-report-companies-using-inaccurate-data-for-big-decisions/#:~:text=for big decisions-,Bad data makes bad decisions: 58% of leaders report companies, to global study from SoftServe.>



HOME | NEWS | AUTOMATION ▼ | SERVICES ▼ | MARKET TRENDS ▼ | PODCAST

Bad data makes bad decisions: 58% of leaders report companies using inaccurate data for big decisions

Sixty-five percent say no one at their organization fully understands the data aggregated or how to access it, according to global study from SoftServe.

By Managed Services Journal Staff - February 15, 2025



<https://anythingispossible.com>

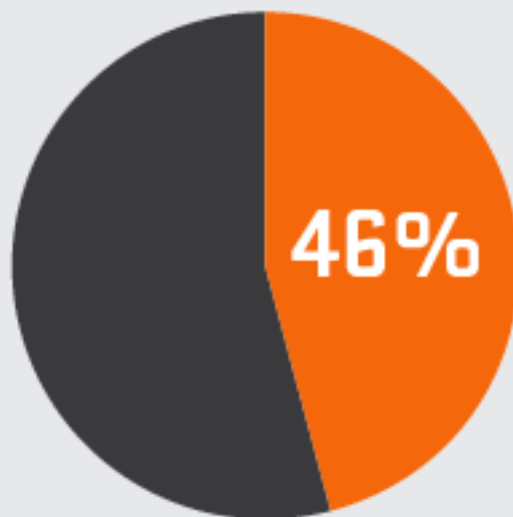
© Copyright 2025 by Peter Allen Slide 4 33

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.

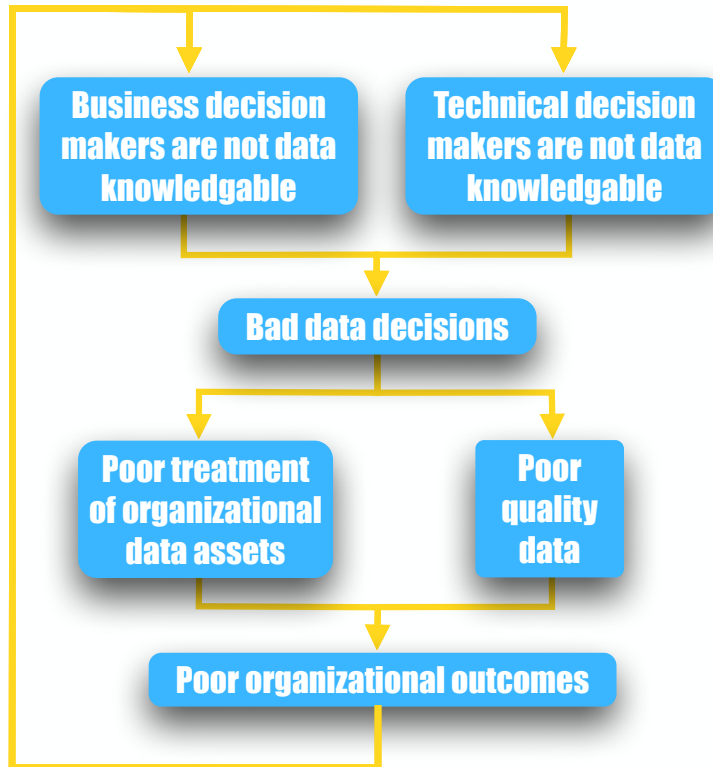
Like Comment Share



<https://anythingispossible.com>

© Copyright 2025 by Peter Allen Slide 4 34

Bad Data Decisions Spiral

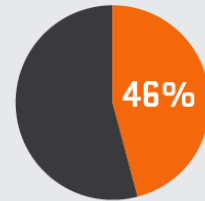


<https://creativecommons.org/licenses/by-nc-sa/4.0/>

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](#)



© Copyright 2020 by Peter Allen Slide 6 35

What are some of the bad data decisions made by walmart recently?

It's important to preface this by saying that I cannot provide specific, confirmed examples of "bad data decisions" made by Walmart without access to internal company information or publicly disclosed incidents with detailed data-related analyses. Companies like Walmart keep many of their data challenges confidential for competitive and operational reasons.

Inaccurate Demand Forecasting

- The Consequences:
- Overstocking: Walmart ends up with excess inventory of certain items, leading to storage costs, markdowns, and potential waste.
- Understocking: Walmart fails to stock enough of popular items, resulting in lost sales, customer dissatisfaction, and customers going to competitors.

Poor Inventory Data Management

- The Consequences:
- "Phantom Inventory": The system shows an item is in stock when it's not, leading to frustrated customers (especially in e-commerce) and order cancellations.
- Inefficient Replenishment: Stores don't get restocked with the right items at the right time, leading to empty shelves and lost sales.

- Increased Shrinkage: Discrepancies in inventory data can contribute to losses from theft, damage, or errors.
- **Ineffective Customer Data Analysis**
- The Consequences:
- Ineffective Marketing: Walmart sends out irrelevant promotions or fails to target the right customers, leading to wasted marketing spend and lower conversion rates.
- Poor Customer Experience: Walmart doesn't understand individual customer needs, leading to a less personalized and satisfying shopping experience.
- Missed Opportunities: Walmart fails to identify emerging trends or customer demands, missing out on opportunities to develop new products or services.

Data Breaches and Security Lapses

- The Consequences:
- Data Breaches: Sensitive data is exposed to unauthorized access, leading to financial losses, reputational damage, and legal penalties.
- Loss of Customer Trust: Customers lose confidence in Walmart's ability to protect their data, leading to decreased loyalty and potential customer churn.

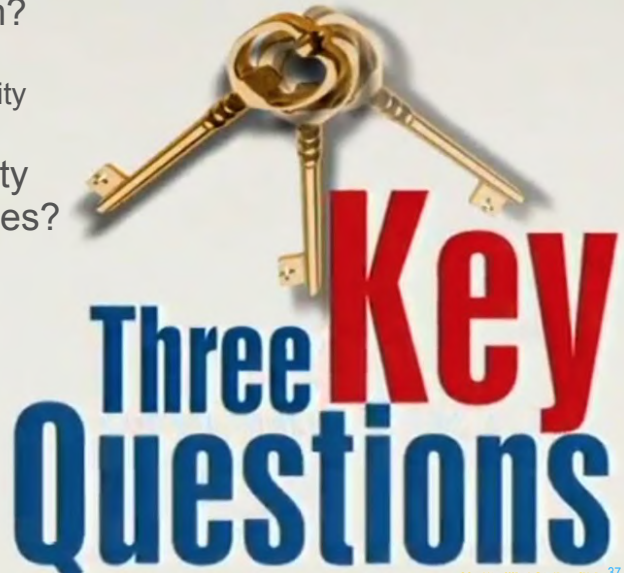


<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 6 36

Three Key Data Governance Questions

1. Is the quality of the data in the new system forecast to be of better quality than the data in the old system?
 - Lift and shift does not improve data quality
2. Are we able to formulate plans to obtain significant new value from this data in the new system?
 - If not we are lacking required specificity
3. Does this afford us an opportunity to consolidate data and data types?
 - Many system capabilities meet 'requirements' - looking at the opportunity to simplify an existing environment provides additional value



© Copyright 2020 by Peter Allen Slide 6 37

Forbes



- \$6b market value
- \$30b AAdvantage

- \$9b market value
- \$22b MileagePlus

EDITORS' PICK | Jul 15, 2020, 09:00am EDT

How Airlines Make Billions From Monetizing Frequent Flyer Programs



JT Genter Advisor Cont
Advisor Contributor Gro
Personal Finance

**How to unlock \$24b
and \$13b in data value?**



<https://www.forbes.com/sites/advisor/2020/07/15/how-airlines-make-billions-from-monetizing-frequent-flyer-programs/?sh=66da87a614e9>

© Copyright 2020 by Peter Allen Slide 6 38

How much data value is locked inside of airline frequent flyer programs and why is it so hard for them to realize it?



The amount of data value locked inside airline frequent flyer programs is immense, potentially worth billions of dollars. It's a treasure trove of information about customer behavior, preferences, and travel patterns. However, airlines often struggle to fully realize this value due to a combination of factors:

1. Data Silos and Legacy Systems:

- **Fragmented Data:** Frequent flyer data is often housed in separate systems from other crucial data like ticketing, customer service interactions, and even ancillary purchases (seat upgrades, baggage fees). Integrating these data silos is technically complex and expensive due to legacy systems.
- **Lack of a Unified Customer View:** Without a unified view, airlines struggle to understand the complete customer journey and personalize offers effectively. They might know a member flies frequently, but not why (business vs. leisure) or what other products they might be interested in.

2. Difficulty in Data Analysis and Interpretation:

- **Data Overload:** The sheer volume of frequent flyer data can be overwhelming. Airlines need sophisticated analytics capabilities to sift through the noise and identify actionable insights.
- **Lack of Skilled Analysts:** Finding and retaining data scientists and analysts with the expertise to extract meaningful insights from complex datasets is a challenge.
- **Limited Investment in Analytics Infrastructure:** Many airlines haven't invested sufficiently in the technology needed to analyze and leverage their data effectively.

3. Privacy Concerns and Regulations:

- **Data Privacy:** Airlines must be extremely careful to comply with data privacy regulations (GDPR, CCPA, etc.) when using customer data. This can limit how they can use and share the information.
- **Building Trust:** Customers are increasingly concerned about how their data is used. Airlines need to be transparent and build trust with their frequent flyers to avoid negative backlash.

4. Business Model Challenges:

- **Focus on Traditional Revenue Streams:** Historically, airlines have focused on traditional revenue streams (ticket sales) and haven't fully embraced the potential of data-driven business models.
- **Resistance to Change:** Changing established processes and organizational structures to fully leverage data can be difficult, especially in large, complex organizations.
- **Measuring ROI:** Demonstrating the return on investment for data-driven initiatives can be challenging, making it difficult to secure funding for these projects.

5. Lack of Clear Strategy:

- **No Data-Driven Vision:** Many airlines lack a clear strategy for how to use their data to drive business value. Without a vision, it's difficult to prioritize investments and initiatives.
- **Lack of Executive Sponsorship:** Data-driven initiatives often require buy-in from senior leadership. Without executive sponsorship, these projects may struggle to get the resources and support they need.

How Airlines Could Realize the Value:

- **Invest in Data Integration and a Unified Customer View:** Breaking down data silos is the first step.
- **Develop Advanced Analytics Capabilities:** Invest in the tools and talent needed to analyze data effectively.
- **Personalization:** Use data to personalize offers, communications, and travel experiences.
- **Ancillary Revenue Opportunities:** Identify opportunities to generate revenue from ancillary products and services based on customer preferences.
- **Partnerships:** Collaborate with other businesses to offer targeted promotions and loyalty programs.
- **Improved Customer Service:** Use data to anticipate customer needs and provide proactive support.

Unlocking the value of frequent flyer data requires a long-term commitment to data governance, technology, and talent. Airlines that can overcome these challenges will be well-positioned to gain a competitive advantage in the increasingly data-driven travel industry.



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 6 39

How to determine what to manage formally?

Too much requires expensive and slow bureaucracy ←

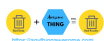
→ Too little misses opportunities

All inputs are data

The
Organizational
Data Machine

All outputs are data

Interoperability is the primary value determinant



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 6 40

Data Governance Goals



Too much requires expensive and slow bureaucracy ←

→ Too little misses opportunities

All inputs are data

Known

Unknown

All outputs are data

Interoperability is the primary value determinant



© Copyright 2020 by Peter Allen Slide 41

Data Governance Goals



Too much requires expensive and slow bureaucracy ←

→ Too little misses opportunities

All inputs are data

Known

Unknown

All outputs are data

Interoperability is the primary value determinant



© Copyright 2020 by Peter Allen Slide 42

Data Governance Goals



Too much requires expensive and slow bureaucracy ←

→ Too little misses opportunities

All inputs are data

All outputs are data

Known

Unknown

Interoperability is the primary value determinant



© Copyright 2020 by Peter Allen Slide 43

What is Strategy?

strat·e·gy

/ˈstrætəj/

noun

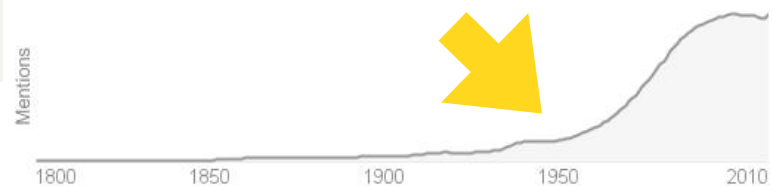
1. a plan of action or policy designed to achieve a major or overall aim.

"time to develop a coherent economic strategy"

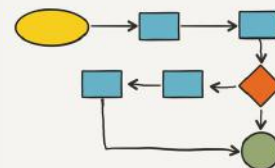
synonyms: master plan, grand design, game plan, plan (of action), action plan, policy, program; More

A thing

Use over time for: Strategy



- Current use derived from military
 - **a pattern in a stream of decisions** [Henry Mintzberg]



PROCESS



© Copyright 2020 by Peter Allen Slide 44

Every Day Low Price



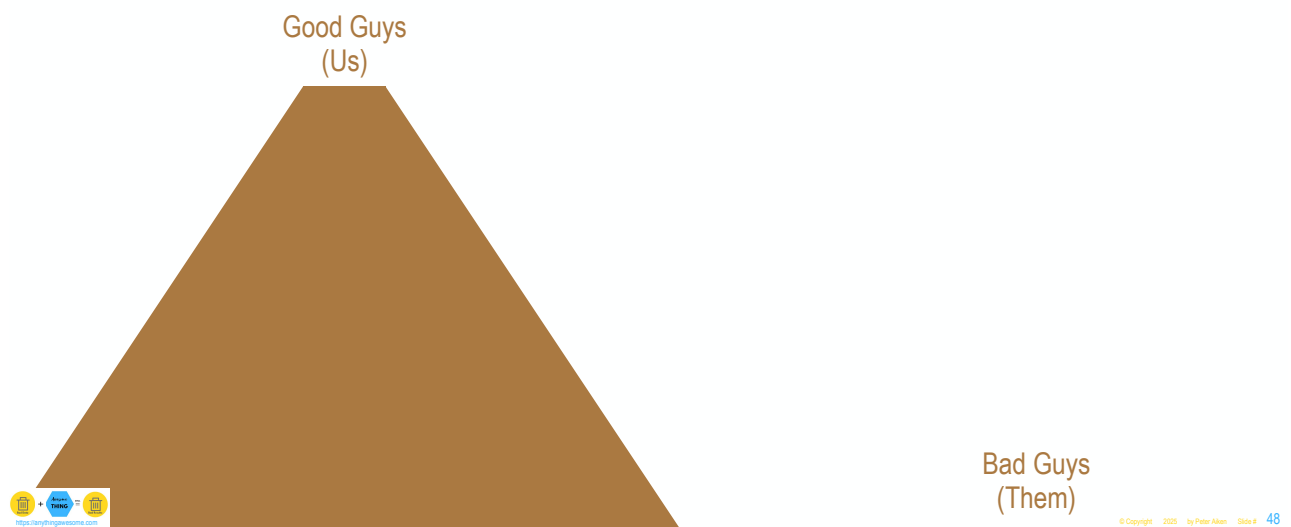
Wayne
Gretzky's
Strategy

He skates to where he
thinks the puck will be ...

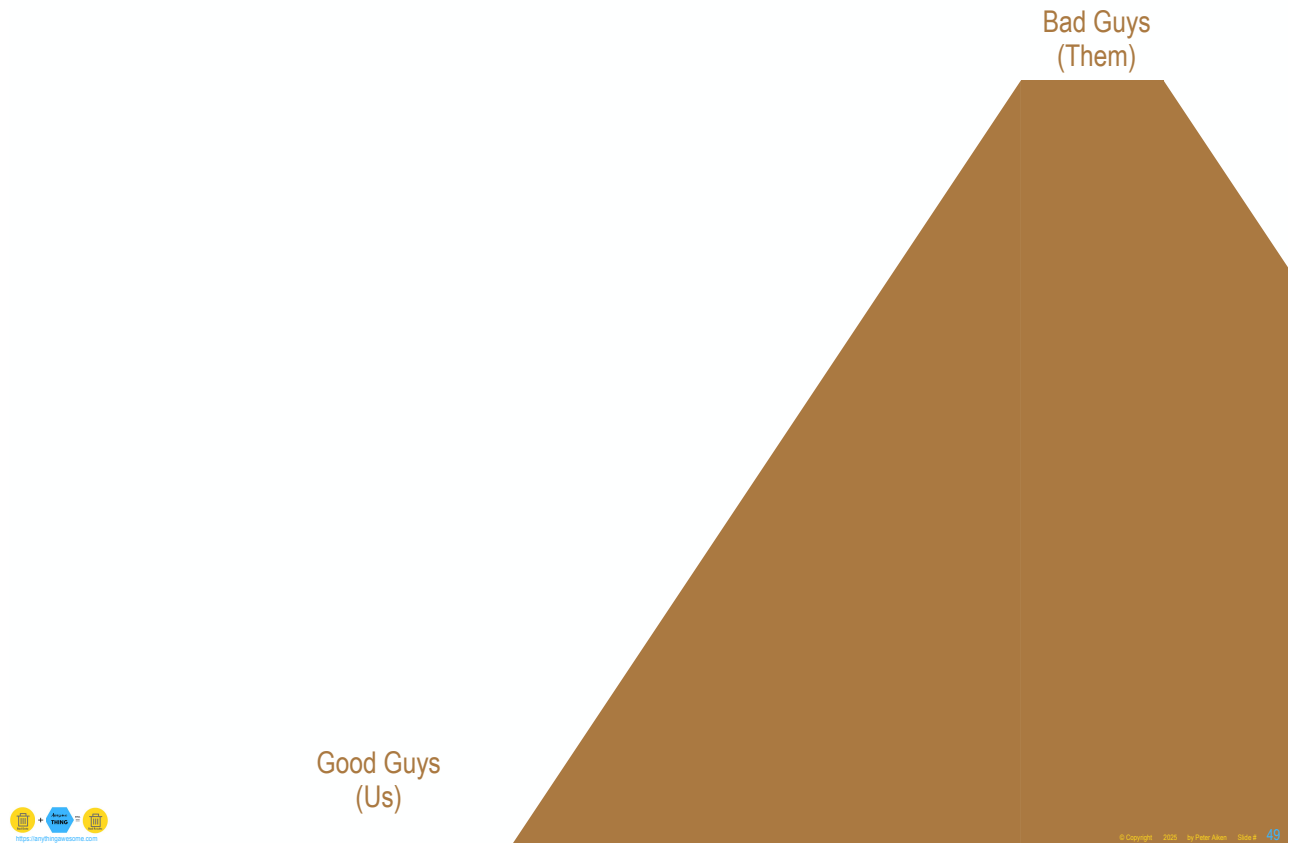
Strategy Example 3



Strategy Example 3



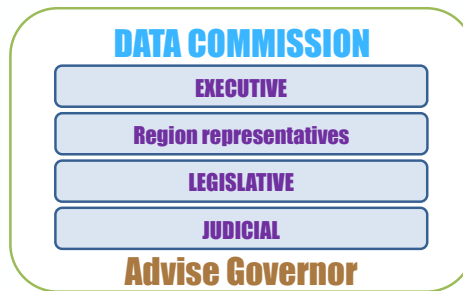
Strategy Example 3



Strategy Guides Workgroup Activities

A pattern
in a stream
of decisions

Complex Data Governance Environment



- Data sharing and analytics
- Identify goals and objectives
 - Prioritize initiatives
 - Study & report
 - Recommend changes to budget and code



execute

- Define, approve, and communicate data strategies, policies, standards, rules, guidelines, & best practices
- Provide a governance, policy, and technology framework
- Define agency data governance responsibilities
- Encourage & facilitate data sharing
- Facilitate coordination to prevent duplication
- Coordinate policy and technology proposals and recommendations
- Administer and manage the commonwealth data trust
- Track and enforce compliance and conformance
- Oversee dissemination of open data

Data Governance Council

- Liaise between agency operations & CDO
- Advise CDO on technology, policy, and governance strategies
- Administer data governance policies set by the board
- Implement data sharing & analytics projects
- Review open data assets
- Report progress & compliance to the Board

Advise CDO

Executive data board

- Translate commonwealth goals to agency performance targets
- provide resources
- Remove organizational obstacles
- appoint data governance council members
- Oversee the data governance council
- Oversee data sharing & analytics projects

Oversee council



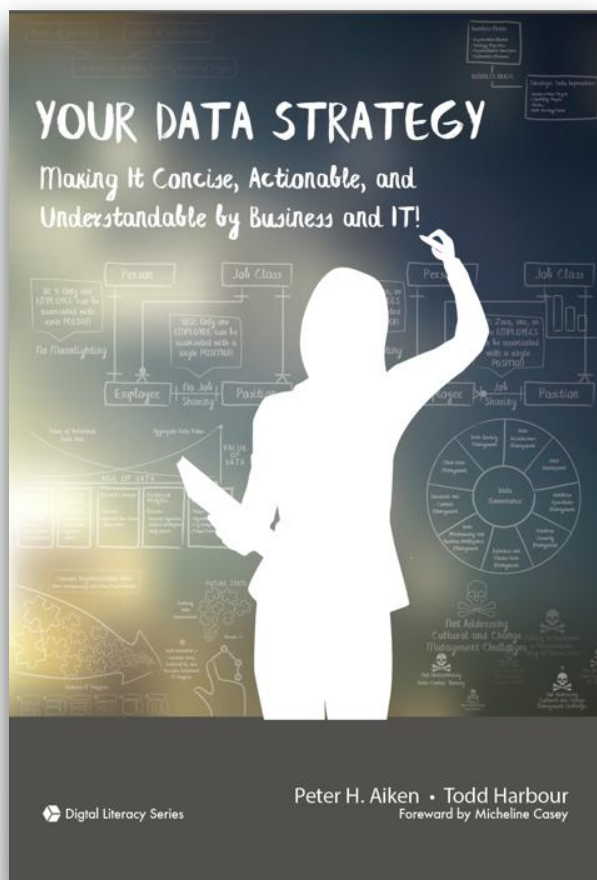
<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Aiken Slide 51

Your Data Strategy

- Highest level data guidance available ...
- Focusing data activities on business-goal achievement ...
- Providing guidance when faced with a stream of decisions or uncertainties

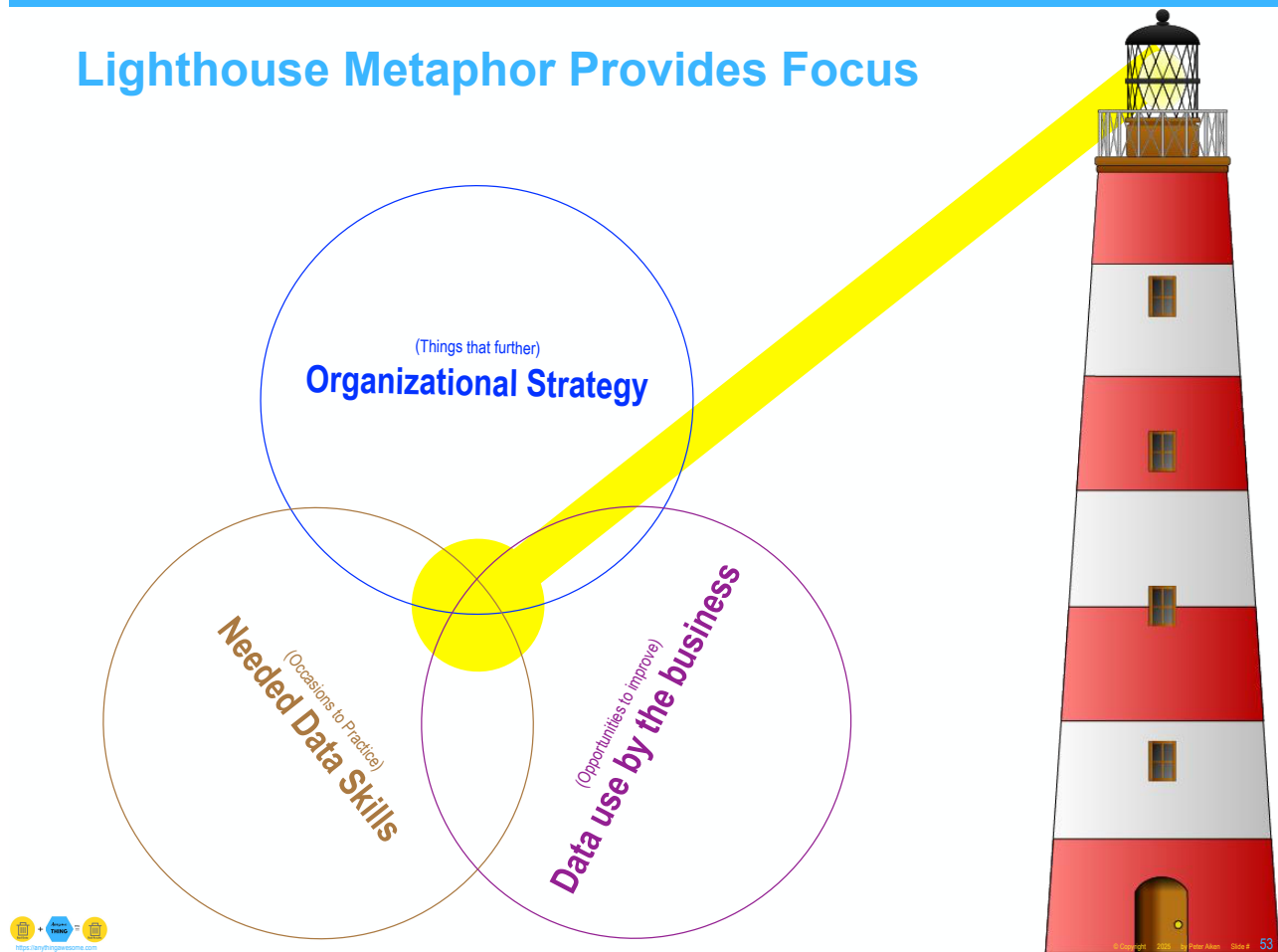
- Data strategy most usefully articulates how data can be best used to support organizational strategy
- This usually involves a balance of remediation and proactive measures



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Aiken Slide 52

Lighthouse Metaphor Provides Focus



Why is organizational strategy such an important driver of data governance success?



Organizational strategy is a critical driver of data governance success because it provides the "why" behind the effort. Without a clear link to strategic goals, data governance can easily become a technical exercise divorced from business value, leading to its ultimate failure. Here's a breakdown of why this connection is so crucial:

1. Alignment and Focus:

- **Prioritization:** Organizational strategy identifies the key objectives and priorities of the business. Data governance efforts must align with these priorities to ensure that resources are focused on the most impactful data domains and initiatives. If data governance isn't tied to what the business is trying to achieve, it becomes a cost center, not a value creator.
- **Strategic Relevance:** When data governance is aligned with strategy, it becomes clear how improved data quality, accessibility, and security contribute to achieving business goals. This makes it easier to justify investments in data governance and secure buy-in from stakeholders.

2. Motivation and Buy-In:

- **Shared Purpose:** Connecting data governance to strategic goals provides a shared purpose and motivates stakeholders to participate actively. When people understand how good data governance helps the organization succeed, they are more likely to support the program and adhere to policies.
- **Executive Sponsorship:** Executive sponsorship is essential for data governance success. When senior leaders see a clear link between data governance and strategic objectives, they are more likely to champion the program and allocate necessary resources.

3. Resource Allocation:

- **Targeted Investment:** Organizational strategy helps determine where to invest resources in data governance. By focusing on the data domains that are most critical to achieving strategic goals, organizations can maximize the return on their investment.

- **Justification for Funding:** Linking data governance to strategic initiatives makes it easier to justify funding requests. When the connection is clear, it's easier to demonstrate the value of data governance and secure the necessary resources.

4. Measurement and Evaluation:

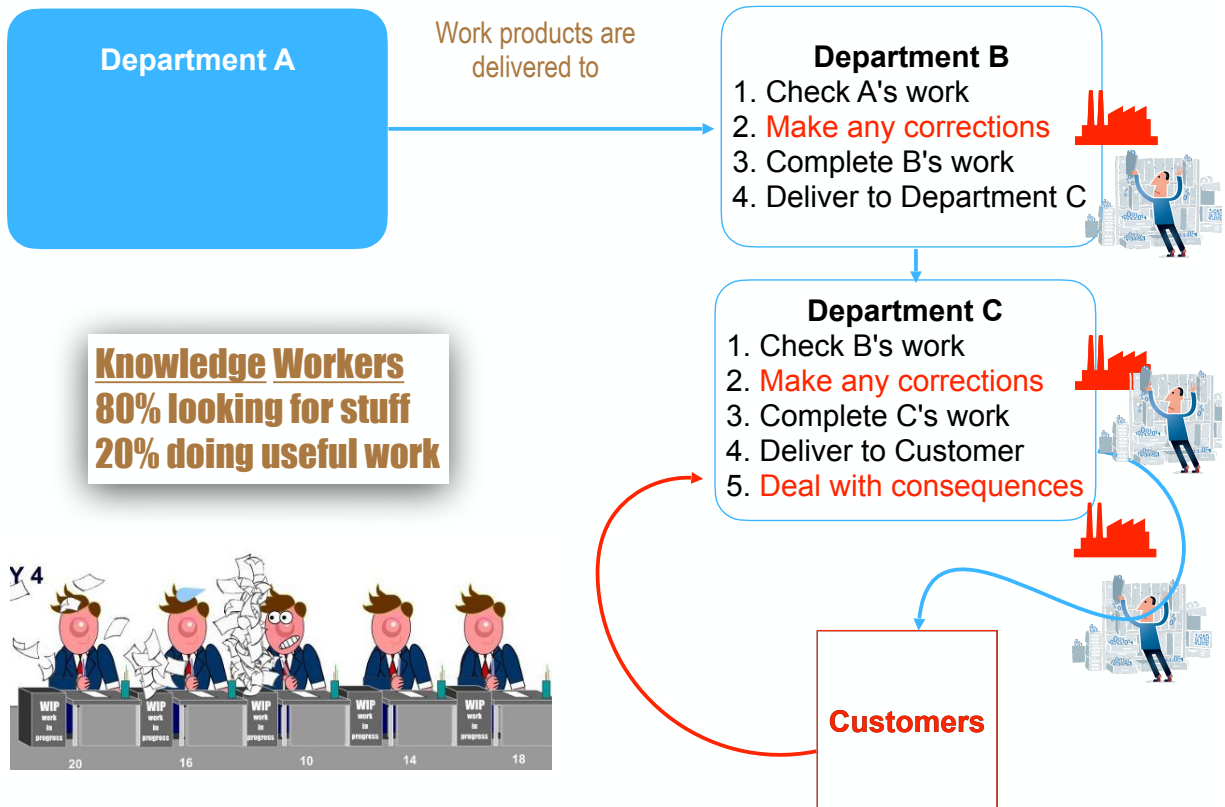
- **Defining Success:** Organizational strategy provides a framework for measuring the success of data governance efforts. By tracking how data governance contributes to achieving strategic goals, organizations can demonstrate the value of the program and identify areas for improvement.
- **Demonstrating ROI:** Connecting data governance to business outcomes makes it easier to calculate the return on investment. This helps justify continued investment in the program and ensures its long-term sustainability.

5. Cultural Change:

- **Data-Driven Culture:** When data governance is aligned with strategy, it reinforces a data-driven culture within the organization. People begin to understand the importance of data quality and use data more effectively to make decisions and achieve their goals.
- **Shared Understanding:** A clear link to strategy helps create a shared understanding of the value of data across the organization. This fosters collaboration and breaks down data silos.

In essence, organizational strategy acts as the compass and roadmap for data governance. It provides direction, focus, and motivation, ensuring that data governance efforts are aligned with the overall objectives of the business. Without this alignment, data governance can easily become a fragmented and ineffective exercise. It's not enough to simply have "good data"; the data must be good for a specific purpose, and that purpose is defined by the organization's strategy.

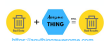
Hidden Data Factories



https://en.wikipedia.org/wiki/Theory_of_constraints

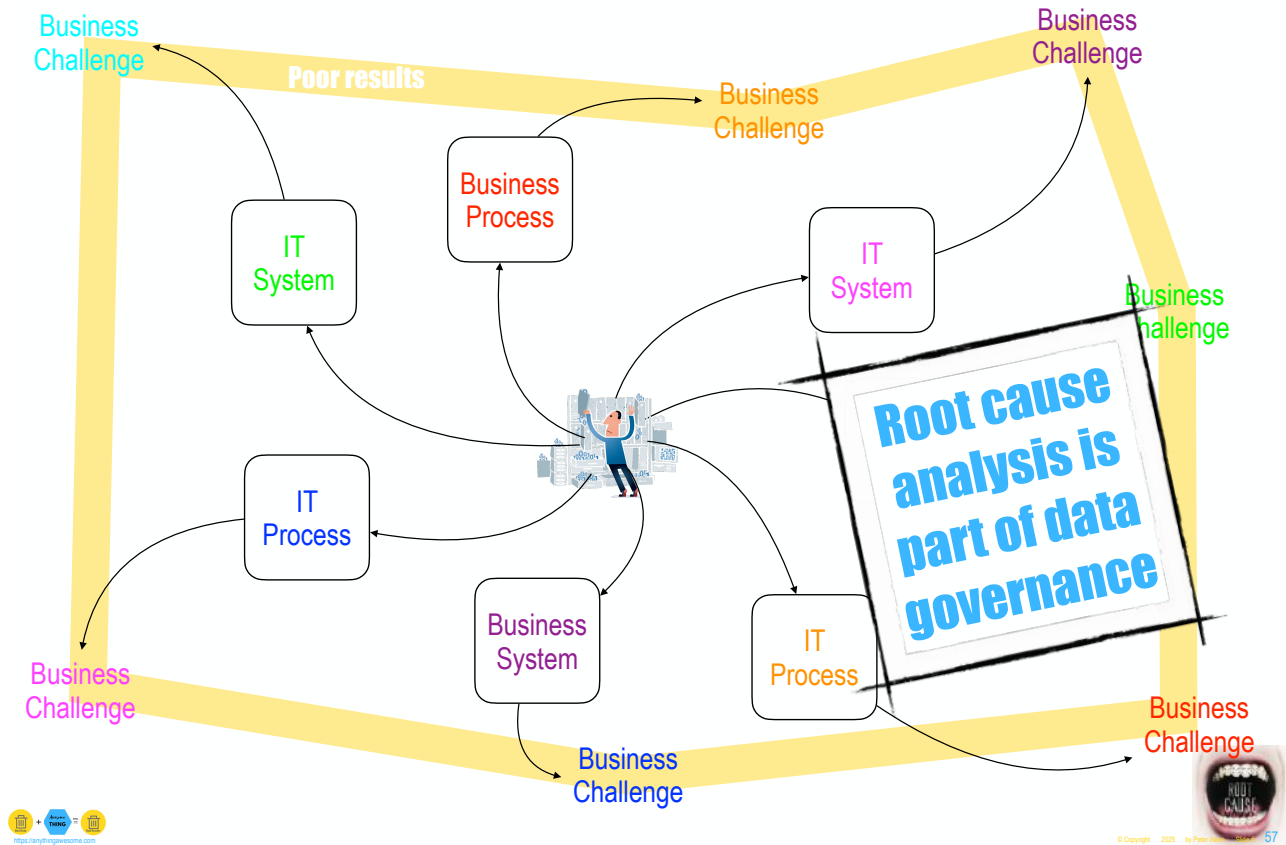
<https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year> © Copyright 2020 by Peter Allen Slide 6 55

Poor data manifests as multifaceted organizational challenges

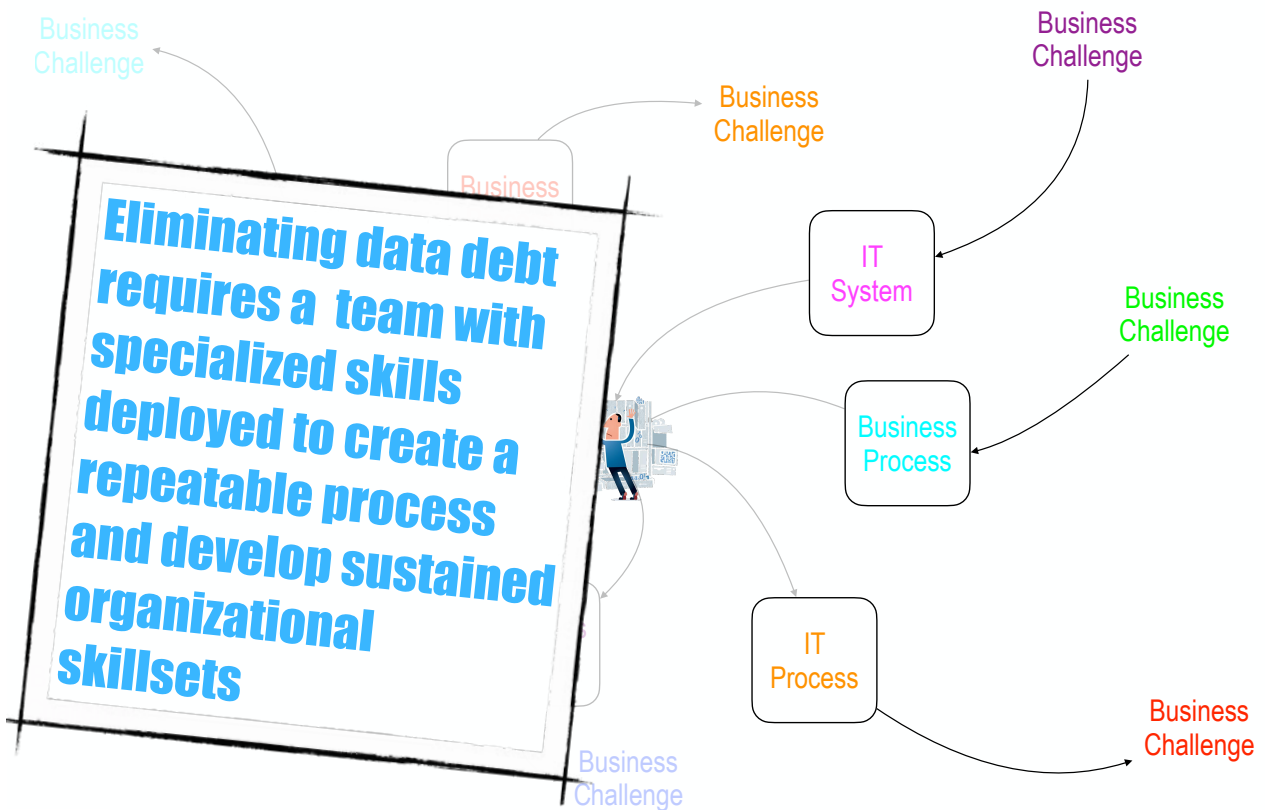


© Copyright 2020 by Peter Allen Slide 6 56

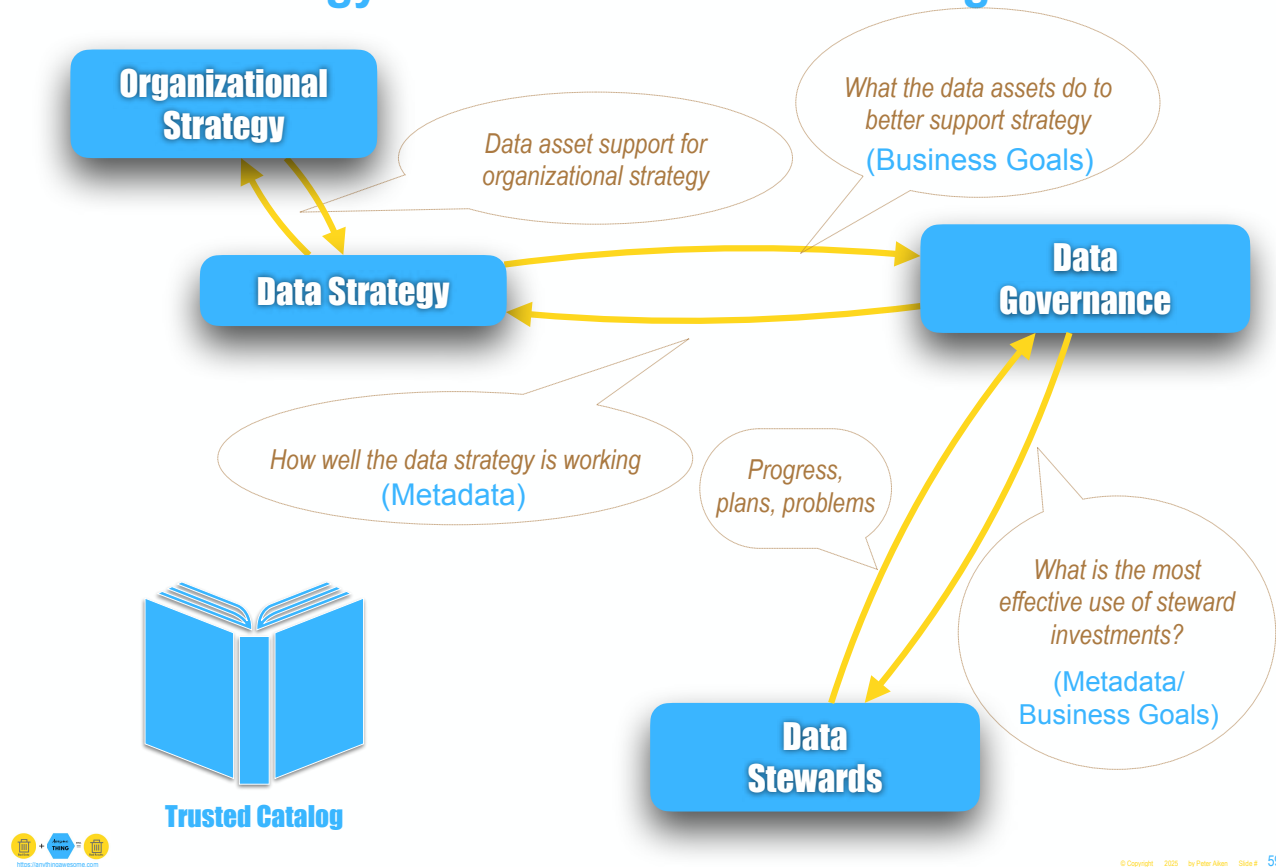
Poor data manifests as multifaceted organizational challenges



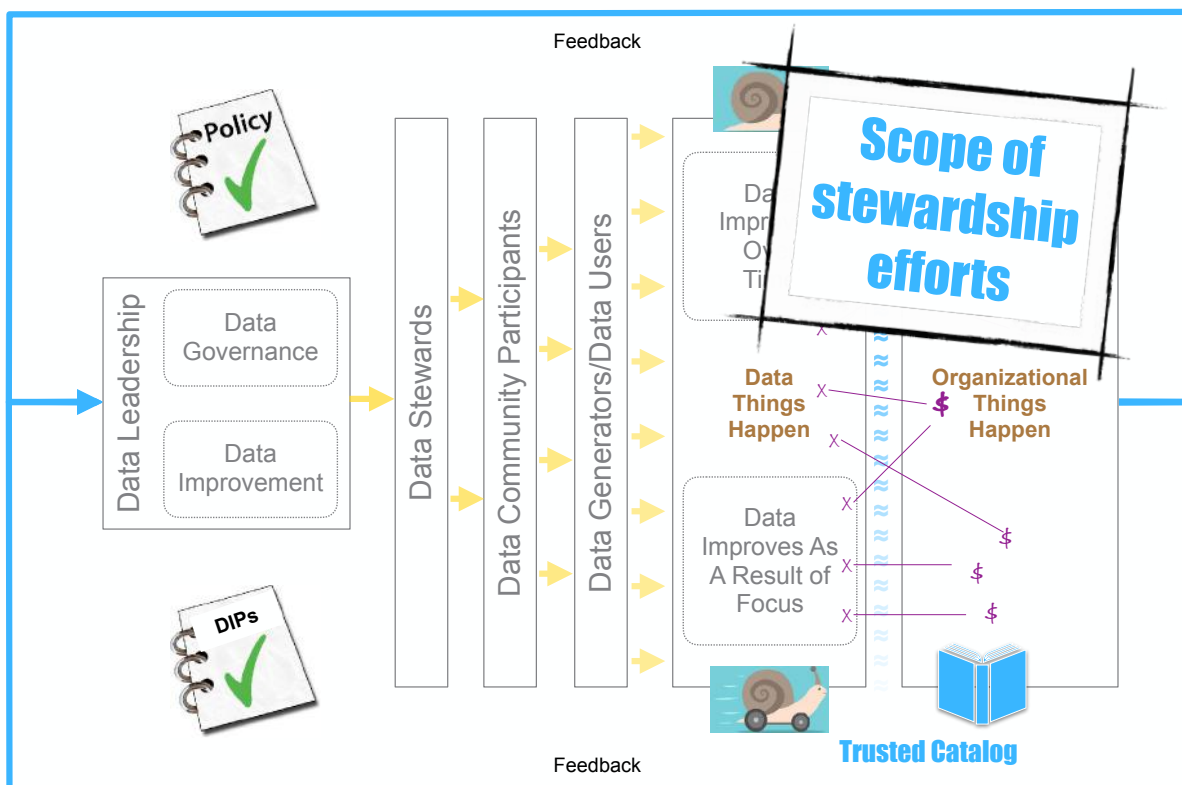
Consistency Encourages Quality Analysis



Data Strategy and Governance in Strategic Context



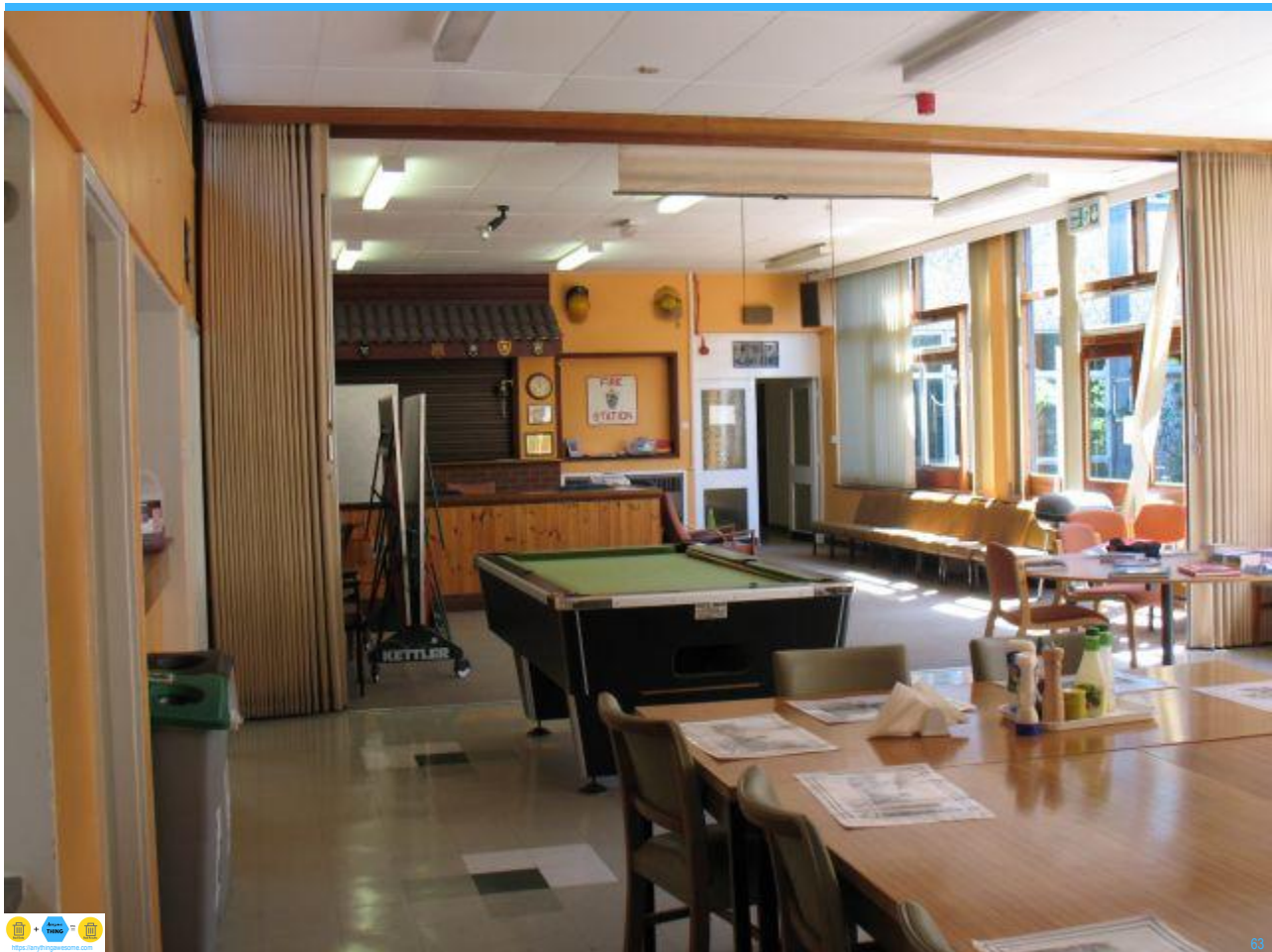
Data Governance Role: Produce systemic organizational changes that impact data and work practices over time





The MacGyver approach to DG uses paperclips and duct tape

MACGYVER



Why is Data Governance important?

- Cost organizations millions each year in
 - Productivity
 - Redundant and siloed efforts
 - Poorly thought out hardware and software purchases
 - Delayed decision making using inadequate information
 - Reactive instead of proactive initiatives
 - 20-40% of IT spending can be reduced through better data governance



Why is data governance important to organizations? ✨

Data governance is crucial for organizations for a multitude of reasons, all boiling down to maximizing the value of data while minimizing its risks. Here's a breakdown:

1. Improved Data Quality:

- **Accuracy:** Data governance establishes standards and processes to ensure data is accurate and reliable. This leads to better decision-making and operational efficiency.
- **Completeness:** It addresses missing or incomplete data, ensuring that information is available when and where it's needed.
- **Consistency:** Data governance promotes consistency in data definitions and formats across the organization, making it easier to integrate and analyze data from different sources.
- **Timeliness:** It ensures data is up-to-date and relevant, providing a current view of the business.

2. Enhanced Decision-Making:

- **Data-Driven Insights:** With high-quality data, organizations can generate accurate reports and analytics, leading to better informed decisions.
- **Strategic Advantage:** Access to reliable data allows organizations to identify trends, understand customer behavior, and make strategic decisions that drive growth and innovation.

3. Reduced Costs and Increased Efficiency:

- **Eliminating Redundancy:** Data governance helps identify and eliminate duplicate data, reducing storage costs and wasted effort.
- **Streamlining Processes:** It improves data management processes, making them more efficient and reducing operational costs.
- **Minimizing Errors:** Better data quality reduces the need for costly data correction and rework.



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

4. Mitigating Risks and Ensuring Compliance:

- **Data Security:** Data governance establishes security measures to protect sensitive data from unauthorized access and use, minimizing the risk of data breaches and cyberattacks.
- **Regulatory Compliance:** It helps organizations comply with data privacy regulations (GDPR, CCPA, HIPAA, etc.), avoiding fines and legal penalties.
- **Reputational Protection:** By ensuring data quality and security, data governance helps protect the organization's reputation and maintain customer trust.

5. Enabling Innovation and Growth:

- **Data as an Asset:** Data governance treats data as a valuable asset, enabling organizations to leverage data for innovation and new product development.
- **Improved Customer Experience:** By understanding customer needs and preferences through data analysis, organizations can personalize offerings and improve customer satisfaction.
- **Competitive Advantage:** Organizations with strong data governance practices gain a competitive advantage by making better decisions, operating more efficiently, and innovating faster.

6. Fostering Collaboration and Communication:

- **Shared Understanding:** Data governance promotes a shared understanding of data across the organization, breaking down data silos and fostering collaboration.
- **Improved Communication:** It establishes clear roles and responsibilities for data management, improving communication and coordination among different teams.

In short, data governance is essential for organizations to unlock the full potential of their data. It ensures that data is accurate, reliable, secure, and accessible, enabling better decision-making, reducing costs, mitigating risks, and driving innovation and growth. It's not just about managing data; it's about managing data for a purpose – to achieve organizational objectives

© Copyright 2020 by Peter Allen Slide 6 65

Program Overview

• Data's Confounding Characteristics

- Uneven understanding
- Has lead fractured views of data and to
- Increasing organizational data debt

1. Keeping DG practically focused on strategy

- This is a young profession and must
- Directly support organizational strategy by
- Improving data and its use in the short and long term

2. DG must exist at the same level as HR

- In order to achieve effectiveness,
- DG is central to DM (and central to digitization efforts)
- Must be de-coupled from IT strategy

3. Gradually add ingredients (practicing and getting better)

- Digital and data are dependent on high speed automation/data processing
- Employ a DG Frameworks to refine focus
- Plan to evolve (PDCA)

4. Data governance in action: Storytelling

- Take Aways/References/Q&A



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 6 66

Data Governance is not a Project

Common Organizational Data
(and corresponding data needs requirements)



Future State



(Version +1)

Evolve

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

Systems
Development
Activities



Create

New Organizational
Capabilities



Data programmes drive IT programs



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

What is the Difference Between?

Data Governance

- Policy level guidance
- Setting general guidelines/direction
- Example: All information not marked public should be considered confidential
- Keep in mind "firehouse" metaphor

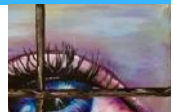
Data Management

- The business function of
 - Planning for,
 - Controlling and
 - Delivering
 - Data/information assets
- Examples: Delivering data to solve business challenges



© Copyright 2020 by Peter Allen Slide 6 69

Sample from: <https://artist.com/kathy-linden/on-outside-looking-in/?artid=4385>



External Comprehension

Everything Else Data

Data Management

Data Governance Program

Data (blah blah blah)

Data Program

Most do not appreciate the difference between Data Governance and the other data stuff that needs to be done



© Copyright 2020 by Peter Allen Slide 6 70

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>

© Copyright 2020 by Peter Allen Slide 6 71

Data Management Body of Knowledge (DM BoK V2)



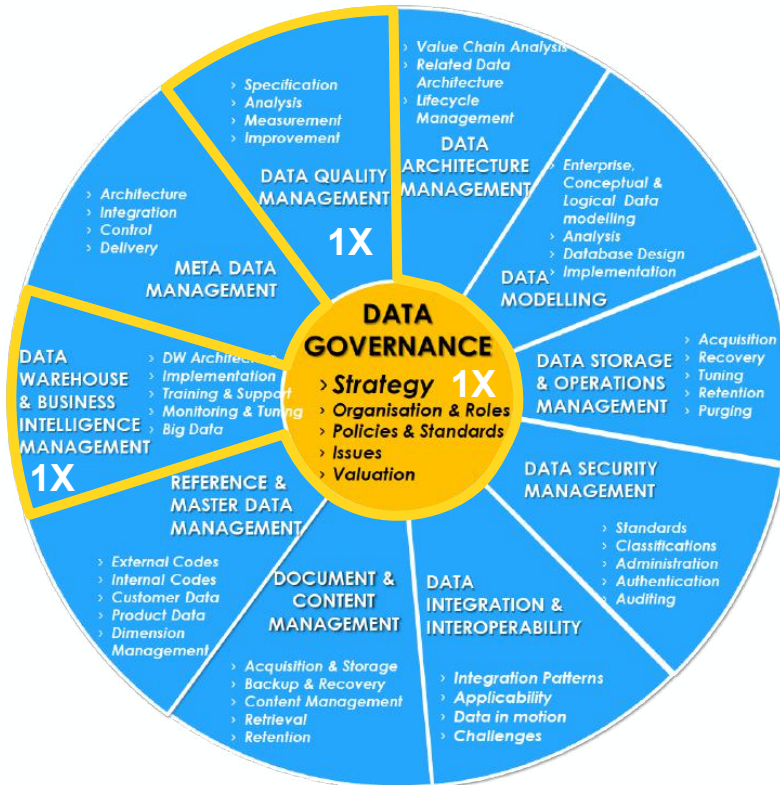
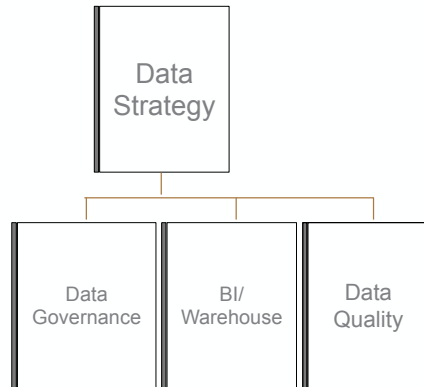
from The DAMA Guide to the Data Management Body of Knowledge 2E © 2017 by DAMA International

© Copyright 2020 by Peter Allen Slide 6 72

Iteration 1



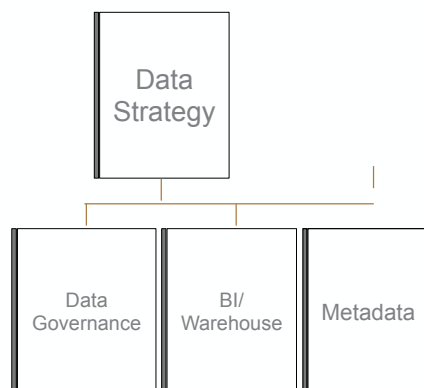
Perfecting operations in 3 data management practice areas



Iteration 2



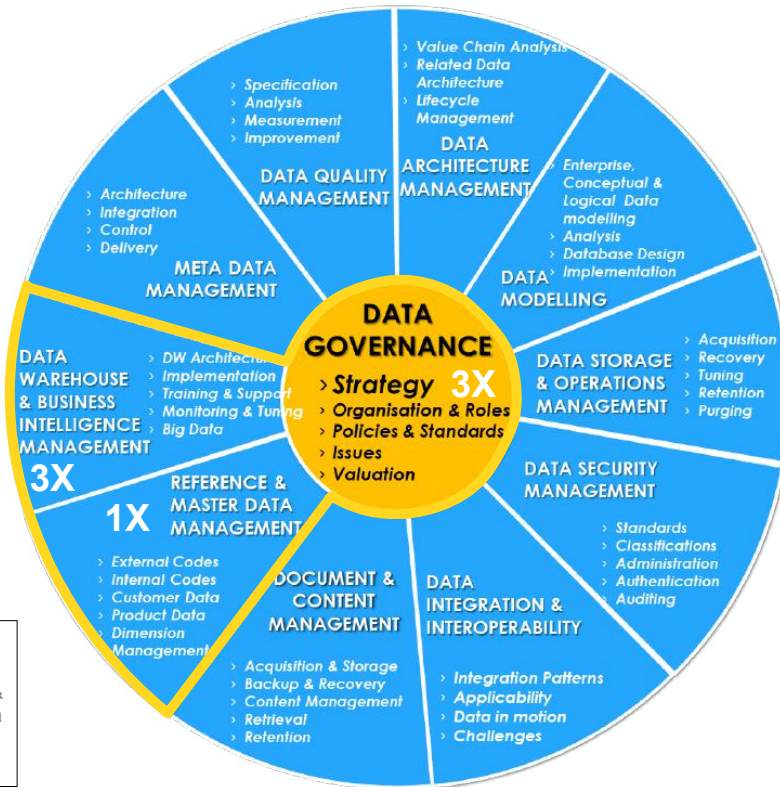
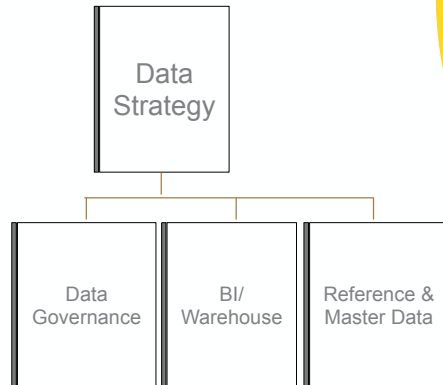
Perfecting operations in 3 data management practice areas



Iteration 3



**Perfecting
operations in 3
data management
practice areas**



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide 6 75

IT Project or Application Centric Development

Strategy



IT Projects



Data/Information



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Original articulation from Doug Bagley @ Walmart © Copyright 2020 by Peter Allen Slide 6 76

This is the wrong way to think about data strategy



Organizational Strategy

IT Strategy

Data Strategy



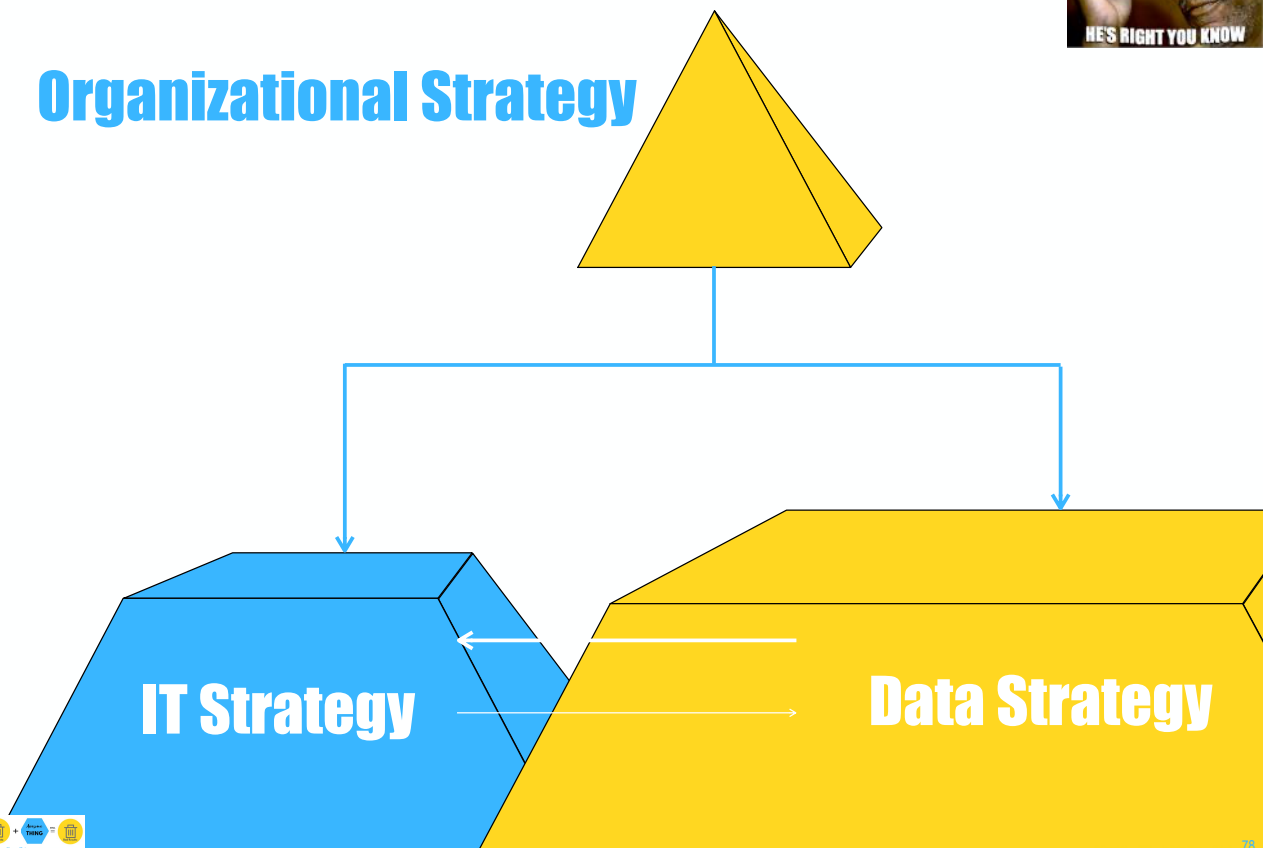
<https://creativecommons.org/licenses/by-nc-sa/4.0/>

77

This is correct ...



Organizational Strategy



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

78

IT Project or Application Centric Development

Strategy



IT Projects



Data/Information



Original articulation from Doug Bagley @ Walmart © Copyright 2020 by Peter Allen Slide 6 79

Data Centric Development

Strategy



Data/Information



IT Projects



Program Overview

- Data's Confounding Characteristics
 - Uneven understanding
 - Has lead fractured views of data and to
 - Increasing organizational data debt
- 1. Keeping DG practically focused on strategy
 - This is a young profession and must
 - Directly support organizational strategy by
 - Improving data and its use in the short and long term
- 2. DG must exist at the same level as HR
 - In order to achieve effectiveness,
 - DG is central to DM (and central to digitization efforts)
 - Must be de-coupled from IT strategy
- 3. Gradually add ingredients (practicing and getting better)
 - Digital and data are dependent on high speed automation/data processing
 - Employ a DG Frameworks to refine focus
 - Plan to evolve (PDCA)
- 4. Data governance in action: Storytelling
 - Take Aways/References/Q&A



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Alken Slide 81

Create a RACI matrix illustrating data governance roles among the business and IT users and data stewards



Activity	Business Users	IT Users	Data Stewards
Data Strategy Development	C	I	A
Data Policy Creation & Approval	C	I	A
Data Architecture Design	C	A	C
Data Modeling	C	R	C
Data Definition & Business Glossary Management	C	I	A
Data Quality Rule Definition	C	C	R
Data Quality Monitoring & Reporting	I	R	A
Data Quality Issue Resolution	C	R	A
Metadata Management	I	R	A
Data Security Policy Implementation	I	A	C
Data Access Management	I	R	A
Data Privacy Compliance	I	A	C
Data Retention & Disposal Policy	C	I	A
Data Training & Awareness	I	C	R
Data Governance Tool Selection & Implementation	C	R	A
Data Issue Escalation & Resolution	C	R	A
Data Governance Program Evaluation	C	I	A



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Alken Slide 82



Digital Insight

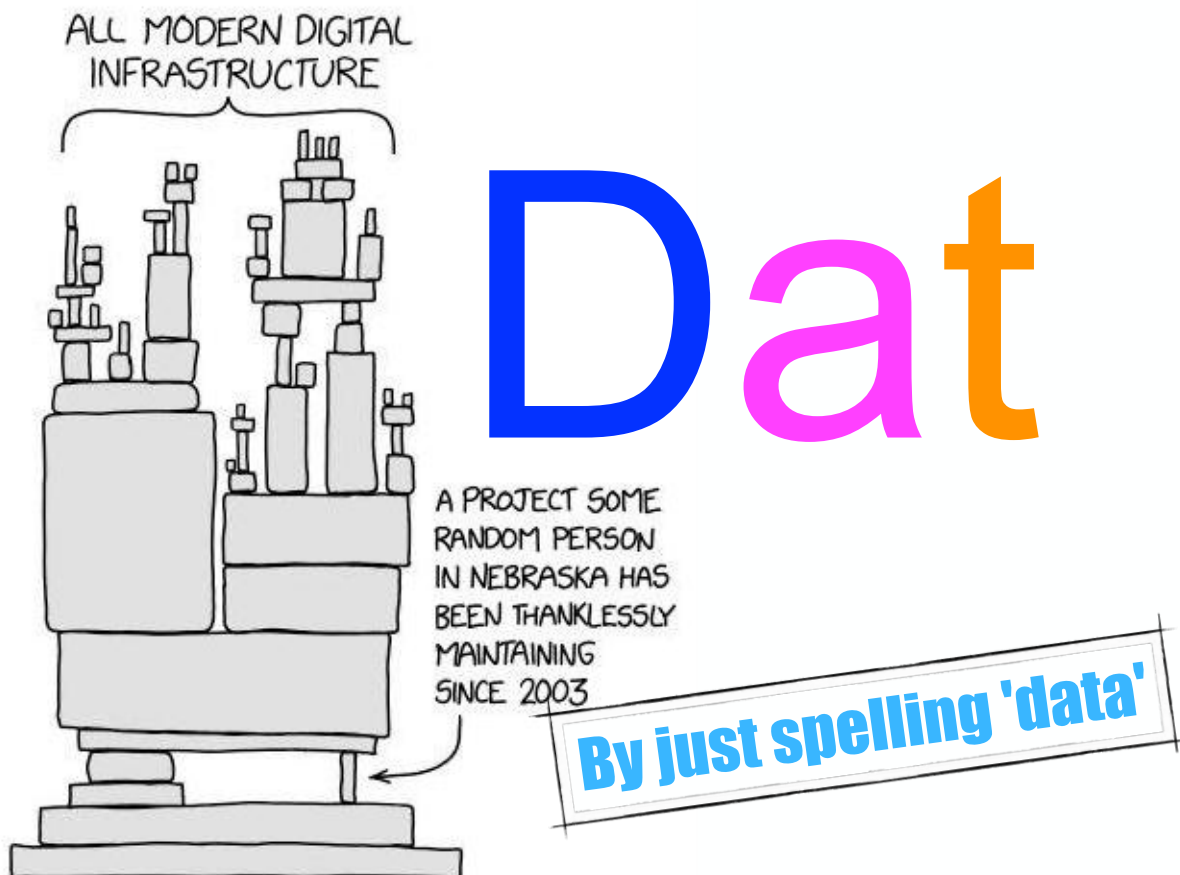


DIGITAL DATA
 — DATA — DIGITAL
 ? DATA

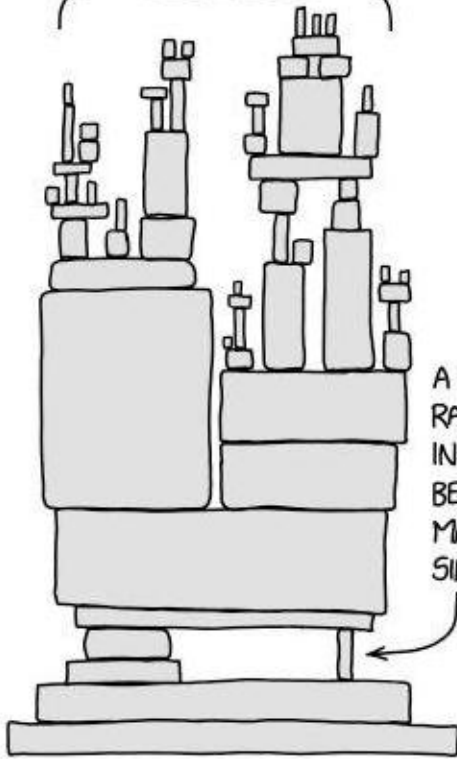
- Subtract data from digital and what do you have?
- Subtract digital from data and you still have data

Digital

It isn't possible to go digital



ALL MODERN DIGITAL
INFRASTRUCTURE



It requires more work!

Data

A PROJECT SOME
RANDOM PERSON
IN NEBRASKA HAS
BEEN THANKLESSLY
MAINTAINING
SINCE 2003

GI→GO!

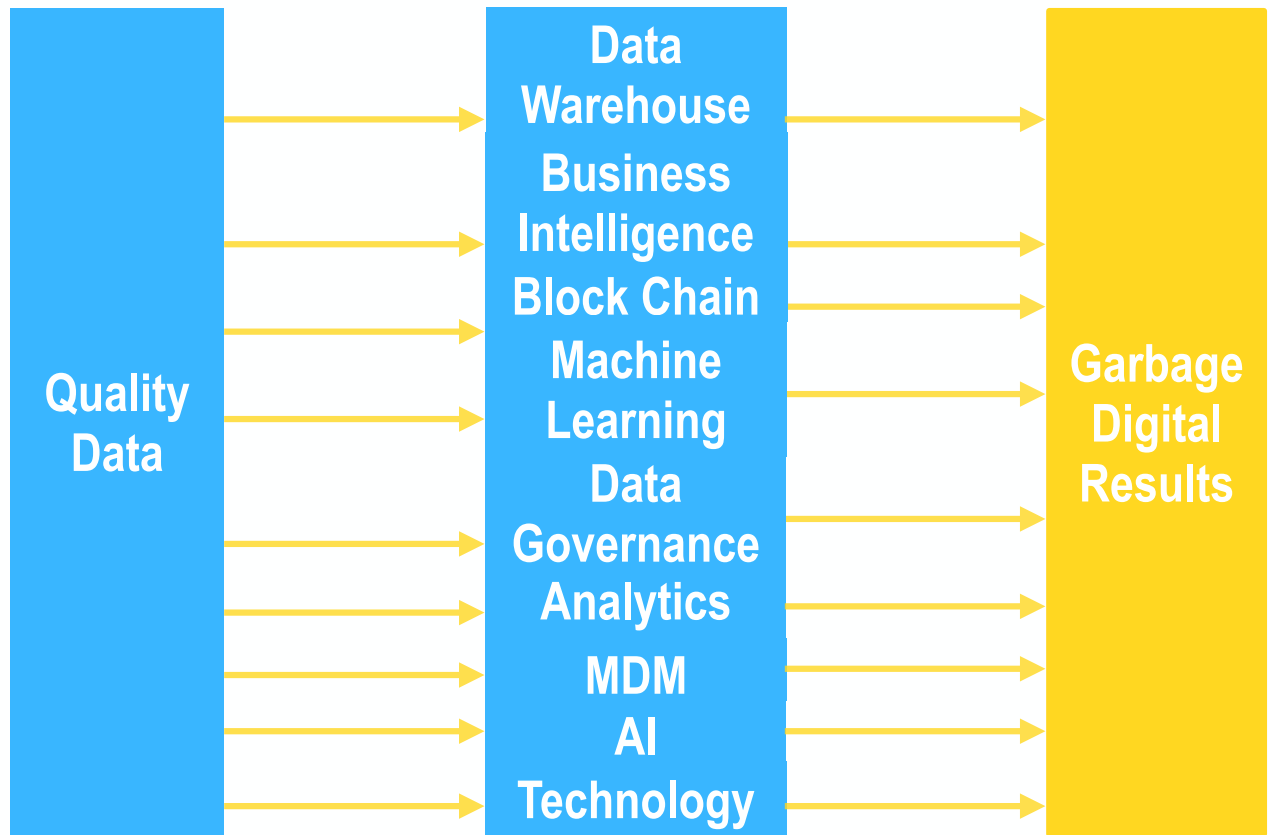
Garbage
Data

Technology

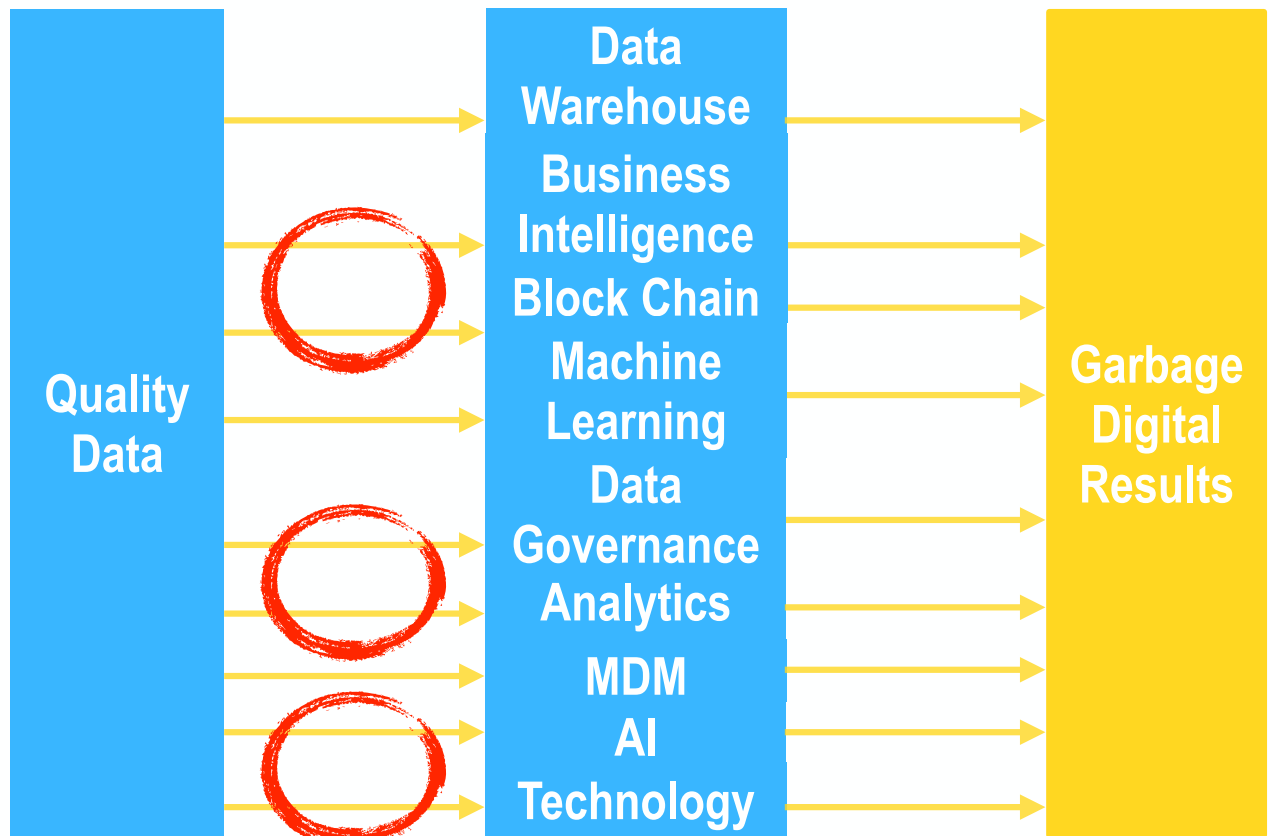
Garbage
Digital
Results

Garbage In → Garbage Out!

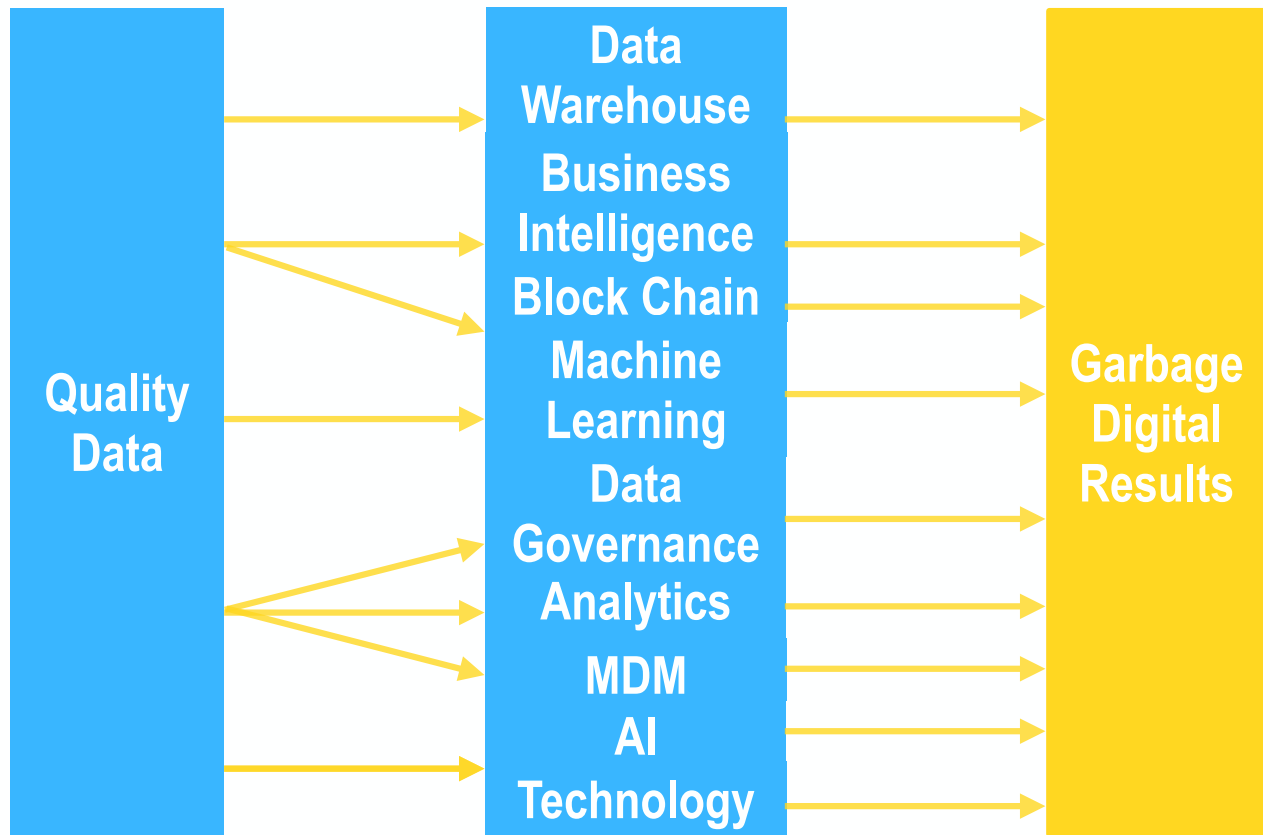
GI→GO!



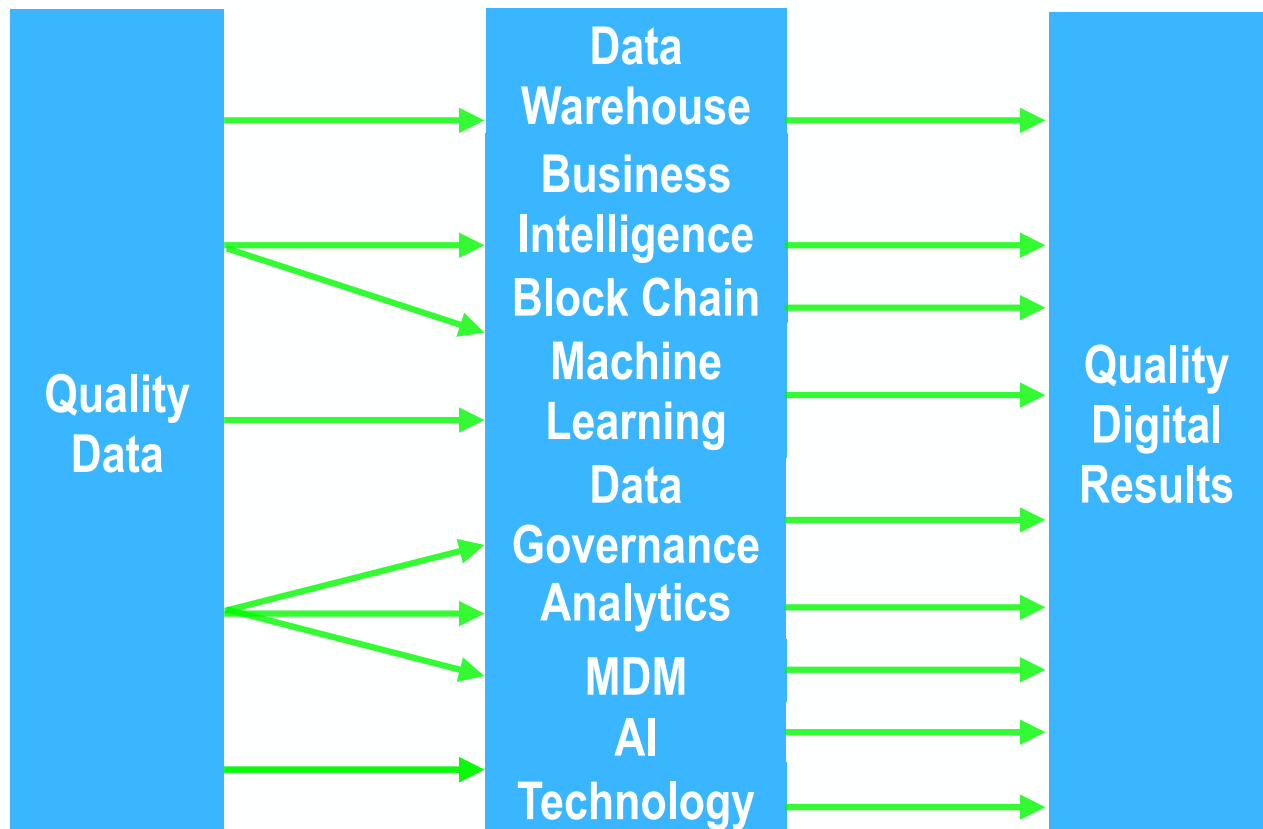
GI→GO!



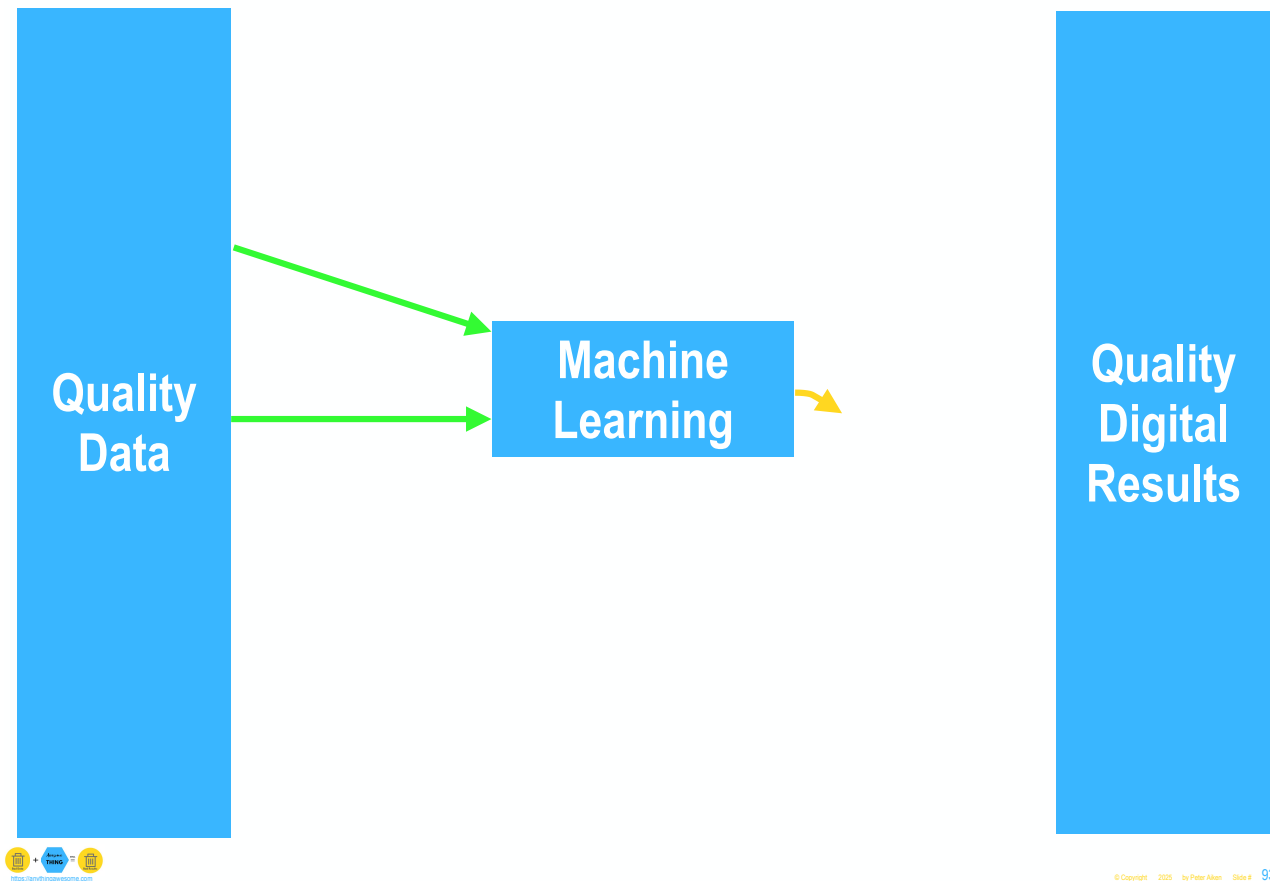
GI→GO!



QI→QO!



Today



Illustrate the concept of garbage in garbage out



Scenario: **Baking a Cake**

- Imagine you're baking a cake. The recipe is your "process," and the ingredients are your "data."

Garbage In:

- Stale Flour:** Your flour is old and has lost its leavening power.
- Rotten Eggs:** You accidentally use eggs that have gone bad.
- Incorrect Measurements:** You misread the recipe and add too much sugar and not enough milk.
- These are your "garbage in" – poor quality ingredients, representing bad data.

Process (Baking):

- You follow the recipe diligently, mixing the ingredients, pouring the batter into a pan, and baking it in the oven. The process itself might be perfect, but it's working with flawed inputs.

Garbage Out:

- Flat, dense cake: The stale flour didn't allow the cake to rise properly.
- Unpleasant taste: The rotten eggs impart a foul flavor.
- Dry and crumbly texture: The incorrect measurements further contribute to the cake's poor texture.
- This is your "garbage out" – a bad cake, representing the undesirable outcome resulting from processing bad data.

Visual Representation:

- [Stale Flour] + [Rotten Eggs] + [Incorrect Measurements] --> [Baking Process] --> [Flat, Unpleasant, Dry Cake]
(Garbage In) (Garbage Out)

Our barn had to pass a foundation inspection

- Before further construction could proceed
- It makes good business sense
- No IT equivalent



Data Governance Frameworks

- A system of ideas for guiding analyses
- A means of organizing project data
- Priorities for data decision making
- A means of assessing progress
 - Don't put up walls until foundation inspection is passed
 - Put the roof on ASAP
- Make it all dependent upon continued funding



Data Governance from the DMBOK

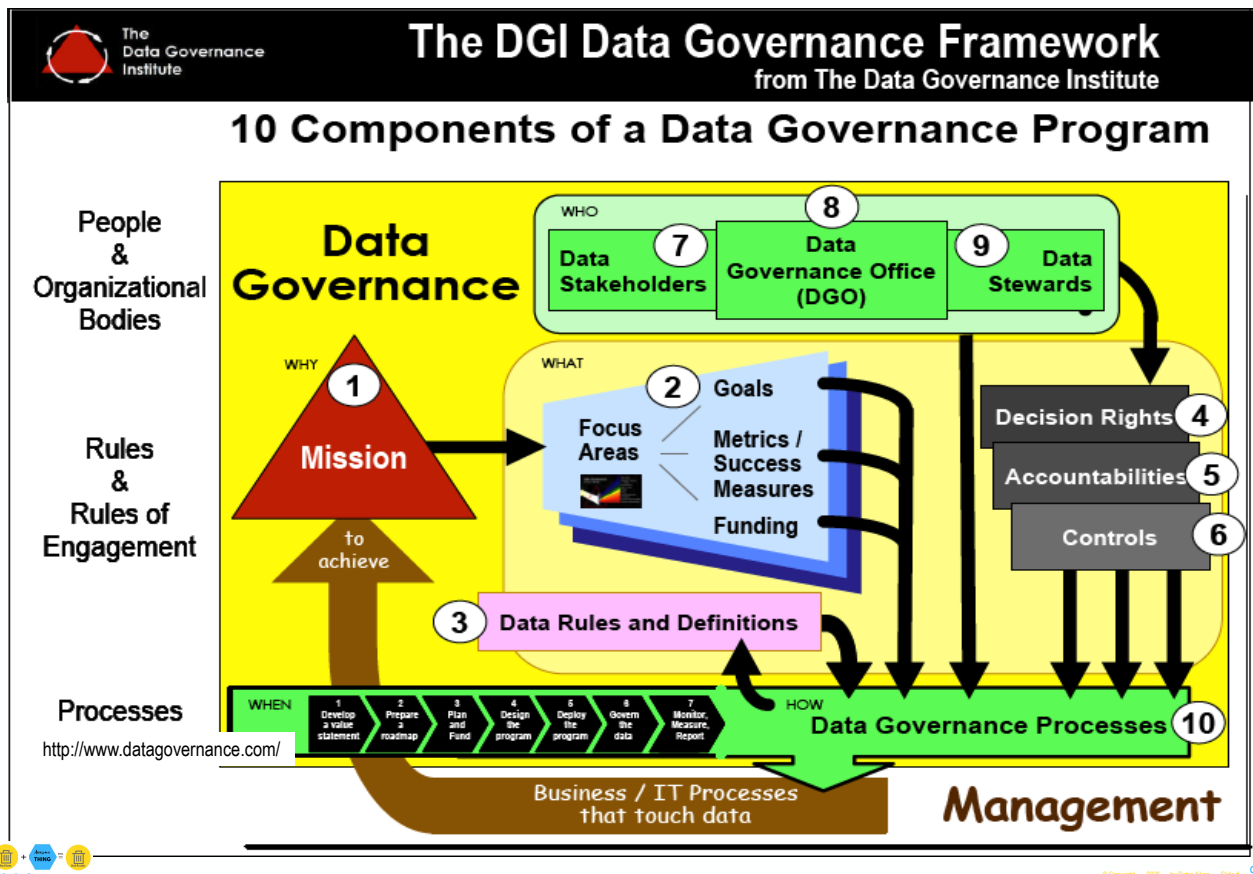
Definition: The exercise of authority and control (planning, monitoring, and enforcement) over the management of data assets.

Goals:

1. To define, approve, and communicate data strategies, policies, standards, architecture, procedures, and metrics.
2. To track and enforce regulatory compliance and conformance to data policies, standards, architecture, and procedures.
3. To sponsor, track, and oversee the delivery of data management projects and services.
4. To manage and resolve data related issues.
5. To understand and promote the value of data assets.

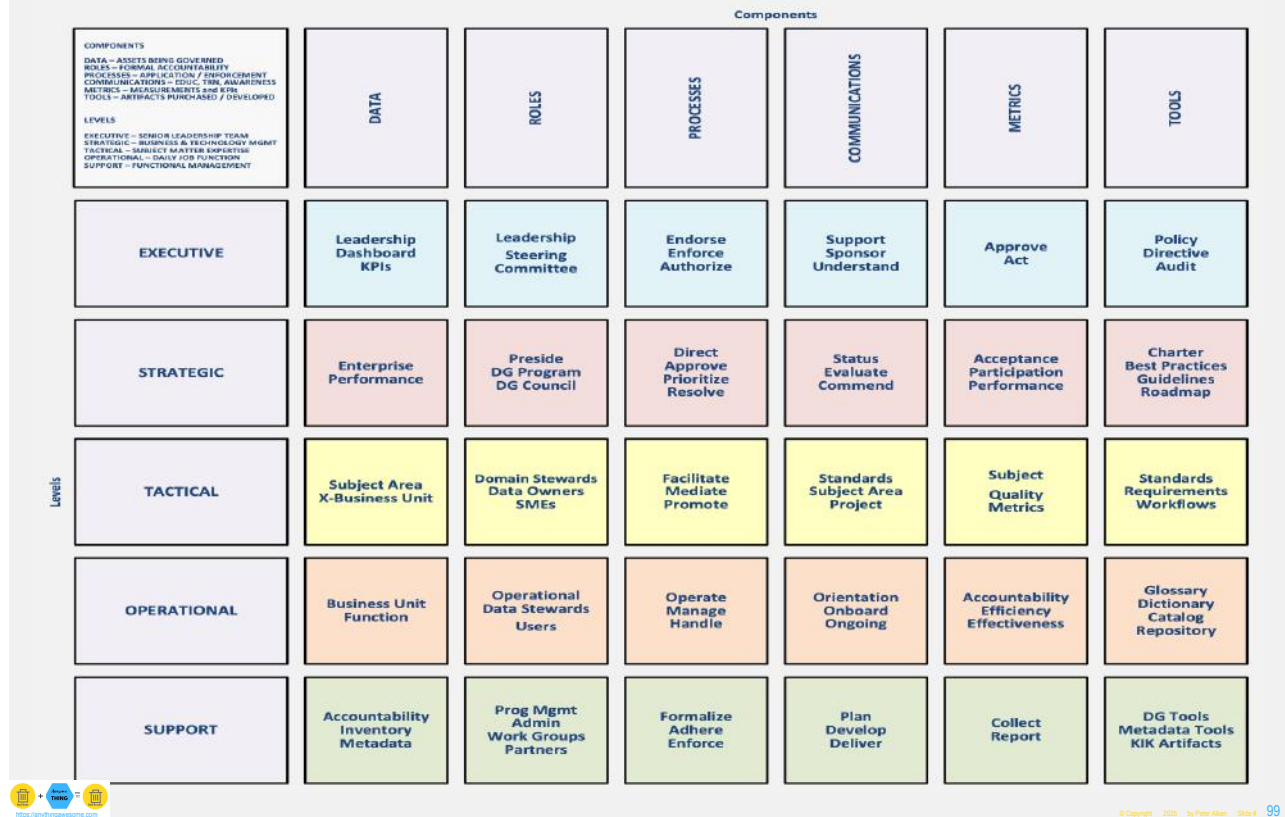


Data Governance Institute



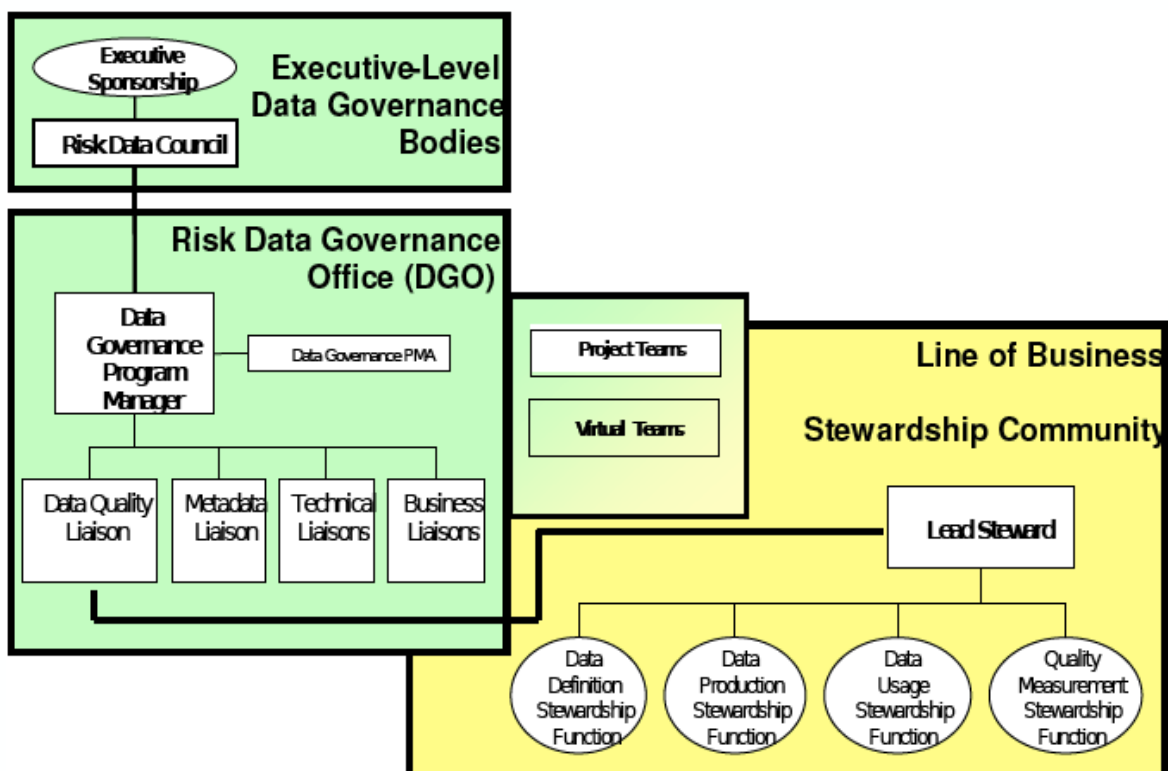
KiK Consulting

NON-INVASIVE DATA GOVERNANCE™ FRAMEWORK

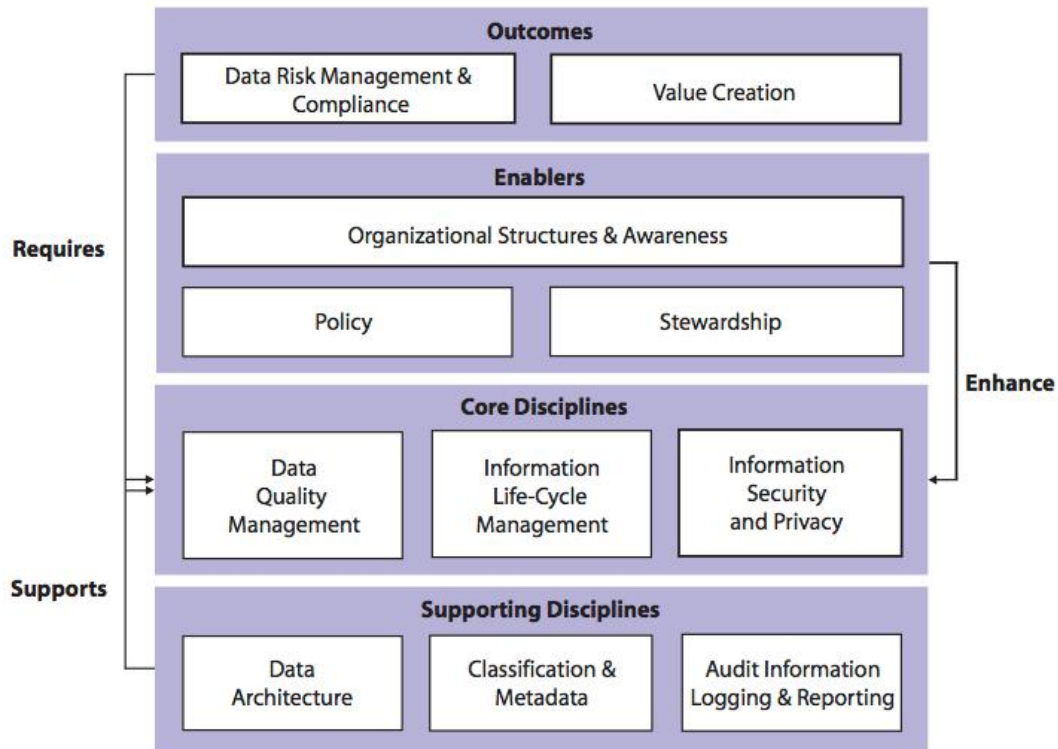


<http://www-01.ibm.com/software/data/system-z/data-governance/workshops.html>

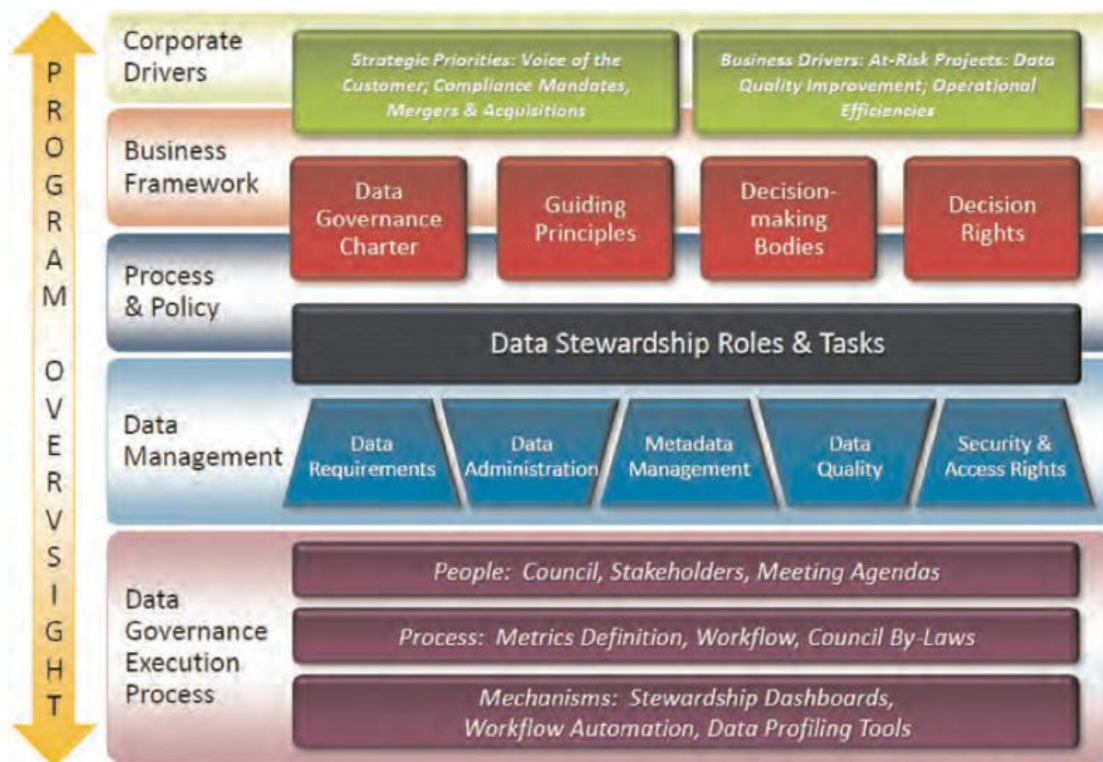
IBM Data Governance Council



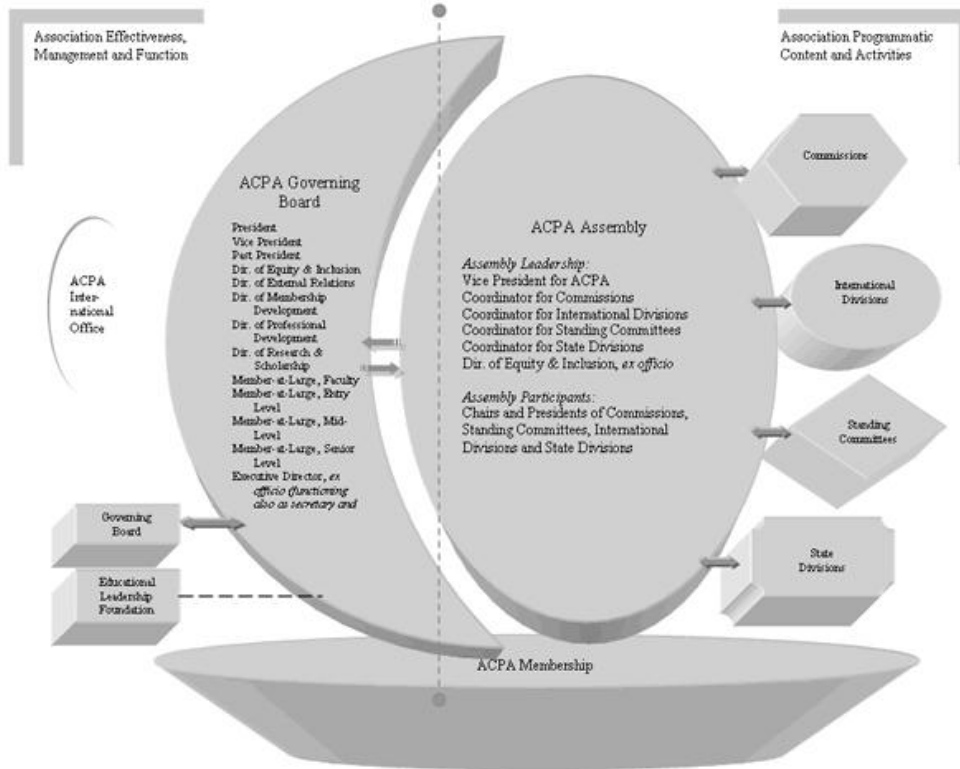
Elements of Effective Data Governance



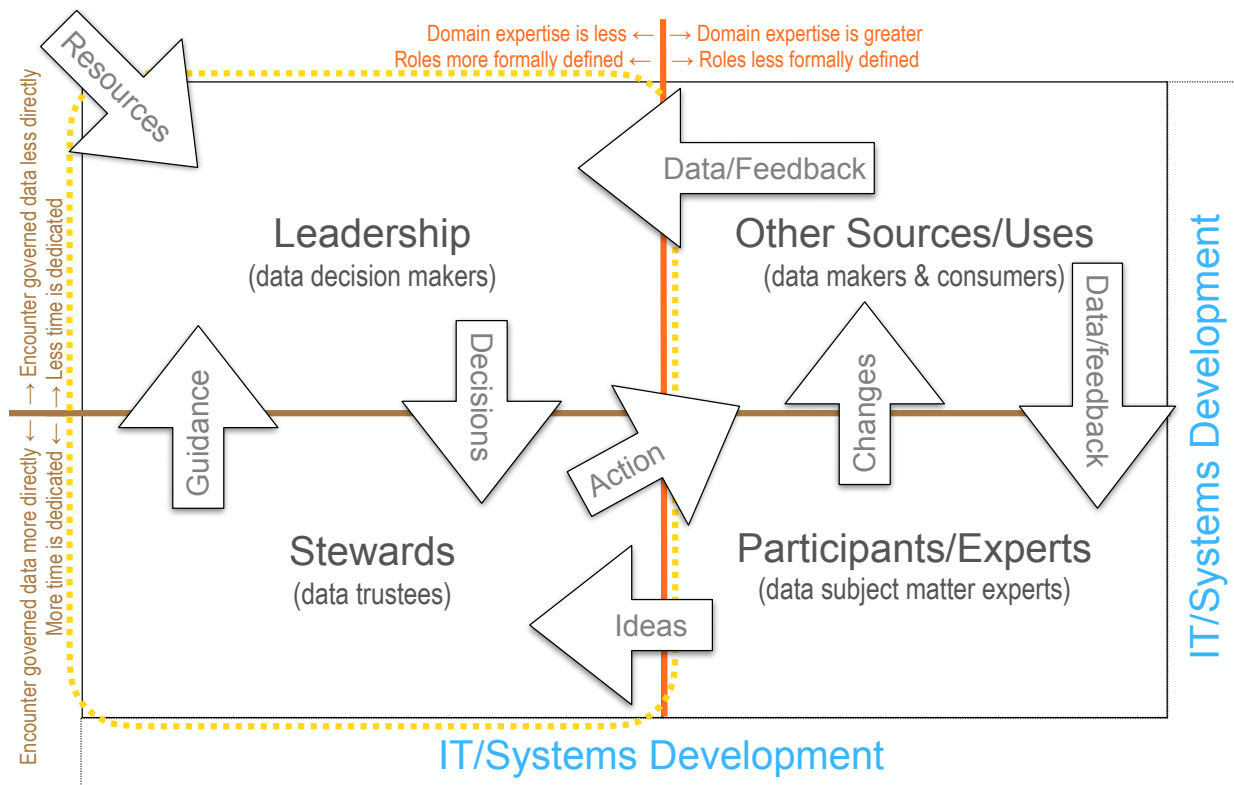
Baseline Consulting (sas.com)



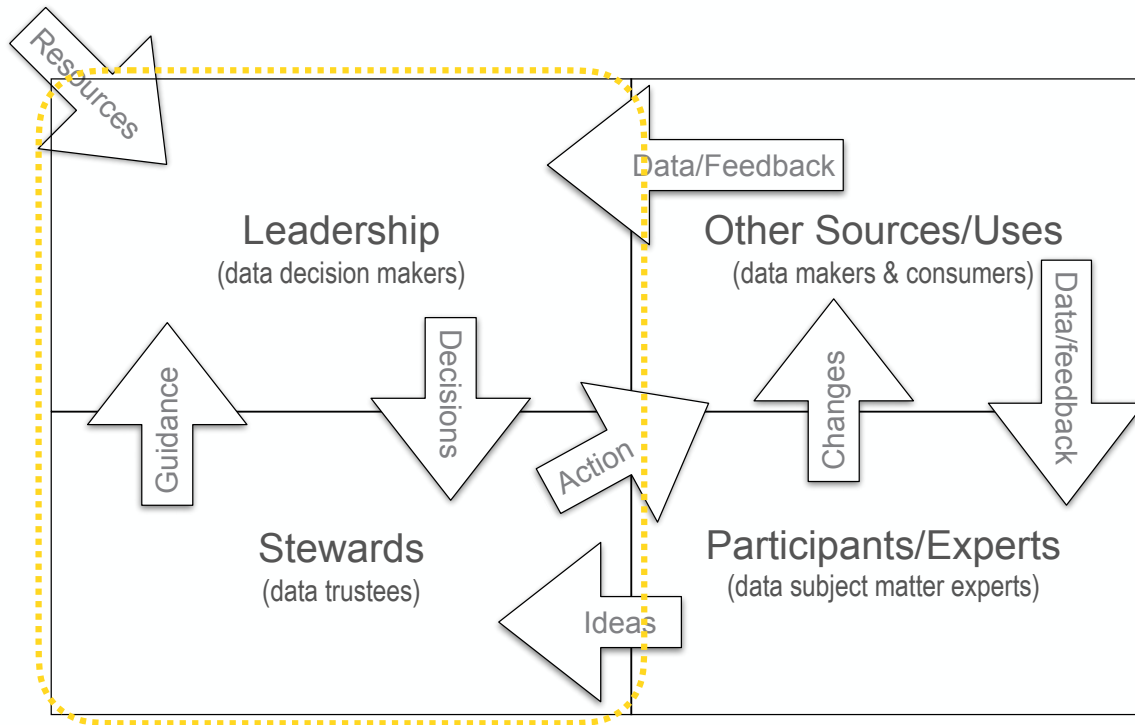
American College Personnel Association



Components comprising the data community



Components comprising the data community



Getting Started with Data Governance



Goals and Principles

- To define, approve, and communicate data strategies, policies, standards, architecture, procedures, and metrics.
- To track and enforce regulatory compliance and conformance to data policies, standards, architecture, and procedures.
- To sponsor, track, and oversee the delivery of data management projects and services.
- To manage and resolve data related issues.
- To understand and promote the value of data assets.

Illustration from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International



<https://creativecommons.org/licenses/by-nc-sa/4.0/>



Primary Deliverables

- Data Policies
- Data Standards
- Resolved Issues
- Data Management Projects and Services
- Quality Data and Information
- Recognized Data Value

Deliverables



Illustration from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Roles and Responsibilities

- Suppliers:
 - Business Executives
 - IT Executives
 - Data Stewards
 - Regulatory Bodies
- Consumers:
 - Data Producers
 - Knowledge Workers
 - Managers and Executives
 - Data Professionals
 - Customers

- Participants:
 - Executive Data Stewards
 - Coordinating Data Stewards
 - Business Data Stewards
 - Data Professionals
 - DM Executive
 - CIO



Illustration from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide # 109

Scorecard: Data Governance Practices/Techniques

- Data Value
- Data Management Cost
- Achievement of Objectives
- # of Decisions Made
- Steward Representation/Coverage
- Data Professional Headcount
- Data Management Process Maturity



from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide # 110

Data Governance Checklist

- ✓ Decision-Making Authority
- ✓ Standard Policies and Procedures
- ✓ Data Inventories
- ✓ Data Content Management
- ✓ Data Records Management
- ✓ Data Quality
- ✓ Data Access
- ✓ Data Security and Risk Management



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2020 by Peter Allen Slide # 111

DG Components

- Practices and Techniques
 - Data Value
 - Data Management Cost
 - Achievement of Objectives
 - # of Decisions Made
 - Steward Representation/Coverage
 - Data Professional Headcount
 - Data Management Maturity
- What do I include in my Data Governance Program?
 - Security and Privacy of Data
 - Quality of Data
 - Life Cycle Management
 - Risk Management
 - Content Valuation
 - Standards (Data Design, Models and Tools)
 - Governance Tool Kits and Case Studies



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Illustration from *The DAMA Guide to the Data Management Body of Knowledge* © 2009 by DAMA International

© Copyright 2020 by Peter Allen Slide # 112

Goals and Principles

- To define, approve, and communicate data strategies, policies, standards, architecture, procedures, and metrics.
- To track and enforce regulatory compliance and conformance to data policies, standards, architecture, and procedures.
- To sponsor, track, and oversee the delivery of data management projects and services.
- To manage and resolve data related issues.
- To understand and promote the value of data assets.

Primary Deliverables

- Data Policies
- Data Standards
- Resolved Issues
- Data Management Projects and Services
- Quality Data and Information
- Recognized Data Value

Roles and Responsibilities

- Suppliers:**
 - Business Executives
 - IT Executives
 - Data Stewards
 - Regulatory Bodies
- Consumers:**
 - Data Professionals
 - Business Executives
 - Data Professionals
 - Customer
- Participants:**
 - Executive Data Stewards
 - Coordinating Data Stewards
 - Business Data Stewards
 - Data Professionals
 - DM Executive
 - Customer

Scorecard: Data Governance Practices/Techniques

- Data Value
- Data Management Cost
- Achievement of Objectives
- Decisions Made
- Steward Representation/Coverage
- Data Professional Headcount
- Data Management Process Maturity

Governance

- Decision-Making Authority
- Standard Policies and Procedures
- Data Inventories
- Data Content Management
- Data Records Management
- Data Quality
- Data Access
- Data Security and Risk Management

What do I include in my Governance Program?

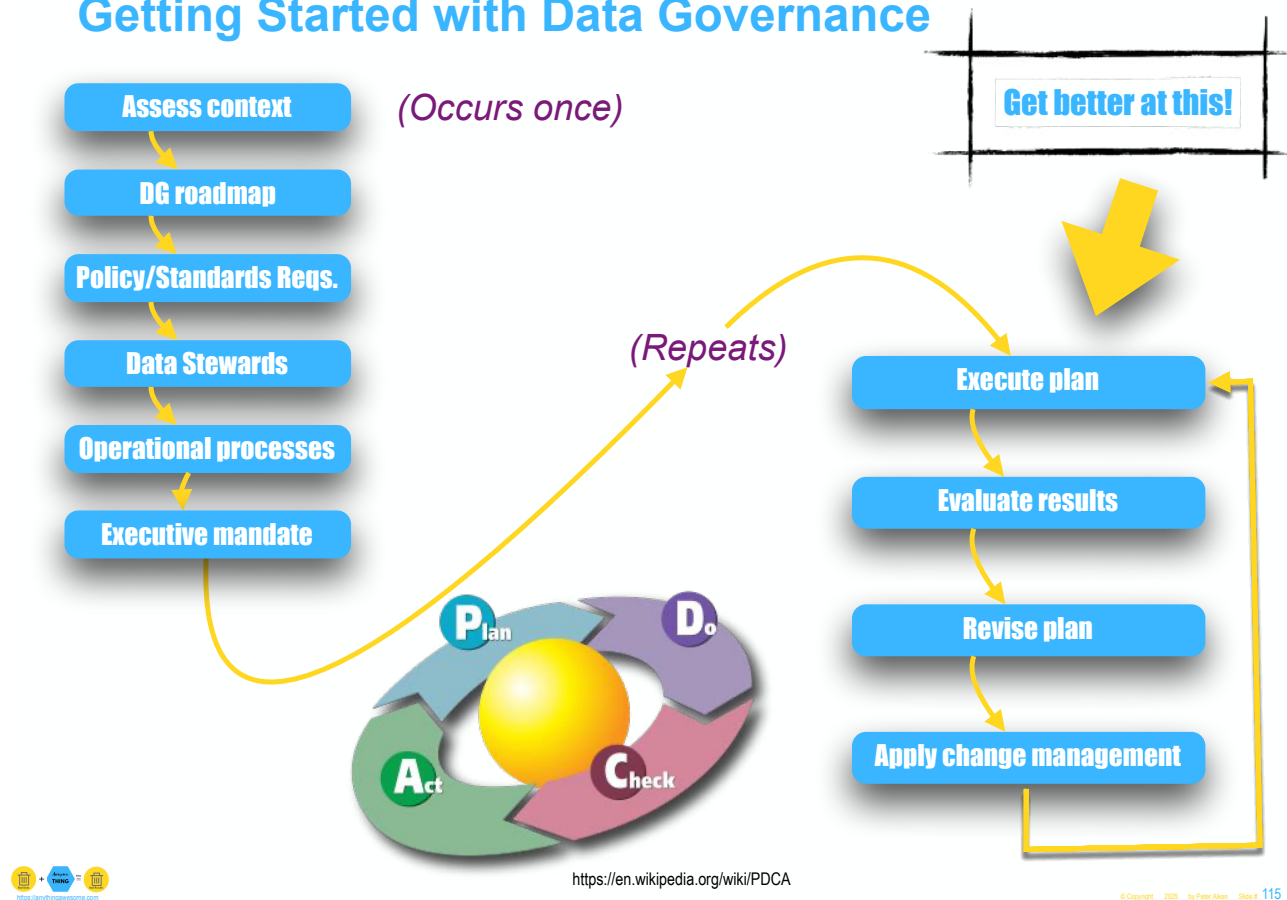
- Security and Privacy of Data
- Quality of Data
- Life Cycle Management
- Risk Management
- Content Valuation
- Standards (Data Design, Models and Tools)
- Governance Tool Kits and Case Studies

https://www.google.com/url?sa=i&url=https%3A%2F%2Fprezi.com%2FEl_1jxok7i-x%2Fevolution-is-not-goal-oriented%2F&psig=AOvVaw3rFrerRanKfaIDIESqab4C&ust=1591700086354000&source=images&cd=vfe&ved=0CAIQJRxqfwTCODBilyH8ukCFQAAAAAAdAAAAABAD

Evolution is not goal oriented

Evolve

Getting Started with Data Governance



Write a Data Governance Charter for a Small Non-profit that Rescues Horses

1. Introduction

- [Non-Profit Name] recognizes that data is a critical asset for achieving our mission of rescuing and rehabilitating horses. Effective data governance is essential to ensure data quality, integrity, security, and accessibility, enabling informed decision-making and efficient operations. This charter establishes the framework for data governance at [Non-Profit Name].

2. Mission and Objectives

- The mission of this Data Governance program is to:
- Ensure data is accurate, reliable, and consistent for effective program management and reporting.
- Protect sensitive data related to donors, volunteers, staff, and the horses in our care.
- Optimize data utilization for fundraising, outreach, and operational efficiency.
- Promote data literacy and a data-driven culture within the organization.
- Comply with all applicable legal and regulatory requirements related to data privacy and security.

3. Scope

- This charter applies to all data created, collected, processed, and stored by [Non-Profit Name], including but not limited to:
- Donor information
- Volunteer information
- Horse records (medical, behavioral, training)
- Financial data
- Website analytics
- Social media data
- Program data (adoptions, rescues, etc.)

4. Guiding Principles

- The Data Governance program will be guided by the following principles:
- Accountability: Clearly defined roles and responsibilities for data management.
- Transparency: Open communication and access to data governance policies and procedures.
- Integrity: Maintaining the accuracy, completeness, and consistency of data.
- Security: Protecting data from unauthorized access, use, or disclosure.
- Accessibility: Ensuring data is readily available to authorized users.
- Compliance: Adhering to all relevant legal and regulatory requirements.

5. Data Governance Roles and Responsibilities

- Data Governance Council: Composed of key stakeholders (e.g., Executive Director, Program Manager, Development Director, Volunteer Coordinator). Responsible for:
- Setting data governance policies and standards.
- Approving major data-related initiatives.
- Monitoring data quality and compliance.
- Resolving data-related disputes.
- Data Steward(s): Designated individuals responsible for managing specific data domains (e.g., Donor Data Steward, Horse Records Steward). Responsibilities include:

- Ensuring data quality within their domain.
- Implementing data governance policies.
- Training staff on data management procedures.
- Data Custodian(s): Individuals responsible for the technical management of data and systems (e.g., IT staff, database administrators). Responsibilities include:
- Implementing data security measures.
- Maintaining data infrastructure.
- Ensuring data backup and recovery.
- Data Users: All staff and volunteers who interact with data. Responsible for:
- Adhering to data governance policies and procedures.
- Reporting data quality issues.
- Using data responsibly and ethically.

6. Data Governance Processes

- Data Quality Management: Implementing processes for data validation, cleansing, and enrichment.
- Data Security and Privacy: Establishing security measures to protect data from unauthorized access and use.
- Data Access and Sharing: Defining clear guidelines for data access and sharing, both internally and externally.
- Data Retention and Disposal: Developing policies for data retention and secure disposal of data.
- Data Training and Awareness: Providing training to staff and volunteers on data governance policies and best practices.

7. Data Governance Policies and Standards

- The Data Governance program will develop and maintain policies and standards related to:
- Data quality
- Data security
- Data privacy
- Data access
- Data retention
- Data dictionary and metadata management

8. Review and Updates

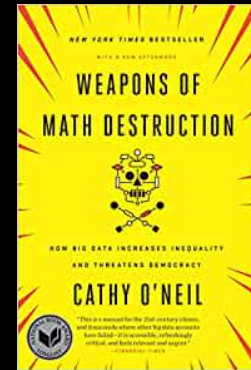
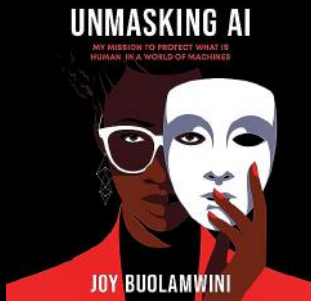
- This Data Governance Charter will be reviewed and updated annually or as needed to reflect changes in the organization's needs or regulatory environment.

9. Contact Information

- For questions or concerns regarding this Data Governance Charter, please contact [Designated Contact Person/Department].
- This charter serves as a foundational document for data governance at [Non-Profit Name]. Its successful implementation will require ongoing commitment and collaboration from all members of the organization.

"Should I-thinking" must replace "Can I-thinking"

(First level capable of being bound by defined behaviors)



Data Ethics



© Copyright 2020 by Peter Allen Slide # 117

Program Overview

- Data's Confounding Characteristics
 - Uneven understanding
 - Has lead fractured views of data and to
 - Increasing organizational data debt
- 1. Keeping DG practically focused on strategy
 - This is a young profession and must
 - Directly support organizational strategy by
 - Improving data and its use in the short and long term
- 2. DG must exist at the same level as HR
 - In order to achieve effectiveness,
 - DG is central to DM (and central to digitization efforts)
 - Must be de-coupled from IT strategy
- 3. Gradually add ingredients (practicing and getting better)
 - Digital and data are dependent on high speed automation/data processing
 - Employ a DG Frameworks to refine focus
 - Plan to evolve (PDCA)
- 4. Data governance in action: Storytelling
 - Take Aways/References/Q&A

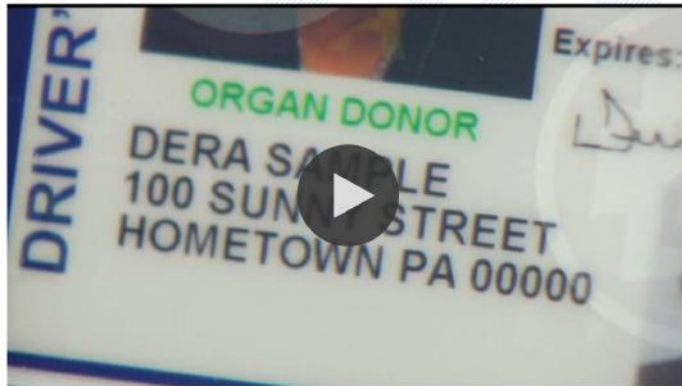


© Copyright 2020 by Peter Allen Slide # 118

PennDOT makes millions selling your personal data, is it safe?

Two Headlines

HARRISBURG, Pa — For years, the Pennsylvania Department of Transportation has been selling drivers information. That information is bought by seven differ...



Author: Jackie De Tore (FOX43)
Published: 10:21 PM EDT May 4, 2016
Updated: 10:27 PM EDT May 4, 2016



HARRISBURG, Pa -- For years, the Pennsylvania Department of Transportation has been selling drivers information.

That information is bought by seven different companies.

"Under the law, PennDOT is allowed to charge \$9 for each record that is accessed. This is a

...ania pays for critical transportation services," says Rich Kirkpatrick, a PennDOT



<https://www.fox43.com/article/news/local/contests/penn-dot-makes-millions-selling-your-personal-data-is-it-safe/521-be4956dc-73a9-4595-ad19-de842f6e769c>

© Copyright 2020 by Peter Allen Size 119

1. PennDOT gets caught selling \$40,000,000 worth of your data to insurance companies
2. PennDOT saves taxpayers \$40M selling data legally derived from your data



Office of the
Comptroller of the Currency

ABOUT

NEWS & EVENTS

PUBLICATIONS & RESOURCES

TOPICS

Home > News & Events > Newsroom

News Release 2020-132 | October 7, 2020

OCC Assesses \$400 Million Civil Money Penalty Against Citibank

SHARE THIS PAGE:

WASHINGTON—The Office of the Comptroller of the Currency (OCC) today assessed a \$400 million civil money penalty against Citibank, N.A., of Sioux Falls, South Dakota, related to deficiencies in enterprise-wide risk management, compliance risk management, data governance, and internal controls.

The OCC took these actions based on the bank's unsafe or unsound banking practices for its long-standing failure to establish effective risk management and data governance programs and internal controls. This failure also resulted in a violation of 12 CFR Part 30, Appendix D, "OCC Guidelines Establishing Heightened Standards for Certain Large Insured National Banks, Insured Federal Savings Associations, and Insured Federal Branches."

The agency also issued a cease and desist order requiring the bank to take broad and comprehensive corrective actions to improve risk management, data governance, and internal controls. The order requires the bank to seek the OCC's non-objection before making significant new acquisitions and reserves the OCC's authority to implement additional business restrictions or require changes in senior management and the bank's board should the bank not make timely, sufficient progress in complying with the order.

The Federal Reserve Board took a separate but related action against Citigroup, the bank's holding company.

The OCC penalty will be paid to the U.S. Treasury.

Media Contact

Bryan Hubbard
(202) 649-6870



<https://www.occ.gov/news-issuances/news-releases/2020/nr-occ-2020-132.html>

© Copyright 2020 by Peter Allen Size 120

'Jack Dorsey's First Tweet' NFT Went on Sale for \$48M. It Ended With a Top Bid of Just \$280

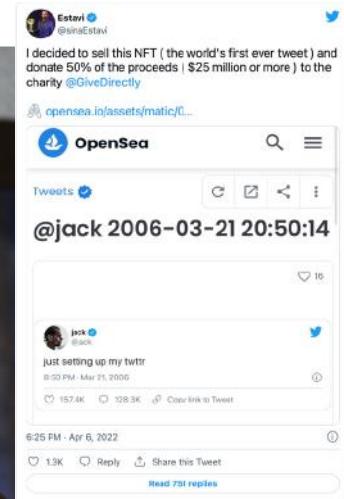


Crypto entrepreneur Sina Estavi bought Twitter founder Jack Dorsey's first-ever tweet as an NFT for \$2.9 million last year. He listed the NFT for sale again at \$48 million last week.

By Sandali Handagama · Apr 13, 2022 at 2:48 p.m. EDT · Updated Apr 16, 2022 at 10:12 a.m. EDT



Twitter founder Jack Dorsey's tweet NFT was briefly on sale again. (CoinDesk archives)



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2022 by Peter Allen Slide 121

Use Their Language ...



- *Getting access to data around here is like that Catherine Zeta Jones scene where she is having to get thru all those lasers ...*



<https://creativecommons.org/licenses/by-nc-sa/4.0/>

© Copyright 2022 by Peter Allen Slide 122

Barclays Excel Spreadsheet Horror

- Barclays preparing to buy Lehman's Brothers assets.
- 179 dodgy Lehman's contracts were almost accidentally purchased by Barclays because of an Excel spreadsheet reformatting error
- A first-year associate reformatted an Excel contracts spreadsheet
 - Predictably, this work was done long after normal business hours, just after 11:30 p.m...
- The Lehman/Barclays sale closed on September 22nd, 2008
- the 179 contracts were marked as "hidden" in Excel, and those entries became "un-hidden" when when globally reformatting the document ...
- ... and the sale closed ...

<https://www.businessinsider.com/2008/10/barclays-excel-error-results-in-lehman-chaos>



<https://daringgawdome.com>

Hidden spreadsheet rows hit Barclays with toxic Lehman contracts

by AccountingWEB



Lehman Excel snafu could cost Barclays dear

Hey, whoa, we never meant to buy that bit

Excel error leaves Barclays with more Lehman assets than it bargained for

Law firm says spreadsheet reformatting error added 179 contracts to bankruptcy buyout deal

© Copyright 2020 by Peter Allen Slide # 123

SCIENCE / US & WORLD / TECH

Excel spreadsheet error blamed for UK's 16,000 missing coronavirus cases

The case went missing after the spreadsheet hit its filesize limit

By James Vincent | Oct 5, 2020, 9:41am EDT

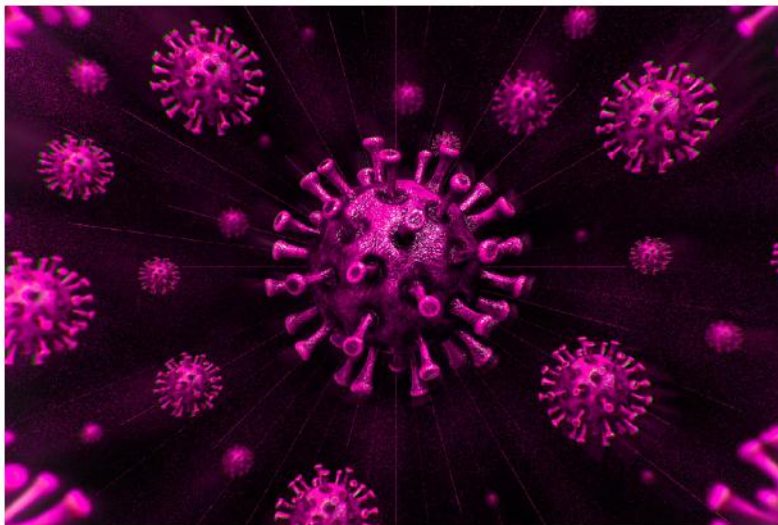


Illustration by Alex Castro / The Verge

The UK failed to add nearly 16,000 confirmed cases of coronavirus to its national track and trace system due to an Excel error. A number of reports, including from The Guardian, Sky News, and The Daily Mail, say the mistake was caused when an Excel spreadsheet used to track confirmed cases of the virus reached its maximum file size and failed to update.

"Failure to upload these cases to the national database meant anyone who came into contact with these individuals was not informed. It's an error that may have helped spread the virus further through the country as individuals exposed to the virus continued to act as normal."



<https://daringgawdome.com>

© Copyright 2020 by Peter Allen Slide # 124

Program overview

- Data's Confounding Characteristics
 - Uneven understanding
 - Has lead fractured views of data and to
 - Increasing organizational data debt
- 1. Keeping DG practically focused on strategy
 - This is a young profession and must
 - Directly support organizational strategy by
 - Improving data and its use in the short and long term
- 2. DG must exist at the same level as HR
 - In order to achieve effectiveness,
 - DG is central to DM (and central to digitization efforts)
 - Must be de-coupled from IT strategy
- 3. Gradually add ingredients (practicing and getting better)
 - Digital and data are dependent on high speed automation/data processing
 - Employ a DG Frameworks to refine focus
 - Plan to evolve (PDCA)
- 4. Data governance in action: Storytelling
 - Take Aways/References/Q&A



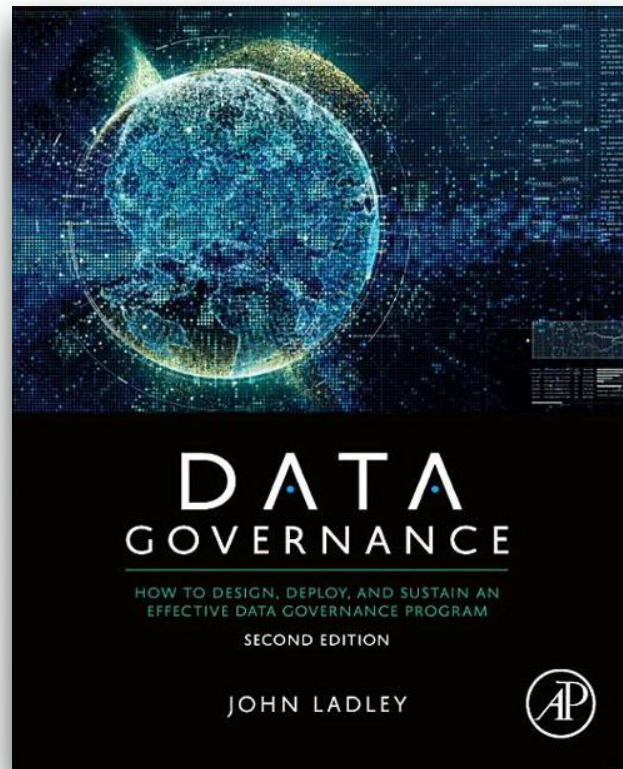
Take Aways

- Need for DG is increasing
 - Increase in data volume
 - Lack of rigorous practice improvement
- DG is a new discipline
 - Must conform to constraints
 - No one best way
- DG must be driven by 4 key elements
 1. Keep DG practically focused on strategy
 2. Implement DG (and data) as a program not a project
 3. Gradually add ingredients
 4. Learn the value of stories/storytelling
- The goal is to improve DG effectiveness and efficiencies (and the data itself) over time
- The more data literate the organization, the easier the transformation



By the book

- Data Governance: How to Design, Deploy, and Sustain an Effective Data Governance Program
- John Ladley
- Amazon Best Sellers Rank: #641,937 in Books (See Top 100 in Books)
 - #242 in Management Information Systems
 - #209 in Library Management
 - #380 in Database Storage & Design



References

Websites

The Data Administration Newsletter (TDAN)—<http://www.TDAN.com>

DM Review Magazine—www.dmreview.com. Note: www.dmreview.com is now www.information-management.com.

EIM Insight, published by The Enterprise Information Management Institute—<http://eiminstitute.org>

SearchDataManagement.com white paper library—<http://go.techtarget.com/r/3762877/5626178>

Compliance Book

Thomas, Gwen. Alpha Males and Data Disasters: The Case for Data Governance. Brass Cannon Press, 2006. ISBN-10: 0-978-6579-0-X. 221 pages.

Bloem, Jaap, Menno van Doorn, and Piyush Mittal. Making IT Governance Work in a Sarbanes-Oxley World. John Wiley & Sons, 2005. ISBN 0-471-74359-3. 304 pages.



IT Governance Books

Benson, Robert J., Tom Bugnitz, and Bill Walton. From Business Strategy to IT Action: Right Decisions for a Better Bottom Line. John Wiley & Sons, 2004. ISBN 0-471-49191-8. 309 pages.

IT Governance Institute. Control Objectives for Information and related Technology (CobiT®). www.isaca.org/cobit

Lutchen, Mark. Managing IT as a Business: A Survival Guide for CEOs. John Wiley & Sons, 2003. ISBN 0-471-47104-6. 256 pages.

Maizlish, Bryan and Robert Handler. IT Portfolio Management Step-By-Step: Unlocking the Business Value of Technology. John Wiley & Sons, 2005. ISBN 0-471-64984-8. 400 pages.

Van Grembergen, Wim and Steven Dehaes. Enterprise Governance of Information Technology: Achieving Strategic Alignment and Value. Springer, 2009. ISBN 0-387-84881-5, 360 pages.

Van Grembergen, Wim and Steven Dehaes. Implementing Information Technology Governance: Models, Practices and Cases. IGI Publishing, 2007. ISBN 1-599-04924-3, 255 pages.

Van Grembergen, Wim and Steven Dehaes. Strategies for Information Technology Governance. IGI Publishing, 2003. ISBN 1-591-40284-0. 406 pages.

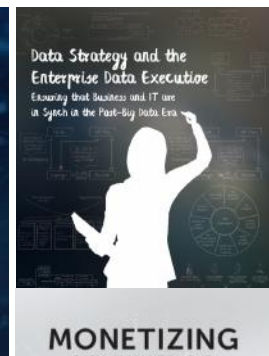
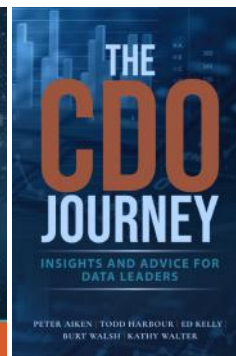
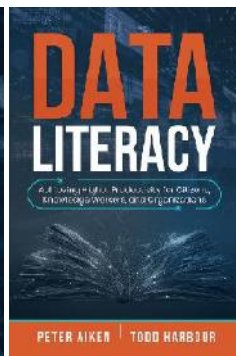
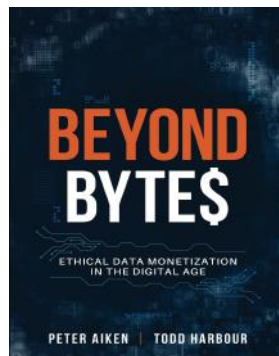
Weill, Peter and Jeanne Ross. IT Governance: How Top Performers Manage IT Decision Rights for Superior Results. Harvard Business School Press, 2004. ISBN 1-291-39253-5. 288 pages.



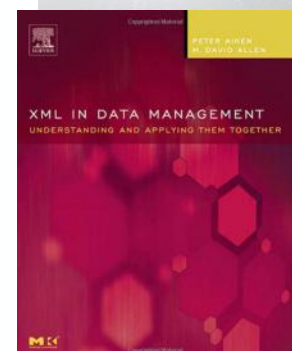
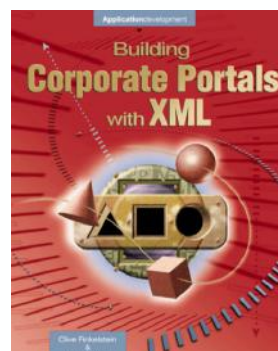
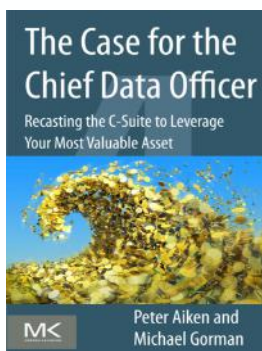
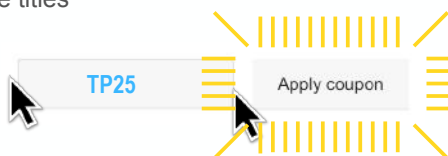
<https://anythingsawesome.com>

© Copyright 2020 by Peter Aiken Slide 129

Event Pricing on Peter's Books



- 25% off directly from the publisher on some titles
- My 'Book Store' @ <https://anythingsawesome.com/books>
- Enter the code "TP25" at the Technics bookstore checkout where it says to "Apply Coupon"



Independent Verification & Validation

Critical Design Review?

Mentoring?

Collaboration?

Executive Data
Literacy Training?



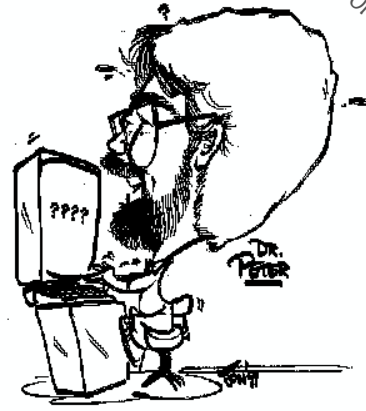
+



=



Peter.Aiken@AnythingAwesome.com +1.804.382.5957



Reverse Engineering Expertise?

Hiring Assistance?

Thank You!

Use your data more strategically?

Tool/automation evaluation?

Book a call with Peter to discuss anything - <https://anythingawesome.com/OfficeHours.html>





CDO Roadmap 2025:

Is Your Data AI-Ready?

An AI-Preparedness Report

cdo CHIEF DATA OFFICER
MAGAZINE

EY

Contents

Executive Summary	3
Insights	4
Methodology	13
Strategic Insights From EY	14

Executive Summary

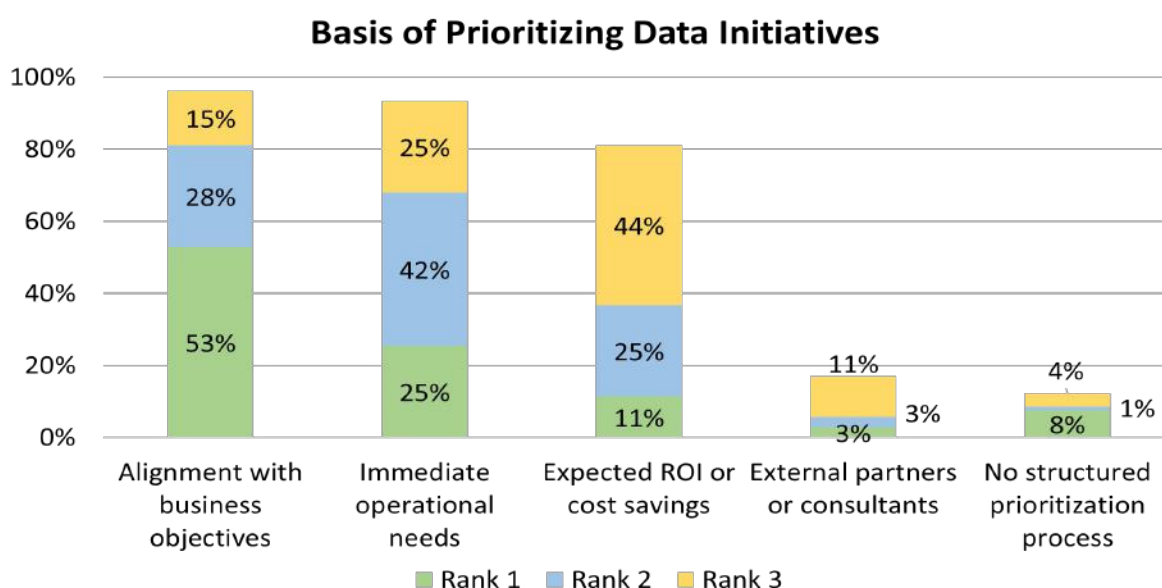
As organizations strive to unlock business value with AI, data readiness has become a critical success factor. This report examines the state of AI-driven data strategies, assessing key areas such as vision, integration challenges, ethical considerations, talent gaps, and infrastructure preparedness.

By identifying obstacles and best practices, it provides a strategic roadmap to enhance AI readiness and maximize data investment returns. The following executive summary highlights key insights from the study:

- **Prioritizing data initiatives:** 53% of organizations rank alignment with business objectives as the top factor in prioritizing data initiatives.
- **Ensuring data readiness for AI:** 61% of organizations are establishing clear data governance processes and policies to ensure data readiness for AI.
- **Obstacles to AI potential:** Poor quality data, the shortage of AI talent, and limited scalability of current technology are the top three obstacles to realizing AI's potential.
- **Challenges integrating data for AI:** 39% of participating data leaders consider data silos across departments as the greatest challenge in integrating data for AI initiatives.
- **Obstacles in using GenAI in business processes:** 37% of respondents rank data security and privacy concerns as the top obstacle to using GenAI in existing processes.
- **Bridging talent gap:** 60% of organizations are upskilling employees through training programs to bridge the data and AI talent gap.
- **Approach to AI ethical risks:** 44% of organizations have a dedicated ethics committee or framework in place to address ethical concerns of AI deployment.
- **Measuring ROI:** 34% of organizations place the highest importance on efficiency improvements and cost savings as to measure ROI on data investments, while 25% place it on revenue growth.

Insights

1. Business objectives and immediate needs drive data initiatives



Alignment with business objectives, immediate operational requirements, and anticipated ROI/cost savings are the three primary factors influencing the prioritization of data initiatives.

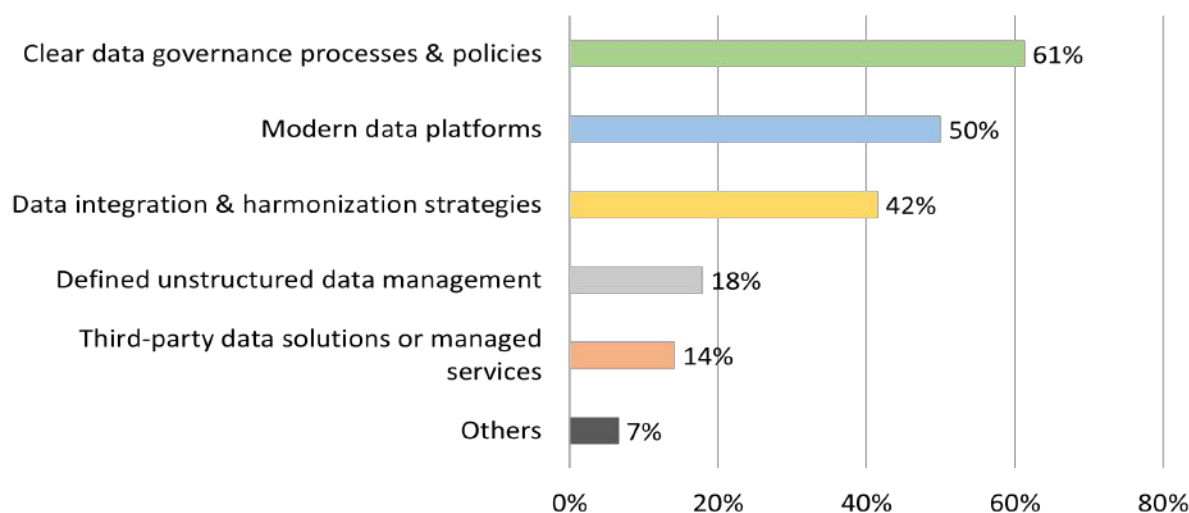
96% of data leaders rank aligning data initiatives with business objectives among their top three factors. Similarly, **92%** of respondents place immediate operational

needs and **80%** consider expected ROI or cost savings within their top three priorities.

Data initiatives are prioritized by overarching business objectives, while also being driven by immediate operational needs and financial gains.

2. Organizations are taking a holistic approach to data readiness for AI

Ensuring Data Readiness for AI



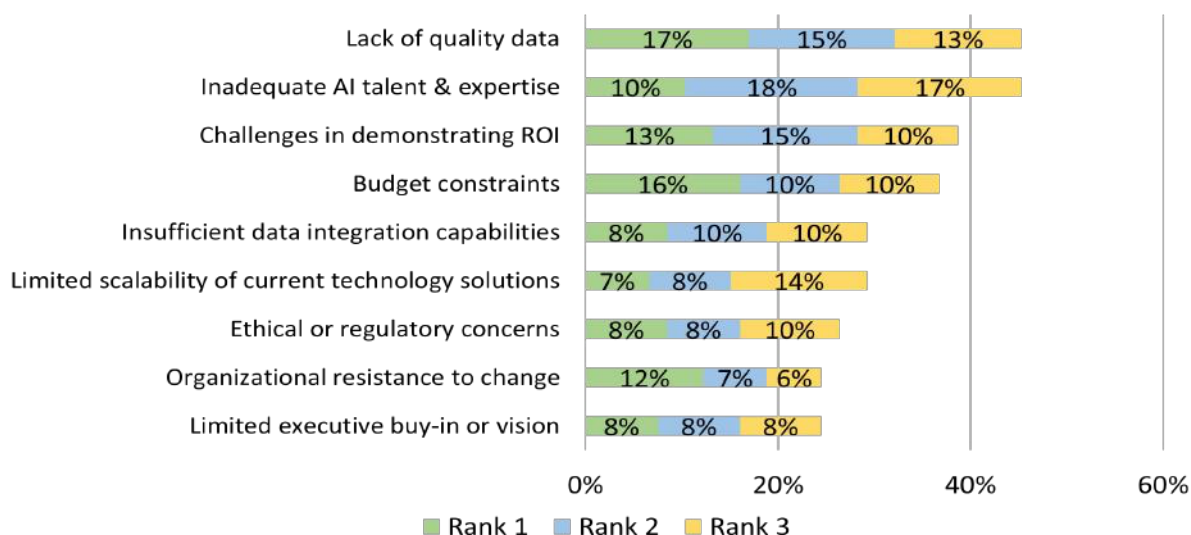
Organizations are actively pursuing multiple initiatives to ensure data readiness for AI. Among their top three approaches, **61%** of respondents are implementing data governance processes and policies, **50%** are modernizing their data platforms, and **42%** are emphasizing data integration and harmonization strategies.

Data leaders also rely on data quality tools and managed services to enhance data preparation efforts.

Data leaders are focusing on defining data governance and policies to create a framework for the usage, integration, and management of data in preparation for AI.

3. Data quality issues and inadequate talent limit AI's potential

Obstacles in Realizing AI Potential



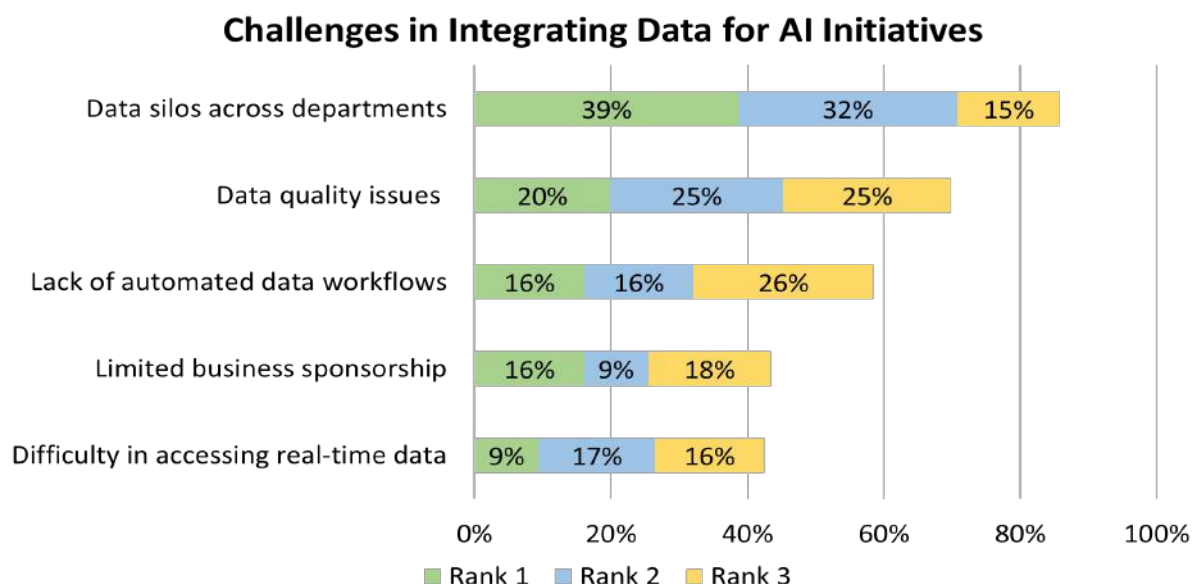
The primary obstacles to realizing AI potential in organizations are the lack of quality data and insufficient AI expertise. **45%** of data executives rank these issues among the top three hindrances.

Additional challenges impeding AI progress encompass difficulties in demonstrating ROI, budget constraints, insufficient data integration capabilities, limited scalability of current technological solutions, ethical or regulatory

concerns, organizational resistance to change, and limited executive buy-in or vision.

Data leaders are facing significant challenges in unlocking AI's full potential due to obstacles such as data quality issues, shortage of skilled talent, and the limited scalability of existing technology.

4. Data silos hinder integration for AI



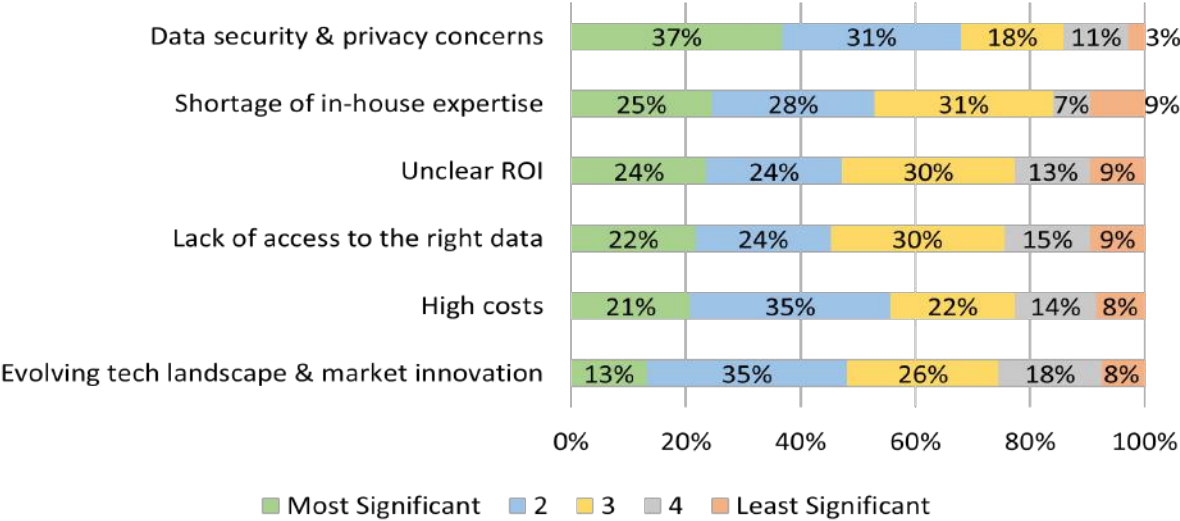
86% of respondents see data silos across departments as the top challenge in integrating data for AI initiatives. **39%** among them identify it as the primary challenge.

Other significant obstacles include data quality issues, lack of automated data workflows, limited business sponsorship, and difficulty in accessing real-time data.

Data silos across departments and the lack of automated data workflows obstruct data integration for AI initiatives.

5. Data security & privacy remains the top obstacle to GenAI integration

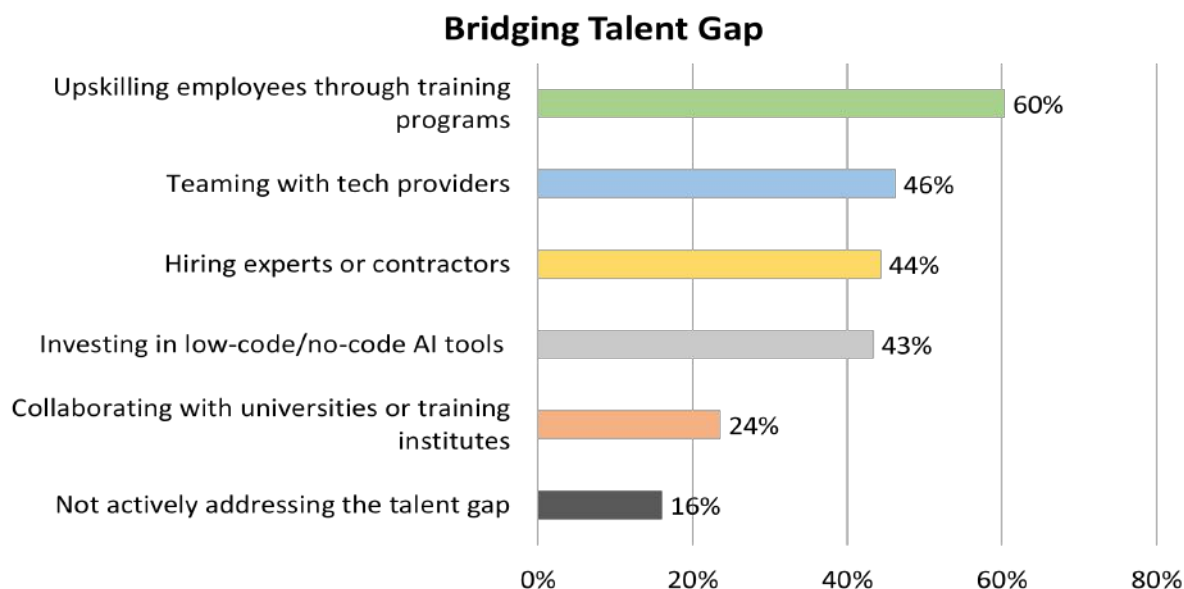
Obstacles Integrating GenAI in Existing Processes



The ranking of obstacles is evenly distributed for integrating GenAI into existing business processes. **37%** of survey participants cite data security and privacy concerns as the main challenge, while **25%** identify a shortage of in-house expertise as a barrier.

Integrating GenAI into organizations' existing business processes presents several significant challenges, including data security and privacy concerns, along with a shortage of in-house expertise.

6. Several measures are bridging the data and AI talent gap



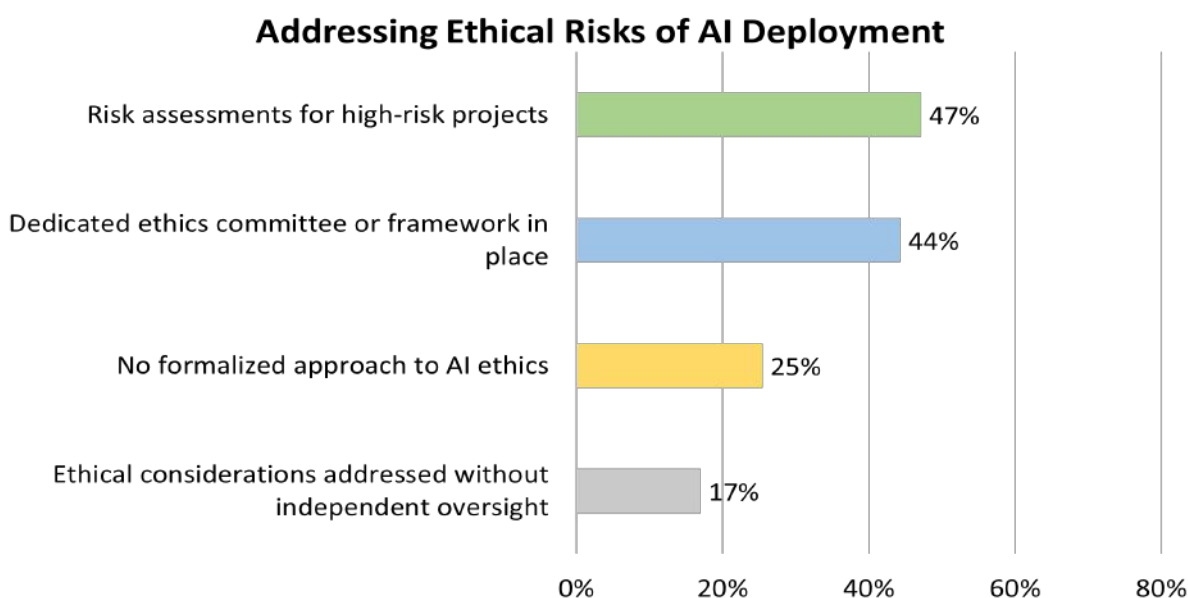
As highlighted earlier, almost half (**45%**) of the organizations identify inadequate AI talent and expertise as the second greatest obstacle in realizing AI potential.

To address this, **60%** of organizations are preferring to fundamentally change the AI talent scenario by running upskilling programs, **46%** are teaming up with tech providers, and **44%** are hiring experts or contractors.

Investing in low-code/no-code AI tools is one strategy **43%** of organizations are using to bridge the talent gap.

Organizations acknowledge the talent gap and are upskilling their teams in AI skills to address it fundamentally. They are also leveraging interim strategies such as engaging experts.

7. Ethical risk mitigation is central to AI deployment

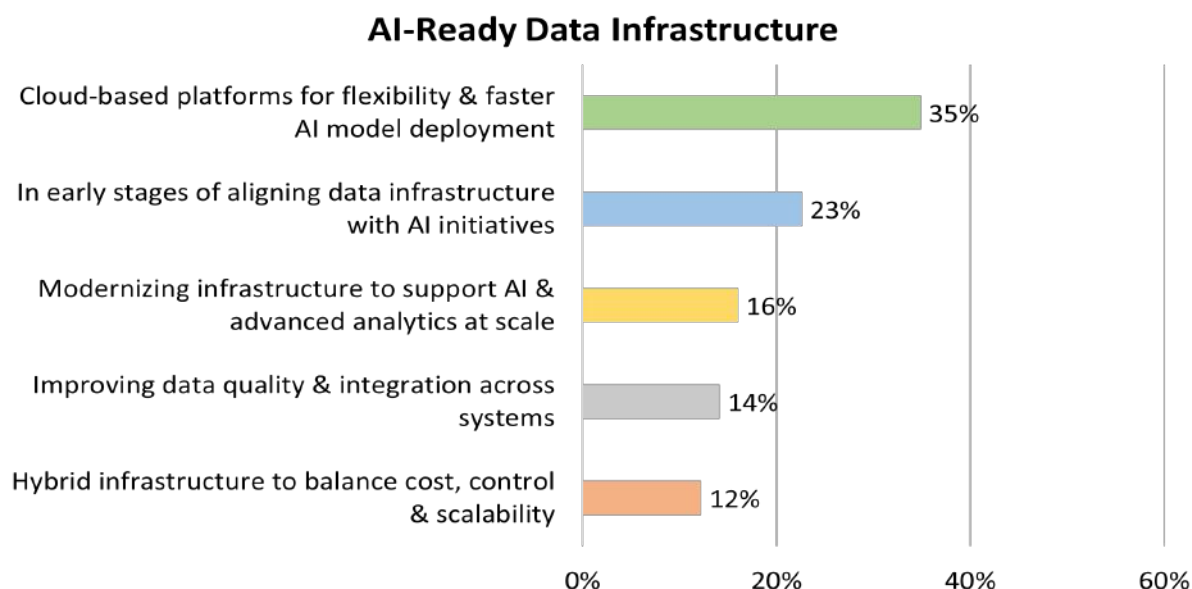


47% of data leaders do a risk assessment of high-risk projects as a measure to address ethical risks in AI deployment. At the same time, **44%** organizations already have a dedicated ethics committee or framework in place.

In contrast, a quarter (**25%**) lack a formalized approach to AI ethics, and **17%** address ethical considerations without independent oversight.

Most organizations have developed formal plans to address the ethical risks of AI deployment, yet many still lack a structured approach.

8. Cloud-based platforms lead the way in AI-ready data infrastructure



35% of organizations say their main investment in AI-ready data infrastructure is cloud-based platforms, taking advantage of their flexibility and quicker AI model deployment.

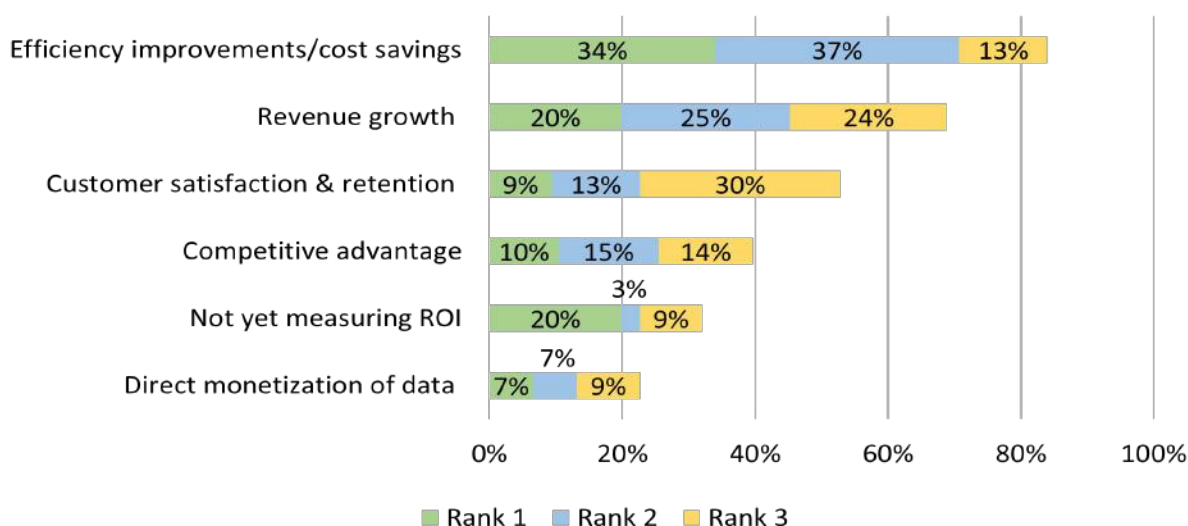
23% are still in the early stages of aligning their infrastructure with AI initiatives.

Meanwhile, **16%** are modernizing their systems to support AI and advanced analytics at scale, and **12%** are opting for hybrid infrastructure to balance cost, control, and scalability.

Cloud-based platforms emerge as the main preference for flexibility and faster AI model deployment.

9. Data investment ROI measured through efficiency/savings, growth, and customer satisfaction

Showing ROI from Data Investments



84% of organizations primarily measure ROI by tracking efficiency improvements and cost savings. Revenue growth is the second most important metric, according to **69%** of data executives.

In addition to these core financial metrics, customer satisfaction and retention are also considered key indicators of the impact of data investments.

Organizations are assessing their ROI on data investments through both quantifiable and qualitative measures.

Methodology

Research objective

The research aimed to provide an in-depth analysis of organizations' data readiness for AI deployment. It focused on evaluating key aspects such as:

- Prioritizing data initiatives
- Obstacles in realizing AI potential
- Measures for ensuring data readiness
- Ethical risks of AI
- Data integration challenges
- Integrating GenAI
- Bridging the talent gap
- Data infrastructure for AI
- ROI of data investments

Survey design and administration

A questionnaire was used to gather insights from data executives on the key aspects outlined in the research objective, aiming to assess data readiness for AI and its potential to drive business value.

Data leaders were invited to participate through multiple CDO Magazine outreach channels. In total, 100+ respondents completed the survey in December 2024.

Strategic Insights From EY

Data plays a critical role in propelling success in today's AI-driven marketplace. This report offers critical insights to help Chief Data Officers (CDOs) harness the power of data to transform their organizations. CDOs who can effectively leverage the full value of their data are poised to align initiatives with core business objectives and drive meaningful outcomes that elevate organizational performance.

Here we highlight five key insights from the research findings that offer strategic guidance for CDOs seeking to unleash the power of data:

1. **Does data have a soul?** Leaders have pointed out that the lack of talent inhibits their ability to successfully deploy AI. In an era when talk about agentic AI is all the rage, this demonstrates that humans are still at the center. Accordingly, 60% of organizations are seeking to address the talent gap by investing in upskilling programs, while 44% have created dedicated ethics committees to manage the ethical risks associated with AI deployment.
2. **Challenges in AI integration:** Nearly four out of ten organizations (37%) report that security and privacy concerns remain significant barriers to adopting Generative AI (GenAI). To overcome this gap, 60% say they are exploring partnerships with tech providers to enhance AI model deployment and address security concerns.
3. **Align data initiatives with business objectives:** More than half of data leaders (53%) say that aligning data initiatives with overarching business goals and operational needs is a critical step for organizations seeking to improve responsiveness to market demands and foster a data-driven culture.
4. **Data governance and quality issues:** While 61% of organizations are establishing robust data governance processes, challenges such as poor data quality (reported by 31%) and data silos (reported by 39%) continue to hinder the effective integration of data across departments.
5. **ROI is not the only key driver for data initiatives:** Only 22% of respondents rank ROI as one of the top requirements for an AI integration. Instead, they are more concerned with focusing on business and operational goals, showcasing a shift in the CDO mindset from the "back room to the boardroom."

About CDO Magazine

CDO Magazine is the leading global publication amplifying the voices of senior executives in data, analytics, AI, and security. Our editorial content delivers boardroom-quality insights, equipping leaders with the strategies needed to drive enterprise-wide transformation in an increasingly digital world.

With a strong commitment to building a connected global community, we bring together data, analytics, AI, and security executives through regional and international forums, fostering collaboration and knowledge sharing at the highest levels. CDO Magazine also recognizes and celebrates the most innovative companies, technologies, and thought leaders shaping the future of data-driven business.

Join our growing community of global data and AI executives—**subscribe today** for exclusive insights and event invitations, and **contribute your opinion** to help shape the conversation at the forefront of enterprise data leadership.

Subscribe or learn more at **cdomagazine.tech**.

About EY

EY is building a better working world by creating new value for clients, people, society and the planet, while building trust in capital markets.

Enabled by data, AI and advanced technology, EY teams help clients shape the future with confidence and develop answers for the most pressing issues of today and tomorrow.

EY teams work across a full spectrum of services in assurance, consulting, tax, strategy and transactions. Fueled by sector insights, a globally connected, multidisciplinary network and diverse ecosystem partners, EY teams can provide services in more than 150 countries and territories.

Learn how EY is shaping the future with confidence at [**EY Emerging Technology**](#).

Acknowledgements

Primary Contributors

- [**Camille Prado**](#), CDO Magazine Global Editor
- [**Abid Hussain Barlaskar**](#), CDO Magazine Managing Editor
- [**Traci Gusher**](#), EY Americas AI and Data Leader
- [**Zakir Hussain**](#), EY Americas Data Leader



Chapter Contents

Chapter Overview	1
Why does data need to be governed?	2
Who needs to be involved in DG?	5
When is it appropriate for organizations to invest in DG?	5
Where should organizations get started with DG?	6
How should organizations apportion their DG efforts over time?	7
What organizational needs does DG fill?	9
Chapter Summary	22

Chapter Overview

This database ain't big enough for the two of us

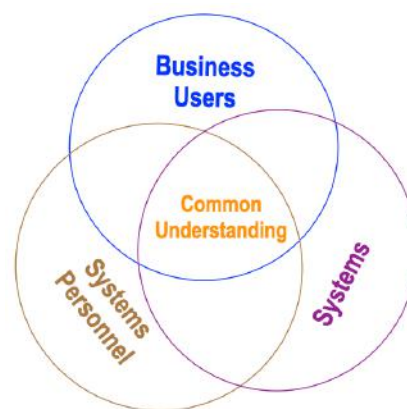
— Bumpersticker seen on an automobile in Texas

The bumpersticker should really have stated “There is no database big enough for two bosses.” Importantly, 1) this has always been true, and 2) it means absolutely nothing to most of the public or much of IT. Let’s address each of these separately.

Just as in any situation where coordination, integration, and information are required, there must be one and only one individual implementing decisions to maintain integrity, continuity, and operational capabilities. Required minimally from a change management perspective, this can always be used to justify DG in general. Ask the skeptical: how can any complex adaptive system function with multiple Chiefs?

The public and unfortunately too many in business and IT do not understand this sort of basic law of (data) nature. Because they are not data literate, when someone proposes having multiple chiefs for database operation, or that group X should ‘own’ dataset Y, or that the DG group should report to the CIO—they do not know these are not a workable concepts!

DG is not focused primarily on databases, clouds, or other technological ephemera. Instead the DG framework must be understood identically by business users, systems personnel, and the systems themselves (as shown to the right). This essential, metadata-based communication is at the heart of any enterprise operation. DG removes barriers to data efficiencies, allowing organizations to function more effectively and efficiently. Resources consumed by bad data practices can now be used to support the mission.



Increasingly organizations are attempting to do 'more' with data. This represents the other strategic dimension, innovation. By definition, most attempts to innovate will fail—so the lessons learned by becoming more effective and efficient, will also help in this innovation dimension. Innovating with data requires programmatic support for the efforts—well supported by data infrastructure and mature organizational data practices.

It is the responsibility of DG programs to manage this and other delicate balancing acts required to successfully contribute to better organizational use of data. DG is a comparatively new, certainly unstandardized, and under-studied topic. While some excellent DG programs are maturing, the majority have not. This leaves individuals and organizations the sequential tasks of:

1. learning about data (and then)
2. learning about their data (next)
3. developing plans to increase the data literacy of their executive leadership (then)
4. their knowledge worker population (before expecting to)
5. make progress faster and further with data.

This chapter takes you through the who, what, where, when, why, and how of DG. It provides a common-basis for building individual and organizational knowledge of this topic—starting with the **why**—the motivation for DG—followed by the **who**, **when**, and **where**. The **how** section is a bit longer and the bulk of the remaining material concentrates on the **what**—a way to successfully start to govern subsets of your data.

Most organizations should not attempt to govern all of their data. Successful DG program goals include subsetting their data into essential and non-essential data. Governing the essential subset and ignoring (or better still removing) the rest reduces the size of the challenge. Since the definition of an organization's essential data will differ from organization to organization, the governed data will also differ among organizations.

One quick word about the use of the term **bespoke** in the title. It is of course deliberate. The only way that **your** organization can use data to better support organizational strategy, is to use **your** data in support of **your** strategy using the capabilities that **you** currently have. Cookie cutter methods will not help **your** organization learn about **your** data!

Why does data need to be governed?

A friend was speaking with an organization on data matters and noticed that the urinals in the restrooms all had unique numbers. Presumably this was in case of malfunction so that the specific instance could be more rapidly identified. Of course my friend used a suitable-for-work (as opposed to not-suitable-for-work) photograph to make a point to leadership that (at least for this organization) it was worthwhile to keep

maintenance histories of this equipment type. Ironically, it was noted that the substance of the discussion for which my friend had been invited was whether the organization should maintain similar information about their organizational data assets. The photo provoked a nice motivational discussion with a decision to proceed with DG as the outcome. After all, if we are going to govern our restroom facilities, shouldn't we also govern our data assets?



Writing as a deeply, industry-immersed university professor, I can say that the academic community has failed its customers with respect to integrated data knowledge. For generations we have graduated students who have become leaders in business and IT. The only class taught about data was really about database development. Smart students who placed their trust in the educational system, were educated that the only concept they needed to learn about data was *how to build new relational databases!* No one should be surprised that one of the major DG challenges is that far too many poorly designed databases clutter most organizations or (more increasingly) their clouds. As Abraham Maslow stated: **"If the only tool you know is a hammer, every problem looks like a nail."**

When considering the asset itself, data has a unique collection of properties including the following from Doug Laney. Data:

- Does not obey all of the laws of physics
- Is not really visible
- Is non rivalrous (many can use it at once)
- Costs of providing an additional copy are zero
- Is non depleting
- Does not require replenishment
- Is regenerative
- Has low inventory and transportation/transmission costs
- Is more difficult to control and own than other assets
- Can be eco friendly
- Is impossible to clean-up if you spill it.¹

When considering career fields and learning experiences, not all data professionals take similar paths. For example, data scientists often **discover** useful data maintenance utilities instead of **learning** that various classes of tools exist and when to apply each as part of their educational programs. For many, data is like the story of the blind men and the elephant and collectively it is DG responsibility to shape this understanding into an organization-wide perspective.

For these and other reasons there continues to be questions as to whether data processing should continue to be part of IT or of the business or of special operations

¹ See *Datanomics* by Doug Laney Routledge Publishing 2017 ISBN 1138090387

such as finance and risk? While the Federal Government resolved this issue correctly with new FEPA legislation, the jury is still out on the rest of the world. Currently it is comprised of $\frac{1}{3}$ of each type: one-third reporting to CIOs; $\frac{1}{3}$ reporting to CEOs; and $\frac{1}{3}$ reporting to CFOs/CROs.

Long lasting consequences of poor data decisions?

Unfortunately, short-term application-centric thinking² has dominated, relegating development of data products to subsets of ERPs, digitization initiatives, or cloud hosted projects (to name just a few types). Virtually none of the popular software integration packages from the major vendors have escaped the long-term consequences of inadequate data Design (big 'D' is used to emphasize the entire lifecycle). These well documented imperfections are locked in for life-wrapped as they are, in a dense set of application constructs interwoven with the imperfect data model. Worse still, the corrections to the organization's data and processing are layered on as additional code—complicating the apps still further. The vast majority of database functionality is not used beyond table-handling. In this manner, developers restrict any subsequent data investment benefits and decrease data leverage potentials. At the very least, DG must illustrate and resolve the 20-40% of IT budgets that are devoted to data evolution:

- Data migration (Changing the data location)
- Data conversion (Changing data form, state, or product)
- Data improving (Inspecting and manipulating, or re-keying data to prepare it for subsequent use)

None of these are accounted for in the usual (and very important) **data storage costs—measure**. DG must also articulate these various costs and tradeoffs associated with increased data rigor (or the risks of not doing so) to the rest of the organization.

Mounting data debt

The failure to do any of this has caused organizations to pay to accumulate large amounts of data debt. (Yes, the indignity that your own organization is creating data pollution that is directly harmful to its operation should be professionally embarrassing!) It is not easy to visualize the cost of data debt but the phrase *many many unnecessary paper cuts* describes the situation well. Data debt slows DG efforts making everything slower, of lower quality, cost more, or present increased risks.

Data debt is like quicksand that mires down all efforts. Defined simply, data debt is: the time and effort it will take to return your data to a governed state from its likely current state of ungoverned. A quick back of envelope calculation of data debit can be

² See *The Data-Centric Revolution: Restoring Sanity to Enterprise Information Systems* by Dave McComb Technics Publications ISBN 1634625404

done using the data storage costs that are perhaps the most tangible and objective data measure. At least 20% of that data is redundant, obsolete or trivial (or ROT).

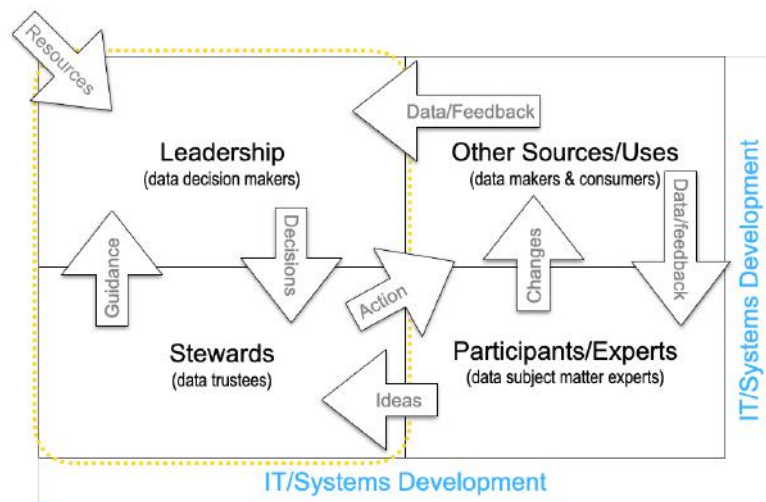
The good news about finding and eliminating data debt is that things can get faster, better, or cheaper. The bad news is that new skillsets are required of the DG team and that diagnostic and analytical systems thinking still requires annual proof of value. The knowledge-base of grey-beards who know how to apply these skills is shrinking as these individuals are judged expensive and encouraged to retire.

In summary, data needs to be governed because society was not taught that it required specific treatment until it was too late. Because individuals do not know that they do not know, it has been difficult to educate them to the need. By focusing on concrete results, organizations have better success making the case that an investment in DG will benefit the organization in specific measurable ways.

Who needs to be involved in DG?

Unfortunately at many organizations, **everyone** has been responsible for data quality and this approach has produced the current unsatisfactory state. It is critical to start DG educational efforts with executives because 1) they are willing to invest in learning and 2) their data decisions have the greatest impact on the organizational data practices. The next goal for all DG programs is to also increase the data literacy of all organizational knowledge workers.

As illustrated, DG efforts are generally built on an IT provided support/foundation/infrastructure. A leadership component provides resources and clears barriers for the effort. Primary functions are (ideally full-time) data stewards who provide guidance and design/implement decisions. Typically these two groups form the basis for DG organizations. Also, highly involved (and incorporated) are various SME or subject matter experts who know the required data and processing details. Then of course there is everyone else. As noted, DG efforts need to be integrated with both organizational and IT governance.



When is it appropriate for organizations to invest in DG?

By now I hope that you agree this is a silly question. The 20-40% of IT costs (referenced previously) is easily gauged. As the maturity of the DG practice matures,

processes can be optimized for key operations. By keeping disciplined measures, organizations have developed expertise in these practices. Keeping the focus on a integrated full time team permits the case to more easily be made when timing investment in a second or third DG team.

Digital and data are dependent on high speed automation/data processing that requires significant amounts of organizational data literacy, data standards use and quality data supplies. Continue to evaluate and evolve DG frameworks to refine the organizational focus. Over time this approach should evolve into the standard Deming plan, do, check, act (PDCA) cycle.³ An incomplete list of potentially useful standards that can be created with the required measurable controls is listed below.

- Access standards
- Change management
- Security
- Storage
- Reporting
- Classifications
 - ➔ Secure
 - ➔ PII
 - ➔ Competitive advantaged
 - ➔ Public

Where should organizations get started with DG?

DG is a rare triple benefit capability that helps refine data strategy, improve the quality of the players, and improves data used to support the mission. However, getting started with DG can be and has been accomplished by a moras of ill-defined and vendor specific methodologies—most of which have no reported research results.

An easily understood model (the theory of constraints⁴ or TOC) views programmatic data support as a manageable system. The system is limited in achieving more of its goals by a small number of constraints. There is always at least one constraint, and TOC uses a focusing process to identify the greatest constraint and restructure the rest of the organization to address it. TOC adopts the idiom that "a chain is no stronger than its weakest link," and processes, organizations, etc., are vulnerable because the weakest component can damage or break them and adversely affect the outcome.

Key is to visualize the various data flows through the organization and understand the value of controls in relation to various processes, risks, outcomes, and performance. The costs of various blockages can be ranked and estimated. What changes made at the data level could most help the organization achieve its strategic goals?

Iterative problem solving provide additional benefits beyond challenge solutions. Team problem solving enables increased organizational data literacy and some go as

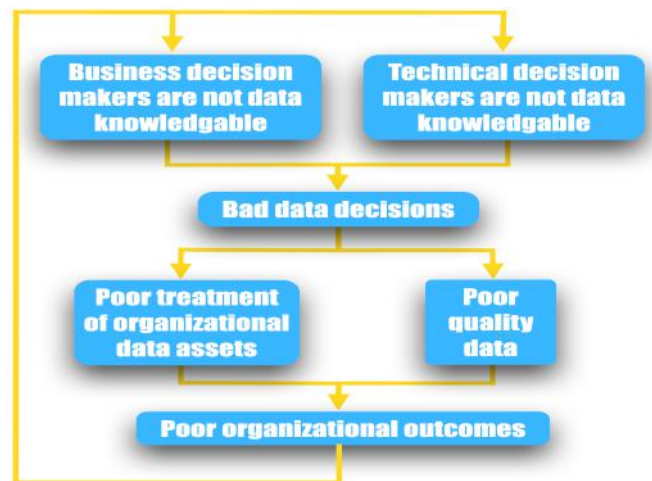
³ https://en.wikipedia.org/wiki/W._Edwards_Deming#PDCA_myth

⁴ https://en.wikipedia.org/wiki/Theory_of_constraints

far as considering these capabilities their 'secret sauce.' It just makes sense to support a group of individuals who possess knowledge of your data and its uses.

Focus first on organizational strategy. Understanding intricately, the data flow supporting increasing performance, decreasing costs, impacting times, and better managing risks. Identify the various types of organizational challenges sharing the same data or (better still) data errors. These become the focus of the first iteration of a data strategy cycle. It is overseen by the DG program and coordinated to be most collectively helpful to organizational as well as IT strategy. Ensure you complete a full cycle to include feedback/improvement/lessons learned/organizational memory/change cycle components. Heavily incorporate the use of 'branded' data checklists and standard control development.

And then (as it says on the shower bottle) lather, rinse and repeat. This is really to only way to escape the bad data cycle. IT and business decision makers are not knowledgeable about data and good data practices. They make poor decisions about data that result in poor treatment of organizational data assets and poor quality data. Both of these lead to poor organizational outcomes.



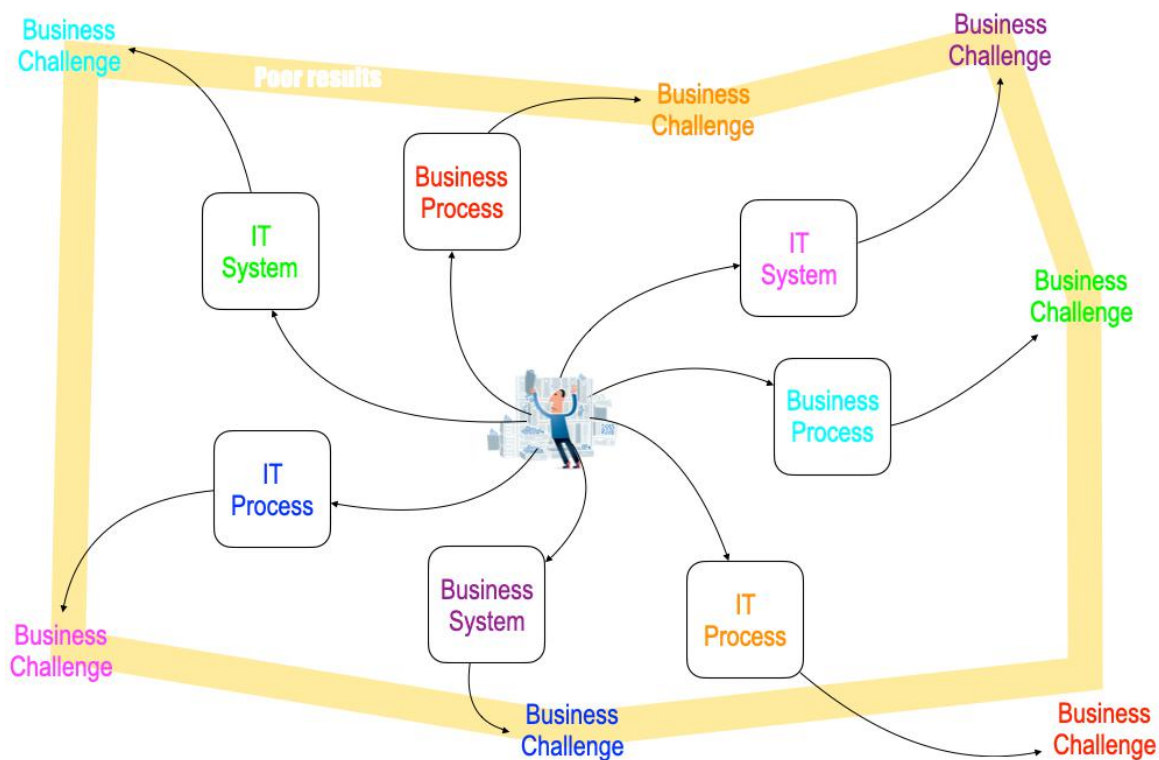
How should organizations apportion their DG efforts over time?

Data debt's impact

Over time, organizational data debt clogs value-adding pathways in a manner similar to the 40% of the internet that is now clogged with malware. Data debt is responsible for inflicting uncounted tiny hidden data factories⁵ on organizational performance-making everything cost more, take longer, deliver less, and at increased risk. Eliminating data debt requires a team with specialized skills deployed to create a repeatable process and develop sustained organizational skillsets.

A major motivation for increasing the data literacy of all knowledge workers comes from the fact that most organizational challenges come filtered through various IT and business practice combinations. The reason for multitude of paper cuts, is that the DG challenges are filtered through various business processes and IT systems. As a result, common challenges go unrecognized with each instance requiring treatment instead of correcting the underlying data challenge.

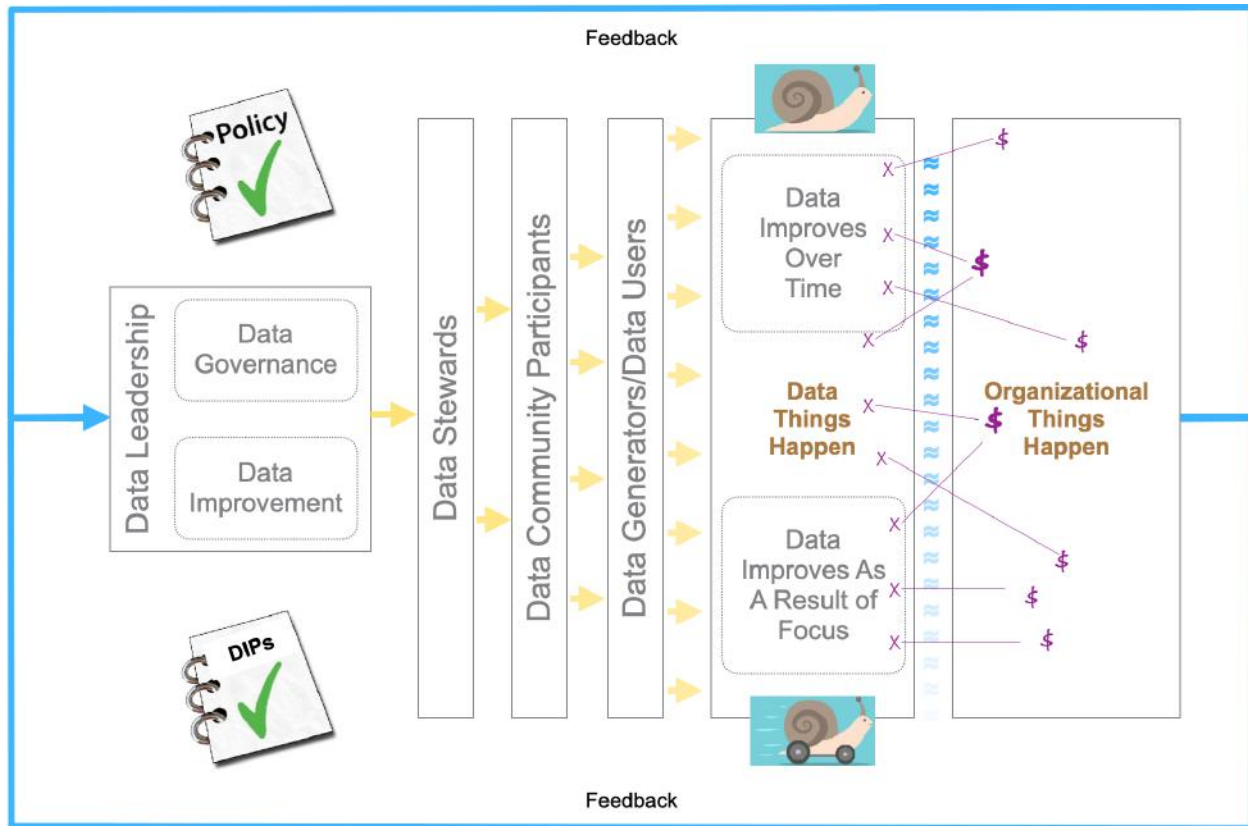
⁵ <https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year>



A key aspect is to evaluate your architectural abilities to build/evolve towards organizational data capabilities in a 3-step process. First, you need to improve the quality of existing organizational data. Too many organizations do not have enough information about the quality of their existing data. These data quality challenges fall into two categories: practice-related data quality challenges and structure-related data quality challenges. Second, the framework must support your efforts to increase the data literacy of literally your entire executive team and knowledge worker population and especially those who already practice data. Finally, only when you have improved your data and your organization's ability to work with data, can you hope to improve the way that data supports your organizational strategy.

Proactive versus reactive DG

One rather traditional realization (almost a rite of passage) is that what ever changes are made to the organizational data practices might take literally years to be able to exploit it. In CIO terms, it can often be a successors, successors, successor that will benefit from DG initiatives. As this realization sets in (that time equals years), DG initiatives come under pressure to 'do something more quickly.' As illustrated, a secondary capability is established to more effectively produce results as a result of direct intervention or Data Improvement Projects (DIPs).



MacGyver abilities

While perhaps not widely acclaimed, the 1980's TV series MacGyver became shorthand for an non-traditional and innovative problem solver who always carried a Swiss Army knife.⁶ In the same manner, the DG program must imagine itself as the 'help desk' for organizational data. Literally all data challenge solutions should be minimally coordinated and, in many instances, led by DG. The key is to develop new data capabilities within a dedicated group focused on organizational data governance. Have this group focus on and conquer a series of DG challenges, producing positive ROI numbers.

What organizational needs does DG fill?

It is useful to describe the organizational needs that DG fills. These include:

- Improving the way that data is treated as an asset
- Available but not widely known research results
- Using data to better to support the organizational mission
- Using data strategically

⁶ <https://en.wikipedia.org/wiki/MacGyver>

Improving the ways that data is treated as an asset?

One of the primary challenges for organizations is to learn how data requires specific considerations. If you consider data as an asset (and currently most business leaders do not yet do so) then one should expect that it would be treated as other organizational assets. I use a series of questions developed by my colleague Dr. Christopher Bradley to help organizations determine whether their data is maintained as an asset. They are:

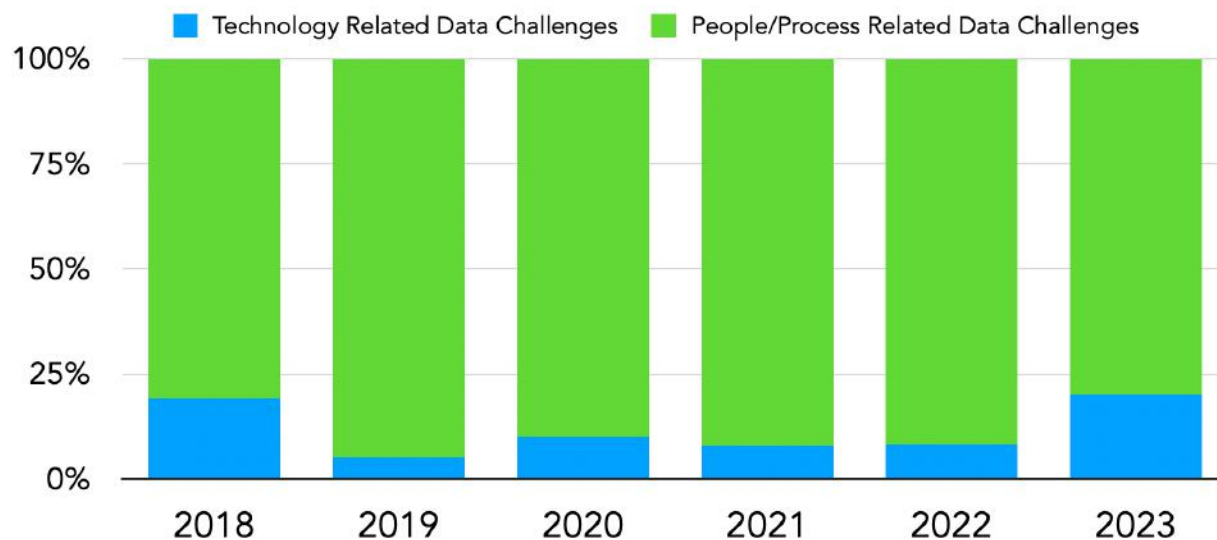
1. Do you have executive positions to support data as an asset?
2. Does the organization track usage of this asset?
3. Are organizational or fiscal controls in place to manage this asset?
4. By and large, are these controls actually executed?
5. Is there general acceptance of the need to manage this asset?
(i.e., do people "get it"?)
6. Do serious discussions about this asset feature on the agenda of senior management meetings?

Using this rather obvious set of criteria, it is easy to determine that most organizations are not treating data as an asset but so far we do not survey results on this particular measurement.

Available but not widely known research results

As referenced above, there is a dearth of knowledge about data much less data governance. On that note however, we do have access to two solid lines of research to which I will refer through out this chapter. The first is in the form of the annual (2013-today) data practices surveys conducted by New Vantage Partnerships and are reference able at: <https://www.newvantage.com/thoughtleadership>. Annually several thousand of the same or similar organizations have been asked the same questions repeatedly providing pictures of how issues are considered over time. Results reproduced here will be referred to as **New Vantage**. A second set of research results come from the collaboration (called the Data Literacy Project) between Accenture and Clique. These results will be referenced as **Data Literacy Project** and are reference-able at: <https://thedataliteracyproject.org/>. These two efforts have provided a good framework that can be used to dive further into research in this area.

One of the New Vantage results has been: *what percentage of your data challenges are people/process related versus technology challenges?* The consistent answer (see figure) continues to surprise: not once since 2018, has the percentage of technology challenges risen to above 20%. This means that for more than six years, **everyone** should have known that the people/process dimension of DG represents the largest challenge. Yet very little organized research beyond surveys has been conducted into this area.



Consider please, what group in your organization is charged with decreasing the number and impact of people and process-oriented data challenges? This is precisely the role that your DG organization must address in your organization. **If not DG then whom** in your organization is responsible for improving the people and process aspects of your data operations?

It is crucial that DGs provide a wholistic view of minimally the above detail but also include data's role in the organization, how individuals can assist, and where to go for more information.

Using data to better to support the organizational mission

This section's title: **using data to better to support the organizational mission** must be the mission of any DG program. But first a specific word about data ownership (bad concept) and Data requirements ownership (good concept).

Avoid a first (and always a major) misstep: trying to assign data 'ownership.' While it is tempting to "establish data owners" as a goal of data governance—it is usually a bad idea. However, many are familiar with the process architecture practice. It correctly embraces and leverages the term "process owner" as the single individual responsible for the integrity of the process design, implementation, and improvement.

While it makes intuitive sense, the concept of data ownership has caused more DG effort to fail than any other. As soon as you allow an under-informed individual (or group) to 'own' any data items, they begin to make decisions about the data that optimize it from their local perspective. If your organization does not formally manage a process architect, skip to the next paragraph. If it does, careful analysis will yield maintainable, high level process/data interaction matrix called a CRUD matrix—showing data/process interaction by access type. (CRUD matrices such as the one illustrated

show business processes and their activity type **C**reating, **R**eading, **U**psdating, and **D**eleting various data items—example also courtesy of Dr. Christopher Bradley).

		Business Processes						
		Product development	Marketing & Sales	Industrial preparation	Order management	Manufacturing	Logistics	Invoicing
Major Entities / Data Subject Areas	Product	C	R	U	U	U		
	Product Part	C	R	R	U	U		
	Manufacturing Plant	U		C	R	R	U	
	Customer	R	C		U	R	U	U
	Sales Item	C	C	C	U		U	U
	Assembly Structure	U		C		U		
	Sales Order		U		R	U	U	U
	Production Order			U	C	U	U	U
	Individual Product					C	R	U
	Shipping						C	
	Customer's Invoice		U					C

If nothing else, these maintainable metadata collections show the interdependencies: data exist only to be consumed by various business processes and only purpose for a business process to exist is to produce data to be consumed by another business process. If you do not have an organization CRUD matrix hand and need to shut down any data ownership conversations, ask the question: *"To whom does the data that accounting stewards belong?"* Since accounting processes data from across the organization, a case could be made that accounting 'owns' much organizational data.

The reason data ownership is such problematic concept is that data persists across business functions. Ownership would only apply to a specific data processing stage. Instead of asking the question, "who are the data owners?" the statement should be that all data belongs to the organization! At best, ownership could only be limited to specific lifecycle phases.

If the organizational culture requires use of the word ownership, then allow ownership of the **data requirements**! Local expertise should be used to specify the size and shape of the specific data items required to perform organizational functions at various stages of data at it is processed.

The role of DG frameworks

All evidence to-date points to frameworks has being useful as:

- System of ideas for guiding subsequent analyses

- Means of organizing measures, project data, and then assessing progress
- Evaluating priorities for data decision making
- Assessing overall functionality
- Moving towards a determination of ROI⁷

For example, a building construction conceptual framework would incorporate bits of wisdom such as:

- Don't put up walls until foundation inspection is passed
- Put the roof on ASAP so that work can proceed in inclement weather
- Make it each construction phase dependent upon continued funding by passing a series of checkpoints

Much has been written about data governance frameworks. I have seen research proposals that anticipate evaluating one type of framework against another. It is far too early to start to 'type' DG frameworks. Non-standard understanding of terms and data concepts lead to 'results' of the sort that were popular at the start of the CDO movement. (Note: researchers have tried and failed to establish correlations between having a CDO and organizational financial performance—similar specious results can be expected until the entire DG profession matures.)

Use the existing DG frameworks to envision what your program should look like given your organizational needs. 'Try each of them on' conceptually and discuss the suitability of each for your organization. Since no two organizations are alike, each organizational DG program must be custom fitted to the organization rather like getting fitted for a suit. The word "bespoke" well describes the design of DG programs that provide good returns on organizational DG investments.

It is quite useful to view representations of various approaches to DG in the same manner that an architect presents sketches of a future building to prospective funders. The utility of DG frameworks generally stops at this point. There are essentially few types of DG frameworks in popular use. (Note: you can see representations of many of these at: <https://anythingawesome.com/DataGovernanceFrameworksCollection.html>) All subsequent are theme and variations on these. Pay no attention to 'proprietary' methods. The goal is to give you something to compare, contrast and consider when designing the first version of your DG organization. (Note: This first version will evolve to a second and third as the organization; DG practices should mature and evolve over time.)

This is where the concepts of stewardship and fiduciary responsibilities come into play. **Stewardship** in this concept is derived from the definition: a person employed to manage another's property. **Fiduciary** is used to describe the nature of the relationship

⁷ Interestingly, ROI means risk of incarceration to most DG professionals.

as involving trust, especially with regard to the relationship between a trustee and a beneficiary. This is accompanied by specific duties.

RELATED TERM DEFINITIONS

It is now time to introduce a few terms to show both the evolution/etymology of the term DG and the most useful definition of DG.

Let's start with the term **governance**: "Governance is the process of interactions through the laws, norms, power or language of an organized society over a social system (family, tribe, formal or informal organization, a territory or across territories). It is done by the government of a state, by a market, or by a network. It is the decision-making among the actors involved in a collective problem that leads to the creation, reinforcement, or reproduction of social norms and institutions" (<https://en.wikipedia.org/wiki/Governance>)

Corporate governance is next. Below are three good definitions highlighting different aspects of this evolving concept.

- "Corporate governance - can be defined narrowly as the relationship of a company to its shareholders or, more broadly, as its relationship to society....", Financial Times, 1997.
- "Corporate governance deals with the ways in which suppliers of finance to corporations assure themselves of getting a return on their investment", The Journal of Finance, Shleifer and Vishnu, 1997.
- "Corporate governance is about promoting corporate fairness, transparency and accountability" James Wolfensohn, World Bank, President Financial Times, June 1999.

Note that the concept of corporate governance is evolving. Just before the pandemic, Jamie Dimon (then head of Chase), led a group of CEOs to proclaim "Maximizing shareholder value can no longer be a company's main purpose."⁸ Similarly, the concept of DG continue to evolve.

Well if corporate governance exists then certainly **IT governance** should be a useful concept? It is and is defined as "Putting structure around how organizations align IT strategy with business strategy, ensuring that companies stay on track to achieve their strategies and goals, and implementing good ways to measure IT's performance. It makes sure that all stakeholders' interests are taken into account and that processes provide measurable results. (https://en.wikipedia.org/wiki/Corporate_governance_of_information_technology)

IT governance frameworks should answer some key questions, such as how the IT department is functioning overall, what key metrics management needs and what return IT is giving back to the business from the investment it's making. Included are typically foci on:

- Strategic Alignment
- Value Delivery

⁸ <https://www.marketwatch.com/story/maximizing-shareholder-value-can-no-longer-be-a-companys-main-purpose-business-roundtable-2019-08-19>

- Resource Management
- Risk Management
- Performance Measures

IT governance is an established discipline with common vocabulary and understanding among those who participate.⁹ Of note is the fact that data practices are not typically included as a topic under IT governance or are lightly treated. This may account for or reflect the current slowly maturing state of DG practices.

Data governance has suffered from both too many definitions and inaccessible (by the business) terminology. However, auditors easily get the concepts. Below are some standard definitions of DG.

- The formal orchestration of people, process, and technology to enable an organization to leverage data as an enterprise asset – The MDM Institute
- A convergence of data quality, data management, business process management, and risk management surrounding the handling of data in an organization – Wikipedia
- A system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods – Data Governance Institute
- The execution and enforcement of authority over the management of data assets and the performance of data functions – KiK Consulting
- A quality control discipline for assessing, managing, using, improving, monitoring, maintaining, and protecting organizational information – IBM Data Governance Council
- Data governance is the formulation of policy to optimize, secure, and leverage information as an enterprise asset by aligning the objectives of multiple functions – Sunil Soares
- The exercise of authority and control over the management of data assets – DM BoK

Technically they are all correct but imagine the following scenario. Stepping onto an elevator for a minute-long ride and an executive enters the car. As the doors close the executive turns and says, “I’ve heard you are working on DG. Can you tell me what it is - I’m confused?” Imagine responding with “DG is the exercise of authority and control over the management of data assets.” Do you think the executive would 1) find the answer useful and 2) think well of your ability to communicate this concept?

I think the answer no to both questions. A better response to the executive is: “**DG is about managing data with guidance.**” Short and to the point, this definition incorporates self explanatory motivation. When I provide this information (the definition of DG) to most executives, their first question to me is: “So we have not been managing our data with guidance?” The answer usually is: “Only recently have we been managing our data with guidance.” Of course the eternal hope is that the executive will be curious to learn more and present an opportunity to become more data literate. Subsequent conversation topics could include:

⁹ https://en.wikipedia.org/wiki/Corporate_governance_of_information_technology

- Why is it generally not a good idea to govern all of your data.
- Why DG will never be complete at our organization.
- Why some decisions that involve data are not considered as such.

The Data Literacy Project reports that four out of five executives surveyed were willing to invest time resources in improving data skillsets. This represents a once in a generation opportunity to reach these executives with good DG education. (Note that anyone offering to improve your organization with DG **training** should be ignored—the process requires **education**, not training.)

A SMALL CONCENTRATED TEAM IS PREFERRED OVER DISTRIBUTED (DISSIPATED) KNOWLEDGE

The next item to consider is what format DG should take. Remember, asking **everyone** to be responsible for (data, data quality, data governance ...) has produced the current state of affairs. Organizations assigning new DG duties to existing personnel have two options: 1) incorporate the new duties along with existing duties or 2) assign these DG duties to full time individuals.

When considering this, it is useful to ask: how long will the need to manage data with guidance exist? The answer turns out to be: **you will need your data program as long as your organization needs to have its finance, HR, and planning operations.** Think about it in the future: Will more or less data exist? Will data collection modes increase or decrease? Will data be found in fewer or more formats? A solid recommendation is to staff with full-time team members dedicated fully to DG. Data literacy and organizational data practice maturity are generally low. Dedicated personnel will interact with each other more—greatly stimulating their individual learning curves. It also makes tracking DG program costs clearer. It is critical to begin to build organizational DG capabilities. This can best be started with dedicated teams with a clear ROI. Against these, results can be evaluated.

Using data strategically

The next question is **on what** do we focus these DG efforts? In regulated environments, these efforts are often compliance driven. Key is to approach these efforts in the same manner. Do we think that regulations will increase or decrease in the future? If increasing, then it seems useful to 'get good' at implementing compliance driven changes. If nothing else, you may gain an implementation advantage over the competition subject to the same data regulations but perhaps not able to implement as efficiently or effectively. Data regulation compliance can become an valued organizational capability with an easily determined ROI.

Outside of compliance, organizations strive to use data strategically with either efficiency/effectiveness or innovation goals. Personal interaction with more than 1,000 organizations indicates that about ½ have clearly articulated strategic goals and

objective measures supporting goal achievement at the organizational level. Absent these, it is not possible to improve the manner in which data supports this Jell-O strategy. I also find universal disdain for 3-5 year plans, most of which fell apart rapidly with the onset of the Covid-19 pandemic. So just a word of caution, check your organizational strategy to ensure it has clear objective and measures before attempting to improve how data can support it.

STRATEGY IS ABOUT WHY

*...it's not **what** you do, it's **why** you do it...*

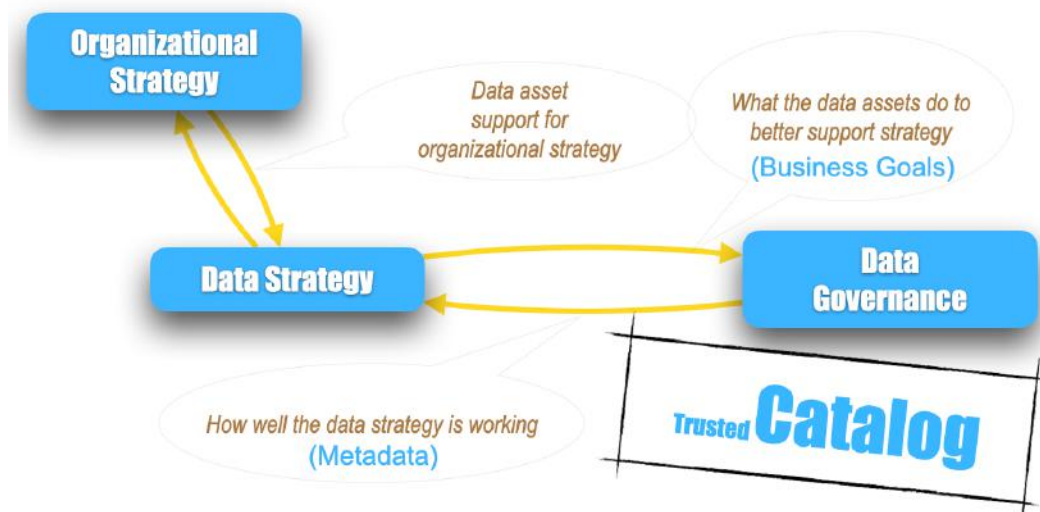
Among many great TED Talks, Simon Sinek's "How Great Leaders Inspire Action" is a favorite. Recorded in 2009, Sinek's talk has enjoyed more than twenty-five million views. His point is quite simple: most of us are very good at describing what we do, and some of us are good at describing how we do things. Not as many of us are good at describing why we do things.

Strategy is the highest-level guidance available to an organization, focusing activities on articulated goal achievement and providing direction and specific guidance when faced with a stream of decisions or uncertainties. More succinctly, strategy is a pattern in a stream of decisions. This pattern must be supported by data or it will not be possible to determine if the strategy is correct or working.

WHAT IS DATA STRATEGY?

Data strategy is the highest level guidance available to an organization, focusing data-related activities on articulated data program goal achievements and providing directional and specific guidance when faced with a stream of decisions or uncertainties about organizational data assets and their application toward business objectives. The data strategy must be understood and supported at the organizational level. Only with this level of scrutiny and involvement can a true systems view be applied to the challenge of improving how data can support strategy.

WORKING TOGETHER: DATA AND ORGANIZATIONAL STRATEGY?



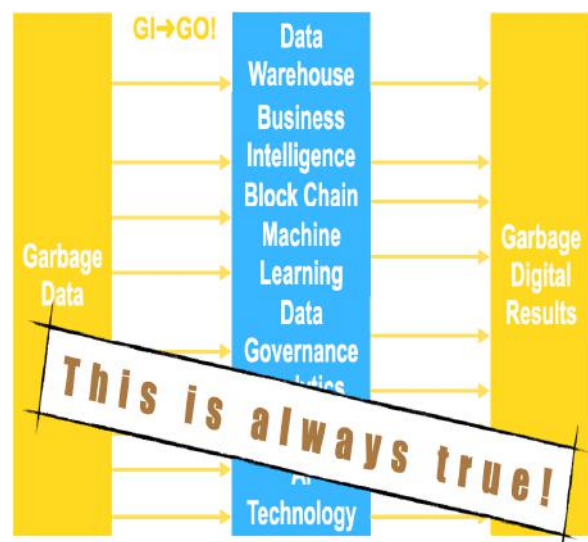
The figure indicates the close relationship among organizational strategy, data strategy, and data governance. Two key aspects of the interaction are: 1) express the data strategy in terms of specific business goals, and 2) ensure that the language of DG is metadata.

STRATEGIC COMMITMENT: PROGRAM VERSUS PROJECT FOCUS

A commonly asked question is: when will you be done? This is a warning that the individual considers DG a **project**. Organizations failing to implement DG at the program level (as a program) are unable to view the totality of their data challenges holistically and the solutions fail. Many organizations require a second or increasingly a third DG 'reset.'

DIGITIZATION

One of the more important areas that DG can be focused to support is 'going digital.' Once again, many vendors have offerings and expertise in these areas. DG sets the standards required to support digitization because, you cannot 'digitize' without a good data capabilities foundation. Garbage in, garbage out is always true. At this point, effective DG is a requirement for digitization, otherwise you can will be unable to trust any digital system outputs.



A WATCHFUL EYE TOWARD THE US FEDERAL GOVERNMENT (FEPA)

Finally on the **what** question (yes—we are still in **what**), it will be useful to observe the progress being made in the US Federal Government. As part of my service as a DoD employee, our group was often sent to ‘learn from the private sector.’ Now the situation has been reversed. In 2019 the **Foundations for Evidence-Based Policymaking Act** was signed into law. Three specific aspects of the law make this especially interesting to aspect of DG to follow. They are:

- Explicitly non-political CDOs must be established separate from CIO roles. From a DG perspective, organizations have been slower to adopt CDOs with non-CIO reporting role.
- Government data is now open by default, and must maintained using open standards. In just a few years, the Federal agencies will have developed a great deal of expertise in these areas.
- Use of open data and open models is required in policy evolution. Policy changes are only permitted with both models and datasets specified prior to the analyses and decisions.

Collectively these efforts, if fully implemented, will improve governmental decision-making and overall effectiveness. More importantly, all impacted Federal organizations are also rapidly, developing and implementing DG as compliance activities still further increasing the pool of DG professionals world-wide.

Breaking through the Barriers of Data Governance

There are a host of barriers to implementing DG. This include the usual failures to include change management and cultural refocusing as key dependencies. While the accounting profession has had literally millennia to develop GAAP, no such guidance exists for DG. There is a vast tendency to depend on technologies that are incapable of acting as silver bullets.

An example of these difficulties was illustrated in 2020 when Forbes ran an article on airline valuations.¹⁰ It purported to show how the airlines were monetizing the data in their frequent flyer programs. However, the buried lede was that in 2020, both United and American Airlines were valued at 10s of billions of dollars less than the anticipated value of the data in these programs. You had better believe that if airline leadership could have unlocked that value during the time most were avoiding flying (the pandemic), they would have unlocked it asap! The fact that they were unable to do so highlights the uphill climb that poorly fitting DG efforts face.

Some basic DG execution principles follow:

¹⁰ <https://www.forbes.com/sites/advisor/2020/07/15/how-airlines-make-billions-from-monetizing-frequent-flyer-programs/?sh=66da87a614e9>

- Ensure that the organization's data strategy is properly aligned with the business strategy. Implement a regular processes with key stakeholders to ensure proper alignment.
- Ensure that data debt is properly being managed and the process is under statistical control.
- Perform a capability maturity assessment or 'reassessment' to determine the required maturity. If the maturity levels are not meeting expectations, ensure that there is a remediation plan with a properly monitored work arounds.
- Consider refresher training for your knowledge workers and data professionals, e.g., Data Stewards, Architects and Engineers as a feedback mechanism for determining needed improvements and remediations.

Based on the organization's strategy, the DG group must determine are they to initially follow a model primarily focused as a:

- Utility-back office, efficiency goal
- Steward-more asset focused, quality goal
- Enabler-strategic partner, innovation goal

This should be determined through the building of the data strategy. If an organization striving toward a modernization transformation, DG should trend to an 'enabler.' To measure the effectiveness of an enabler, DG standards should be repeatable and statistically stable. The focus can be changed at a later stage but can usefully focus effort and discussions during initial phases.

Hopefully your organization will be spared major data catastrophes. It is more likely you will experience one or more in the future. In this event, attempt to learn as much as possible from the event. Take for example, the story of two major banks in the process of consummating an arranged marriage. The deal came down to a single spreadsheet containing many rows, each representing an asset. If an asset on the spreadsheet was to not be transferred, that row was hidden with agreement by both parties. After final agreement was reached, the spreadsheet was handed to a junior associate who was told to 'make it look nice for the Judge tomorrow.' Unfortunately late in the evening junior accidentally unhid hundreds of rows and did not notice! Presented to the Judge as the golden copy, the judge would not reverse-even on appeal.¹¹ As you might imagine, DG practices around the use of spreadsheets is quite extensive. I assisted one organization with the elimination of more than 400,000 legacy systems of a certain type. The list of preventable spending continues.

Unfortunately, the conversations have been generally unsatisfactory. Key to getting started with **data valuation** is to add up 'at least' instead of attempting to master the entire costs. I justified an investment into an organizational repository at one

¹¹ <https://www.businessinsider.com/2008/10/barclays-excel-error-results-in-lehman-chaos>

organization with a business case built on the premise of saving everyone in IT **1 hour annually**. The organization conducted surveys asking if the one hour savings was achieved. It was!

When determining the internal and external value of data two prerequisites exist: first, business and data strategies must support data monetization and second, DG must be effective and properly measured. Components of data value can include:

Internal

- Properly managed data debt
- Efficient usage of cataloging and master data management
- High trust in supplier and customer data integration
- Measured positive ROI

External

- Organizational data monetized in a public market or exchange
- Organizational data becomes a profit center
- Organizational data becomes a band-aid of adhesive strips

Sometimes it is easier to highlight the value with unfortunate examples with clear costs to society. Early COVID monitoring was inhibited because health care workers did not know to save MS Excel data sheet and workbooks as .xlsx instead of .xls files. The difference, unknown to the users, was that the older .xls files dropped all rows beyond the 16,000th or so row without warning. We will likely never know how much better performing the early monitoring systems were because all the errors are in one direction.

On a more cheery note, an agency charged with home evaluation/intervention, discovered that 40 questions on its evaluation assessment were immaterial. This shortened each interview by half and ultimately shifted more than \$1 million from overhead to service delivery.

In terms of execution, DG should be viewed as an iterative process that the organization is striving to **get better at!** Each cycle focuses on aspects of the various data challenges with a goal of eliminating or reducing the impact of a specific constraint. To understand the importance of this shift in thinking about DG, consider the circumstances where a plan was the goal. It was former President and General Eisenhower who said:

"In preparing for battle I have always found that plans are useless, but planning is indispensable"¹²

Mike Tyson's version is that *everyone has a plan until they get punched in the face*. A team knows how to react to unforeseen challenges and efficiently address the ones they have planned for. The PDCA cycle provides operational context.

¹² <https://quoteinvestigator.com/2017/11/18/planning/>

Chapter Summary

The word **bespoke** has evolved from a verb meaning 'to speak for something', to its contemporary usage as an adjective. Originally, the adjective bespoke described tailor-made suits and shoes. Later, it described anything commissioned to a particular specification. [Wikipedia](#)

The difference between data analysis capabilities and data requiring analysis is increasing. DG will continue as a maturing and growing field and can only be assisted by increased research into the various challenges outlined. Practice standardization and improvement are clearly the next steps on this industry's maturity curve. As a new discipline, DG works best directly addressing the manner in which data is used to support achievement of organizations strategy. There is no one best way and right now there isn't agreement on terminology, hence anything. Consequently, the only way to obtain a positive ROI on investments in DG is to ensure that **your** data is successfully leveraged using methods (**your** data strategy) that **your** knowledge workers and **your** executives understand.

The goal is to improve DG effectiveness and efficiencies (and the data itself) over time. The more data literate the organization, the easier the transformation. Perhaps now is more understood about the phrase quoted at the beginning of the chapter:

This database ain't big enough for the two of us

— Bumpersticker seen on an automobile in Texas



Acknowledgement

My colleague Rob Greaves made many helpful suggestions that were incorporated into this chapter.

I appreciate Dr. Christopher Bradley granting me use of the two concepts note herein.