“Age of Data”

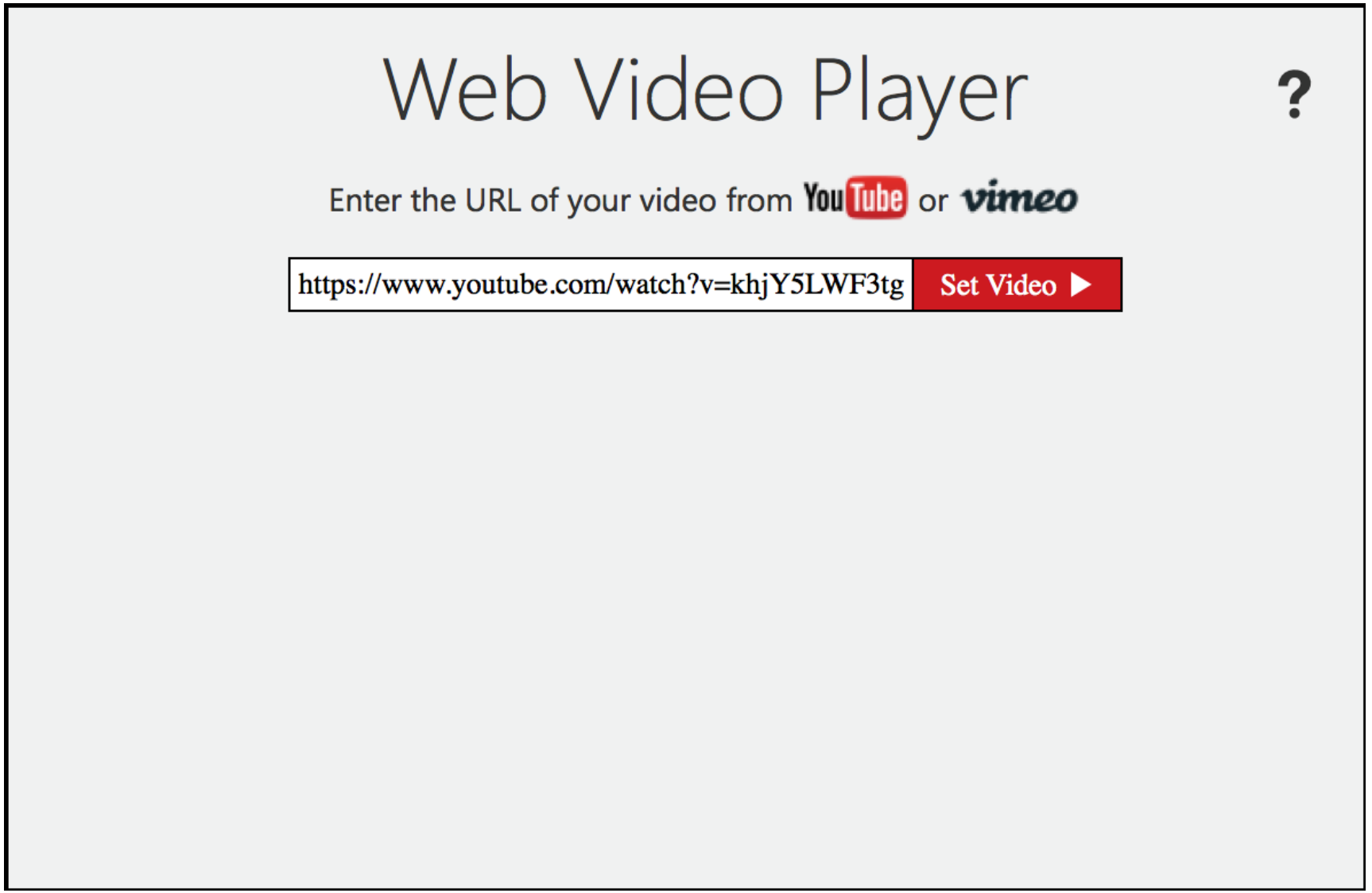
Creating Value and Managing Risks

@JayZaidi

DAMA – Minneapolis

Dec. 14, 2016

AlyData
Data Value Unleashed



Introduction

We live in the "Age of Data" and organizations that have mastered the art of generating deep insights from internal and external data are gaining a competitive advantage and thriving.

Data Management, Data Analytics, and Data Science are complex fields and not well understood. **What organizations need is a data culture coupled with a data savvy staff and a strong data management foundation**. That's what this presentation is all about.

6 Things I Shall Discuss

AlyData
Data Value Unleashed

1. Mission and Risk Alignment

4. Strong Data Foundation

2. Age of Data & Leadership 2.0

5. Data Management and the Supply Chain Pattern

3. Data Culture

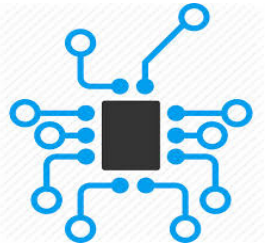
6. Operational Efficiency, Agility, and Cost Reduction

6 CEO Priorities

(http://blogs.workday.com/6-priorities-ceos-care-most-about/)

AlyData
Data Value Unleashed

Growth



Technology

Risk

Innovation

Regulatory Management

People and Culture

Data and Insights - The "Strategic Assets" that enable the six CEO priorities.

Age of Data – Points to Ponder

AlyData
Data Value Unleashed

"Facebook has relationships with 2.4 billion users. The Roman Catholic Church 1.2 billion. ***Facebook has more relationships on the planet than God.***"
- **Scott Gallaway, Marketing Professor, NYU**

The Fourth Industrial Revolution "We stand on the brink of a technological revolution that will ***fundamentally alter the way we live, work, and relate to one another. In its scale, scope, and complexity, the transformation will be unlike anything humankind has experienced before.*** This is the Fourth Industrial Revolution or the digital revolution that has been occurring since the middle of the last century. It is characterized by a fusion of technologies that is blurring the lines between the physical, digital, and biological spheres."
– Klaus Schwab, Executive Chairman of The World Economic Forum

"Succeeding with data isn't just a matter of putting Hadoop in your machine room, or hiring some physicists with crazy math skills. It requires you to ***develop a data culture that involves people throughout the organization***."
- **DJ Patil, Chief Data Scientist of the U.S.**

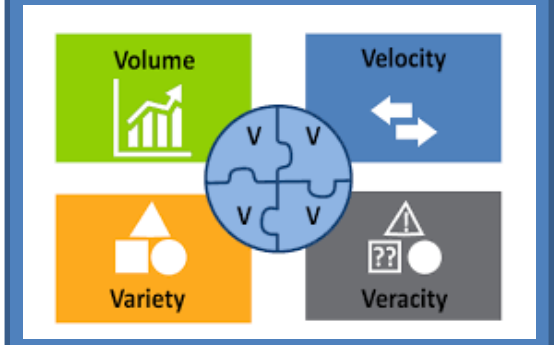
"***Culture eats strategy for breakfast, technology for lunch, and products for dinner, and soon thereafter everything else too.***"
– Business Management Guru Peter Drucker

The data deluge caused by Social, Mobile, Analytics, Internet-of-Things, and Cloud (SMAIC) is a reality and is fundamentally impacting every individual and organization. Organizations that wish to succeed with data must develop a **data culture** .

3D's of Data – Industry Trends

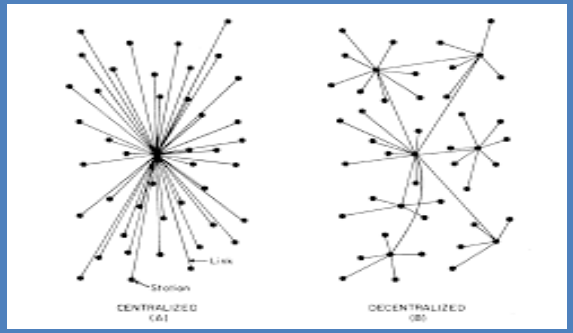
AlyData
Data Value Unleashed

Data Deluge



40 to 60% Average Annual Data Growth (Strains existing data infrastructure & resources)

Data Decentralization



Cloud & Distributed Data (Exposes organizations to cyber theft)

Data Democratization

Self Service Analytics (Exposes organizations to Insider threats and data leakage)

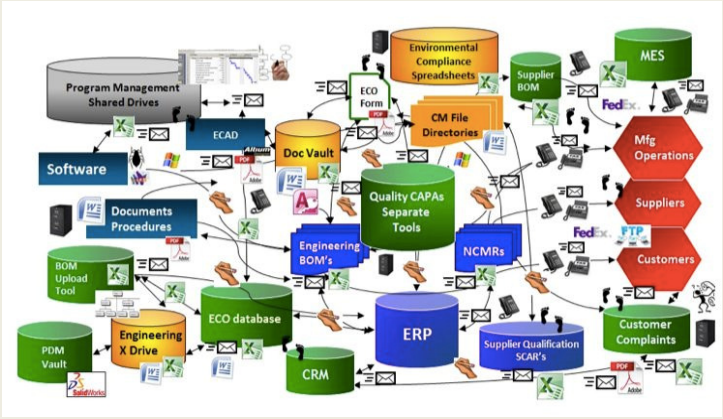
The 3D's of Data create an opportunity for improving data infrastructure and processes – focused on data platforms, data quality, data security & privacy, data governance, and analytics.

Mission Focus - Value Creation and Risk Mitigation

AlyData
Data Value Unleashed

GOALS

REALITY

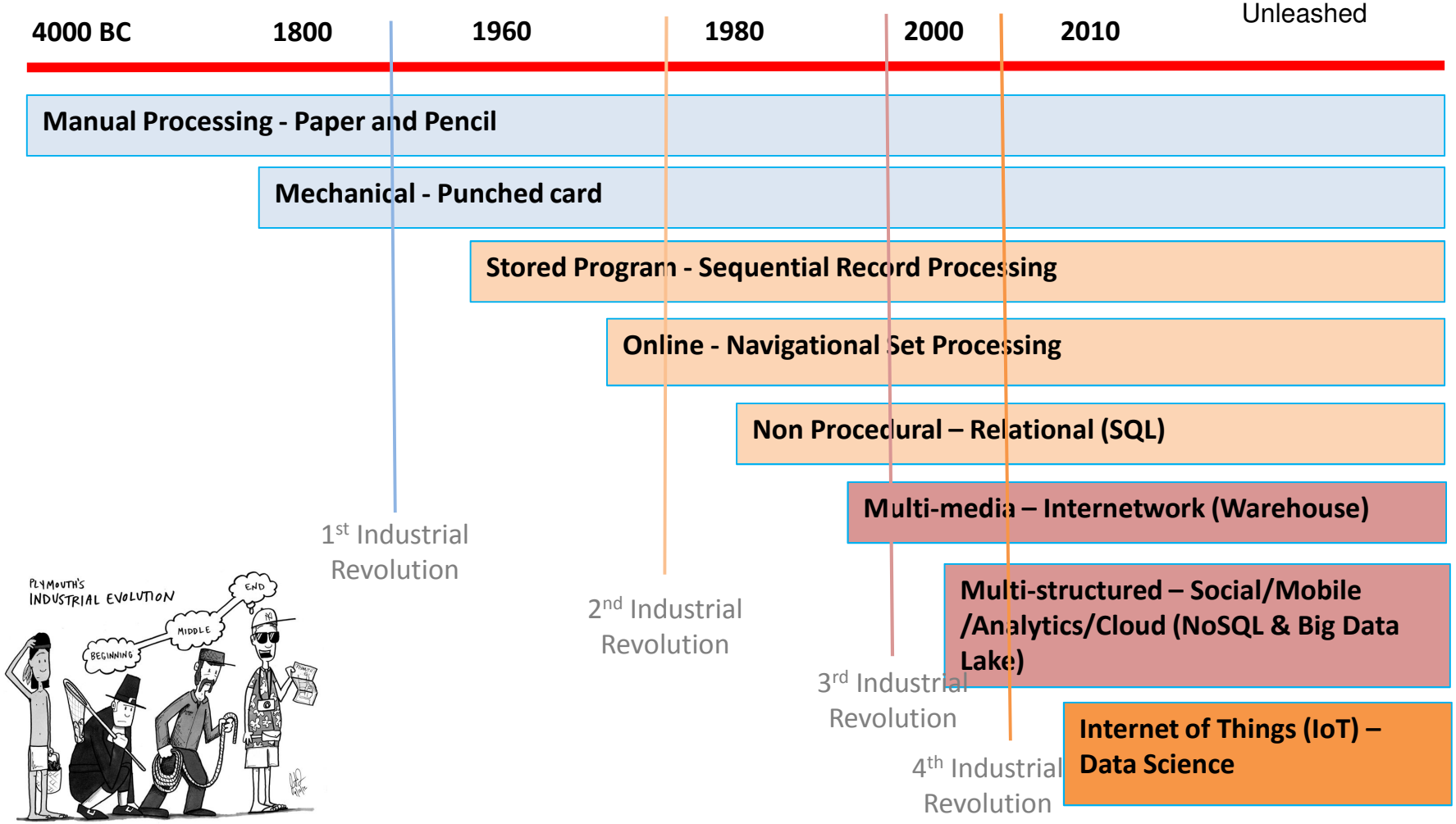


DATA LIFECYCLE

Getting from raw data to insights is the goal, but it requires a 5-step process and maneuvering through a complex data ecosystem.

A Historical Perspective of Data Management

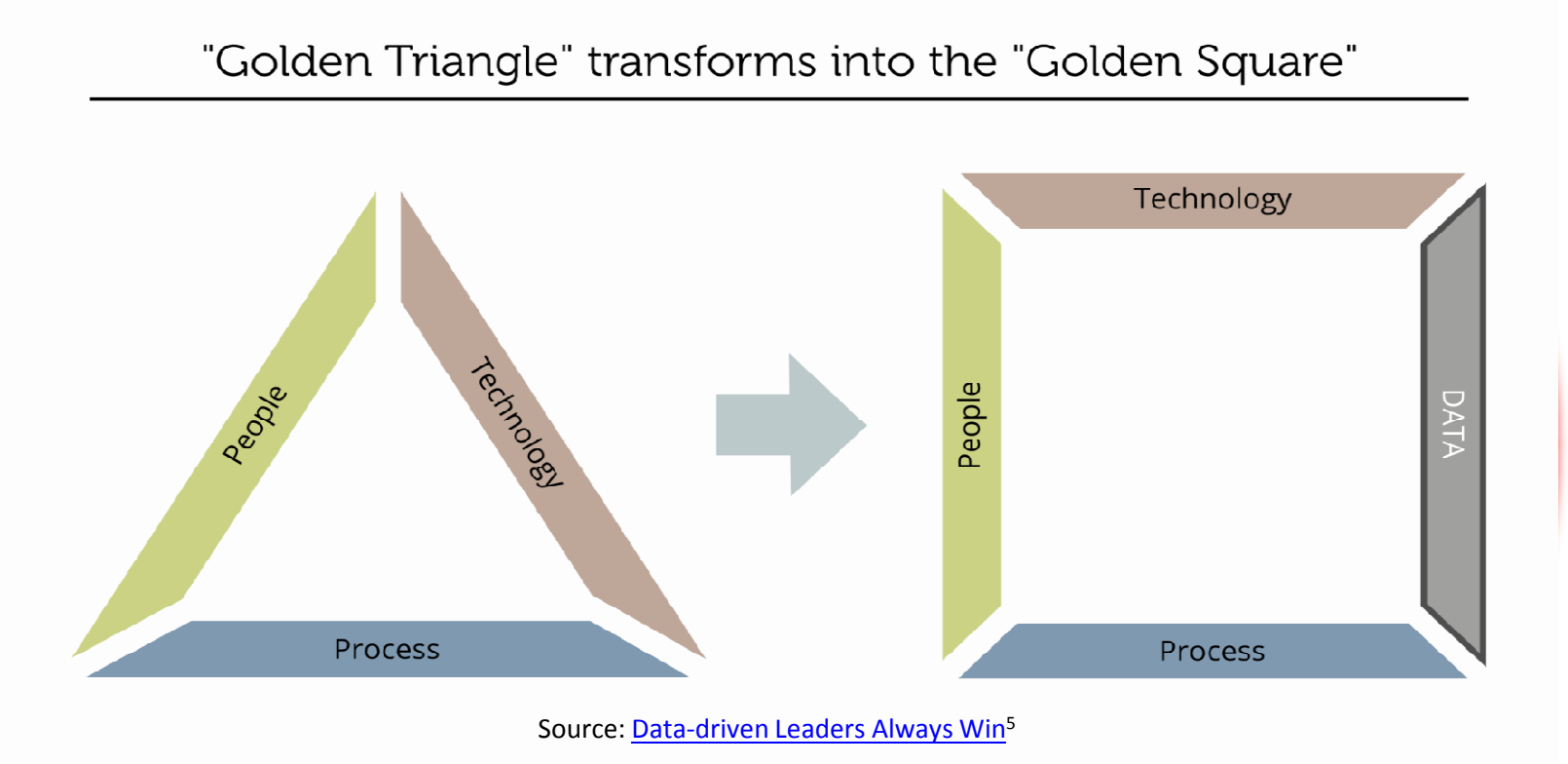
AlyData
Data Value Unleashed



Data Management has evolved from the mechanical world to the Internet-of-Things – to address business needs and our ability to process multi-structured data. Organizations are at various levels of data management maturity and using technology from the 2nd, 3rd, and 4th industrial revolutions.

Leadership 2.0 – A New Management Model for the "Age of Data"

AlyData
Data Value Unleashed



Source: Data-driven Leaders Always Win[5]

In this new management model data will play a critical role in organizations where decisions are made and power and authority are distributed on the basis of knowledge, rather than organizational hierarchy.

11 Characteristics of a Mature Data-Culture

AlyData
Data Value Unleashed

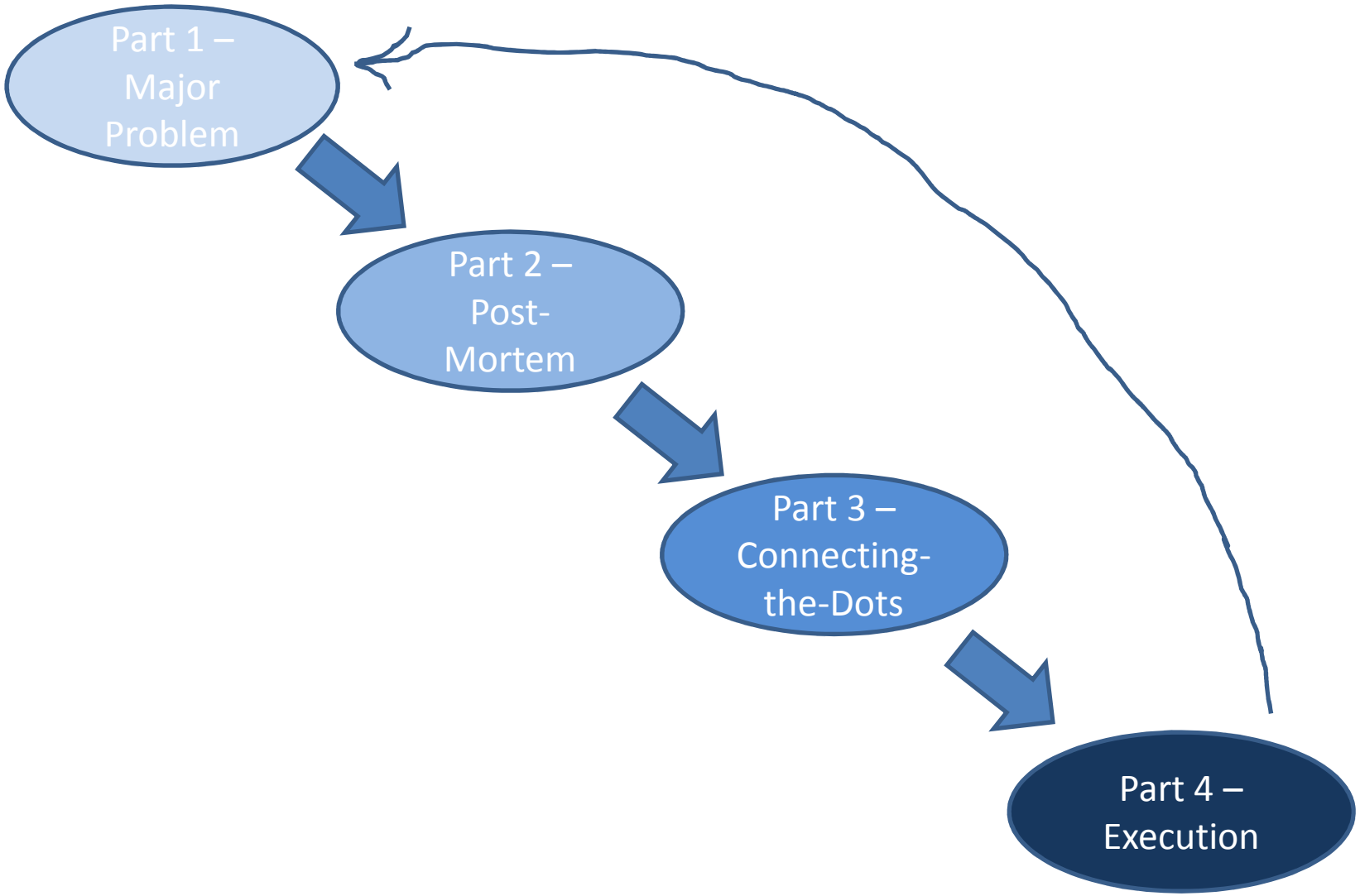
1. **Mission Alignment**: Data's role in the company's overall mission and goals is clearly articulated. Openly discussing strategies and innovation goals provides employees with a clear view of data's role in the company's overall mission and reinforces their connection to the larger organization.
2. **Behaviors**: Everyone makes evidence-based decisions (not based on gut) and leaders lead by example.
3. **Right Incentives and Alignment**: Cross-functional solution teams are completely aligned across goals and incentives between IT, Data, and Business staff
4. **Right Questions**: Leaders and staff are empowered to ask the right questions such as – what is the system of record for data? what's been done to it? can I trust it? who is accountable for specific data? etc.
5. **Information Supply Chain**: Departmental silos of information are the nemesis of thriving data cultures. To promote the view of data as a flexible asset that's usable by multiple departments, organizations need to educate employees on how the data they use daily ripples through other parts of the organization. Employees need to see the big picture.
6. **Rewards and Recognition**: Data successes are shared and individuals and teams responsible for them are rewarded and recognized.
7. **KPI Transparency**: Availability and use of key data metrics and measures via comprehensive dashboard – data quality, data issue management, data governance, data security and privacy, data lineage, etc.
8. **Data Sharing**: There is sharing of data and information between departments and total transparency – no data hoarding. A thriving data culture depends on an environment in which everyone can share information without being perceived as negative.
9. **Data Management Savvy**: Management and staff are data management savvy – understand each of the eleven foundational elements of data management and why data is critical for their company's success.
10. **Data Quality Processes**: Define quality requirements, measure quality, and proactively address quality issues.
11. **Robust Data Platform**: A robust data platform has been built and it supports the types of analytics required to make decisions, manage risk, and innovate.

More at: https://www.linkedin.com/pulse/11-characteristics-highly-mature-data-culture-jay-zaidi?trk=mp-author-card

Organizations with a mature data culture possess all the eleven characteristics.

A Real Life Business Problem

AlyData
Data Value
Unleashed



A Real World Business Problem

AlyData
Data Value
Unleashed

This scenario occurred at a Fortune 50 firm:

- Regulator requested major changes to disclosures post-2008 market crash to provide much granular data
- Per fiduciary responsibility company had to issue financial disclosures regularly
- Upon first issuance financial disclosures were deemed incorrect by major investors
- CEO on the hook to provide answers quickly
- Corporate Reputation, Legal threats, and Financial Risk impacts. Regulator paying close attention.

This particular scenario occurs at every organization, but takes different shapes.

Sequence of Events

AlyData
Data Value Unleashed

7:30 am - The CEO receives a call from one of the largest investors that disclosures aren't correct. CEO calls head of operations after getting off the phone with the investor to get answers. Is told that a post-mortem of the project will be conducted to determine root causes. Regulator is notified.

8:30 am - Head of operations contacts all directs and the key staff involved in development, testing, deployment of the project.

9:00 am - The project team is told of the major problem and asked to drop everything to conduct a post-mortem and identify root cause(s). All hands on deck.

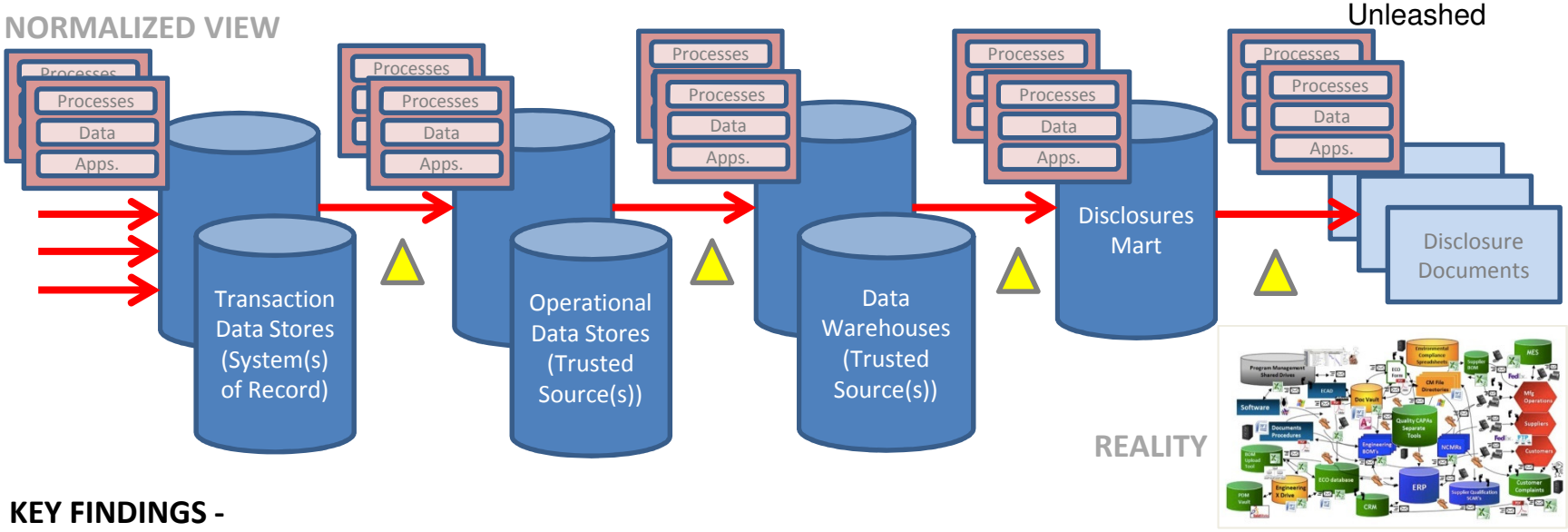
12:00 noon - A consulting company is asked to run this high profile engagement. They deploy resources onsite the next day, start developing a plan, holding status update meetings multiple times a day and send daily updates up the chain – to the CEO.

2:00 pm - A cross-functional team comprising of IT, Data, Operations, Business, and Management staff (over 200 people) drops everything to work on this emergency. They spend two weeks, over $4 Million to narrow down the root causes.

A sense of urgency is established and the right players are deployed.

Post-Mortem Findings

AlyData
Data Value Unleashed



KEY FINDINGS -

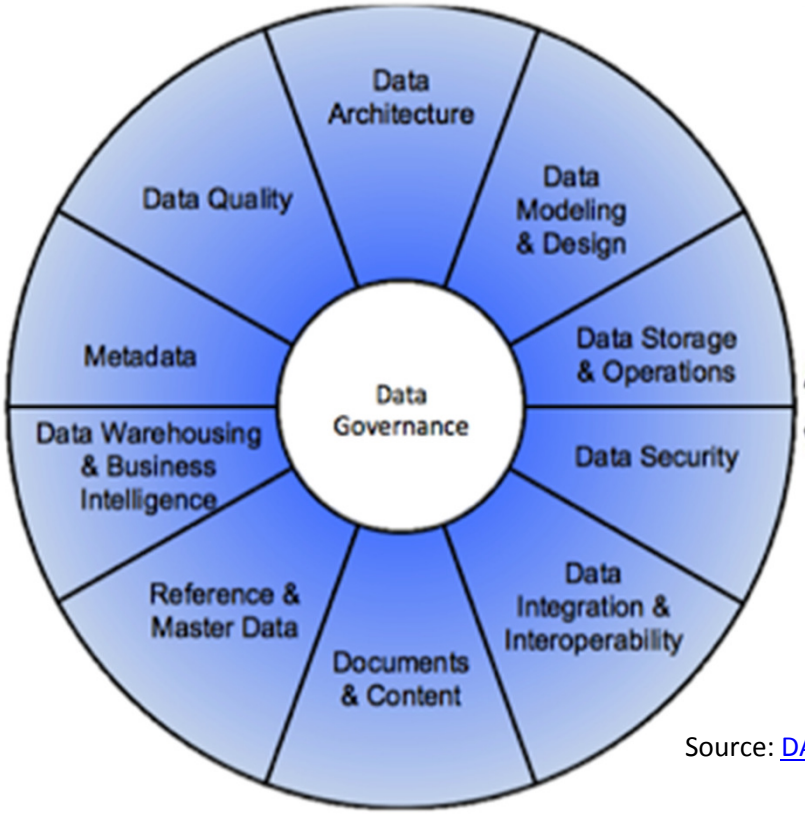
- Data was acquired via multiple channels – files directly loaded into data stores, web portals, and proprietary applications.
- Once acquired, data flowed through multiple data stores via file and message-based interfaces. Data was transformed throughout – from transaction data stores to operational stores, to the warehouses, and finally within the "disclosures" mart.
- The Information Supply Chain from the source to target systems wasn't documented and no one on the project team had any insight into the end-to-end big picture.
- The data lineage from source (transaction system of record) to target (Disclosures Mart) wasn't documented.
- The quality of all critical data elements that were disclosed to investors wasn't certified at the target. Disclosure data wasn't reconciled against the source (e.g., Transaction System(s) of Record). It was determined that the "head of operations" was advised to certify all critical data elements but decided to only certify a sub-set to save $80k. In hindsight, this was a very poor decision.
- Regulatory mandates and business policies changed over time and affected data requirements (e.g., mandatory vs optional etc.). The impacts of these changes weren't documented chronologically and neither were the corresponding data quality and business rules. The lack of documentation and corporate memory impacted the validation of data at the target (e.g. Disclosures Mart).
- The system(s) of record for the critical data elements to be disclosed were incorrectly assigned.
- Incorrect assumptions were made regarding all channels via which data was acquired. The post-mortem found that some new processes were introduced into the data flow that had the ability to update data.
- There was a lack of accountability for the quality of the data – since the data governance program hadn't adopted at the departmental level – in spite of much effort made by the Enterprise Data Governance team.
- An enterprise Data Dictionary was available, but hadn't been maintained. The same was true of Data Glossaries.
- The right Subject Matter Experts that had knowledge of the historical context of business policy changes and their impact on data requirements weren't consulted during the project.

The post mortem identified improvements across the Information Supply Chain.

Data Management 101

AlyData
Data Value
Unleashed

Data Management Knowledge Areas



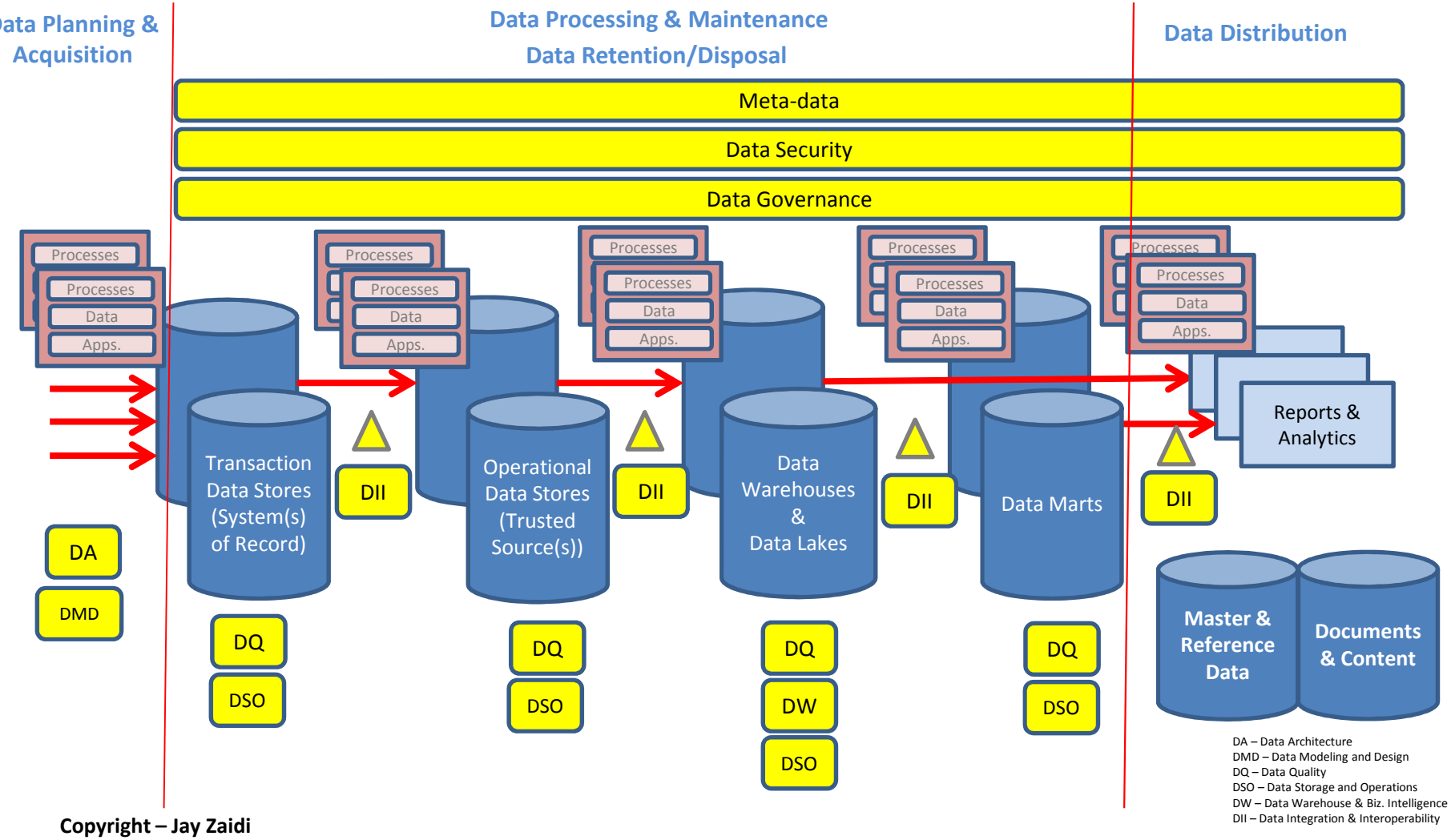
Source: DAMA

Data Management Environment Elements

DAMA International has defined the Knowledge Areas (KA) and Environmental Elements (EE) used in data management. KAs and EEs focus on the data management foundation and organizational factors respectively since both are required components of an effective data management program.

Connecting the Dots

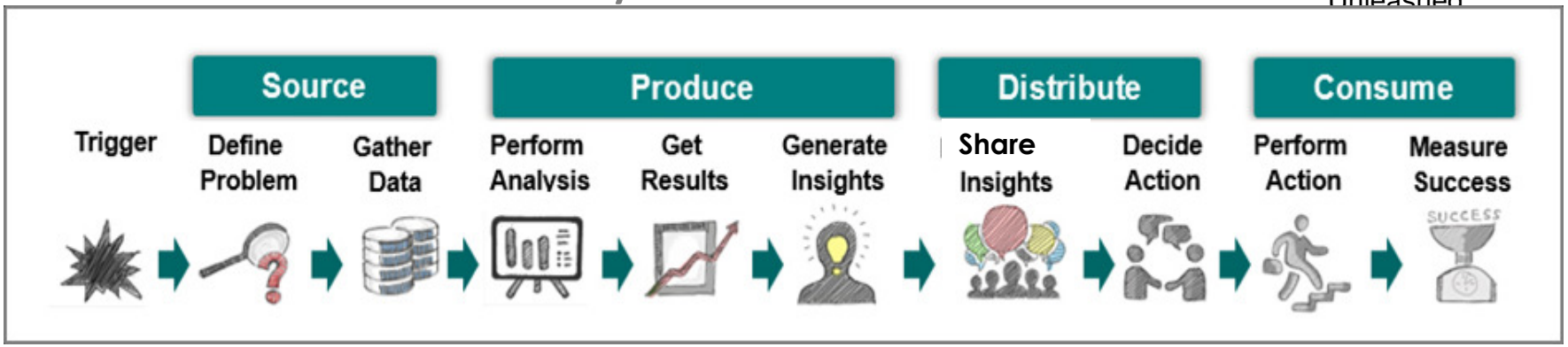
AlyData
Data Value Unleashed



The DAMA Knowledge Areas Applied to the Information Lifecycle.

10 Research Findings Across The Data Life Cycle

AlyData
Data Value Unleashed



85% of corporate data isn't visible ("Dark Data"[9]) and not used for generating value for the organization.

1/10/100 Rule of Data Quality: If fixing quality issues at source cost $1, then it is $10 to fix it on the next hop and so on[14].

80% of data processing effort spent on preparing data and only 20% on analysis[10].

6% to 10% of an organization's operating budget is spent on getting quality right[5].

Many organizations don't invest in a corporate data dictionary and glossaries that are available in electronic format and maintained to stay current[16].

Decentralization and democratization of data exposes organizations to significant risk due to higher probability of exposure of private data and data leaks[2,3].

Data producers typically don't certify data for quality and consistency and do not provide consumers with the necessary context related to the contents shared. Consumers have to expend significant resources to validate the data and to make sense of it.

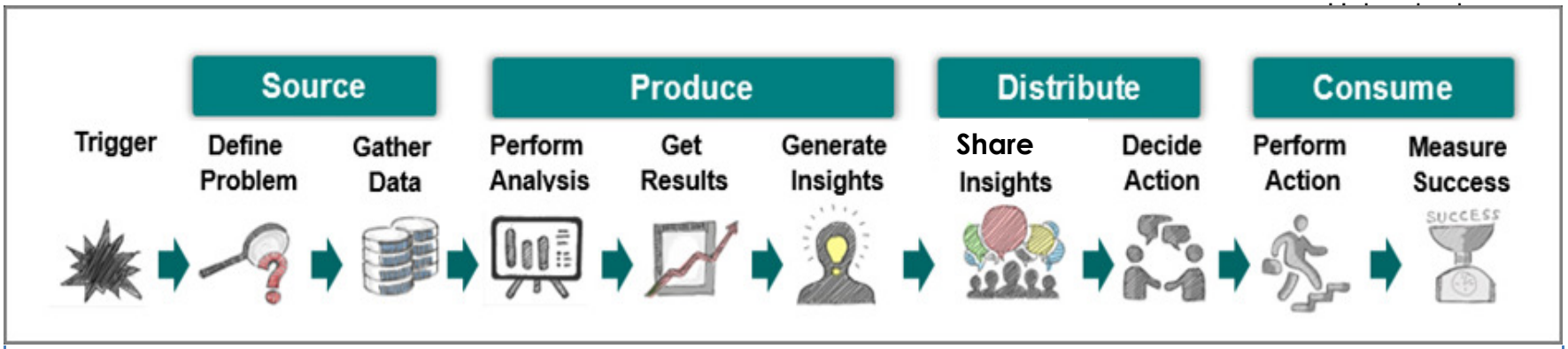
Most organizations don't have transparency into the lineage of data or its quality across the information supply chain[11].

Many organizations maintain data silos that create additional risk, data redundancy, and operating costs[15].

Data Governance across the enterprise can become rather cumbersome and add significant overhead, thus slowing down implementation. There is seldom alignment between IT Governance and Data Governance causing undue burden on various business and IT teams[12,13].

Best Practice Recommendations

AlyData
Data Value



Build **departmental** and **enterprise data inventory** and maintain it, so that all corporate data assets are visible and can generate value for the organization.

Identify and fix **data quality issues** at the source using data quality services or tools, to significantly reduce costs to downstream consumers.

Introduce **automation in data processing** to reduce effort spent on preparing data and increase time spent on analysis.

Profile data and address data quality issues proactively. Introduce re-usable rules and repeatable processes.

Capture "context" or "**metadata**" related to critical data elements and share this with reports/analytics. Invest in maintaining a **enterprise data dictionary and glossaries**.

Implement a robust **data governance program** to ensure that there is accountability for data across the data life cycle and proper data controls are incorporated.

Ensure that **sensitive data** is handled and distributed per federal government guidelines and policies.

Enforce **data sharing agreements** between data producers and consumers.

Data producers must **certify data** for quality and consistency and provide consumers with the necessary **context** related to the contents shared.

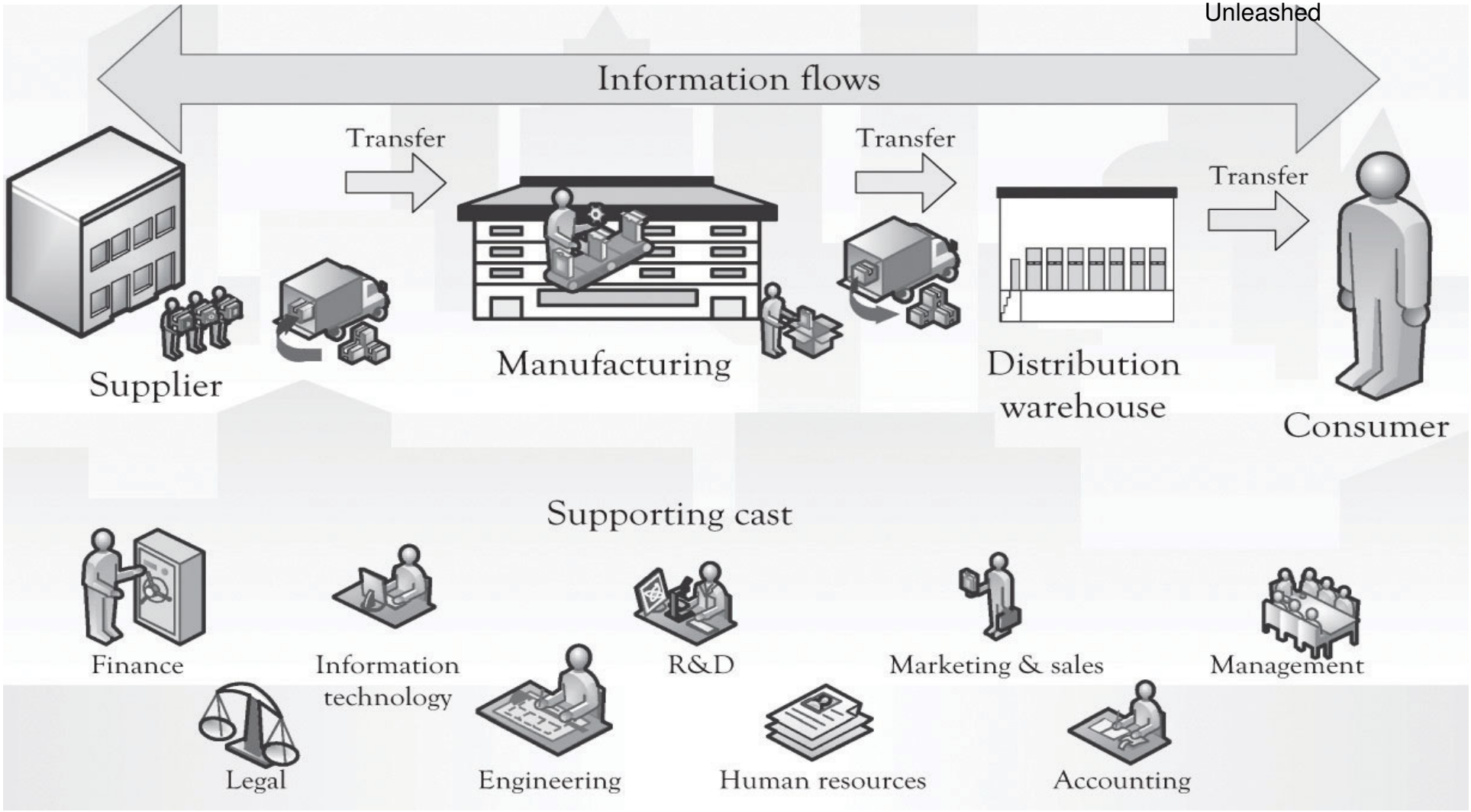
Capture **data lineage** as data flows from a System of Record to downstream consumers. This will increase auditability and ensure no material changes are introduced.

Introduce a **data culture** where transparency and collaboration across departments is encouraged and physical data silos are reduced.

A **federated Data Governance Model** is desirable. Data teams at Enterprise level (e.g. CDO) must not "boil the ocean". They must focus on enterprise critical data and let individual departments focus on their critical data. IT and Data Governance must be aligned.

Physical Supply Chain Concepts

AlyData
Data Value
Unleashed



Risk detection in a physical supply chain has three steps – analysis, validation, and assessment. This approach combines data analytics, supply chain expertise, and confirmation.

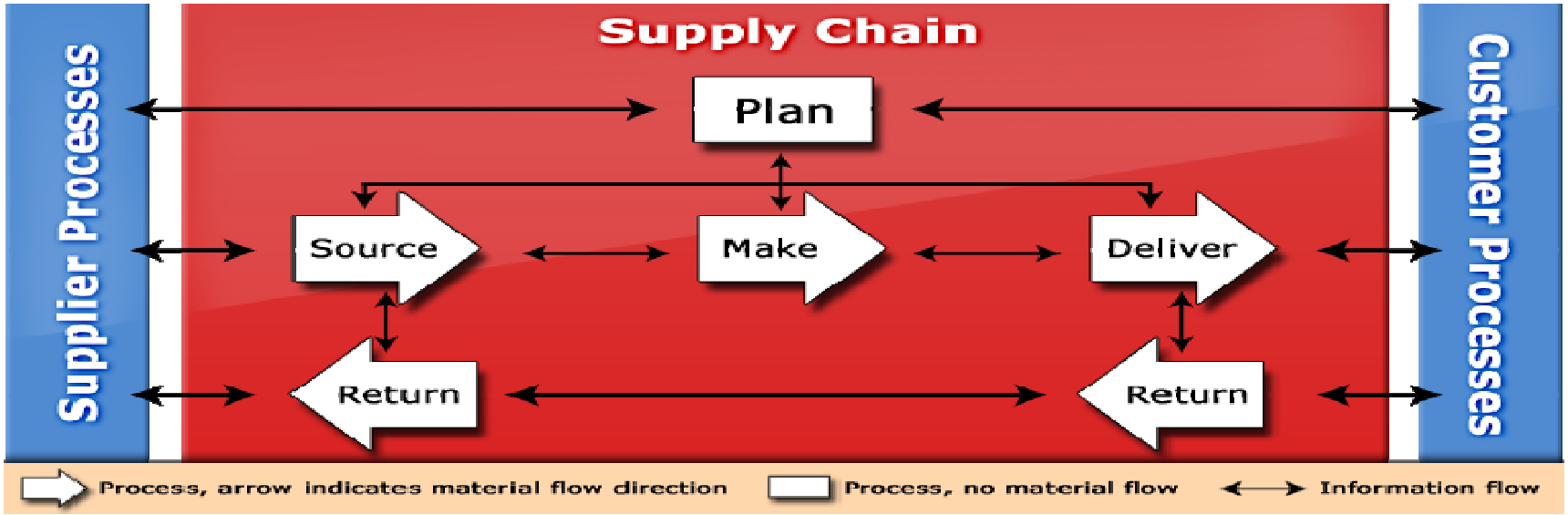
Data Management - Implemented Across Information Supply Chains

AlyData
Data Value Unleashed

Key Supply Chain Concepts:

- Data and Information flows between Suppliers and Consumer Processes
- There is a pre-defined agreement between supplier and consumer on delivery schedules, product quality, and acceptance criteria
- A product return/defect management process is agreed upon, implemented, and Service Level Agreements are tracked.

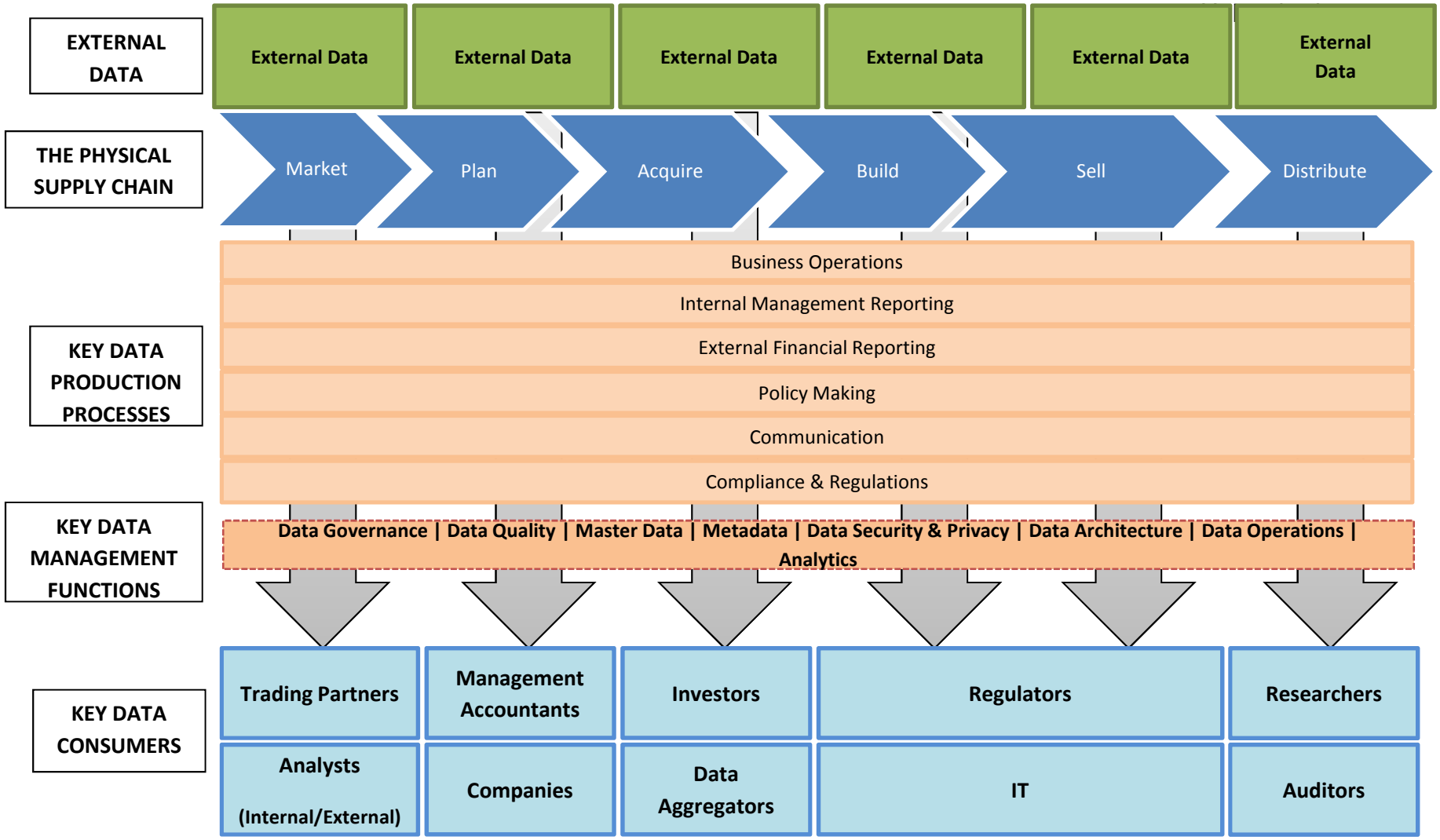
Supply Chain Processes and Information Flow



Data Management is implemented across Information Supply Chains. Therefore, all the key supply chain concepts such as Suppliers, Customers, Delivery Schedules, Product Quality, Acceptance Criteria, and SLAs apply to it.

The Information Supply Chain

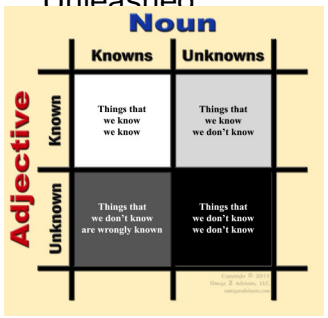
AlyData
Data Value



The enterprise Information Supply Chain (ISC) Clearly Shows That Data Flows Across Departments and External Parties. Managing the quality and privacy of the data as it flows through the ISC is paramount and so are the Insights gleaned from it – since they are used for decision-making and reporting to consumers.

The Data Value Pyramid - Built On A Robust Data Management Platform

AlyData
Data Value Unleashed



DATA MATURITY

Data Science

(Statistics, Machine Learning, & Artificial Intelligence)

- Uncover hidden patterns in data
- Understand complex behaviors, trends, and inferences
- Apply advanced algorithms and machine learning to solve complex problems

Analytics

(Descriptive, Diagnostic, Predictive, & Prescriptive)

- What happened?
- Why did it happen?
- What is likely to happen?
- What should I do about it?

Reporting

(Transactional, Operational, Ad-Hoc, & Self Service)

- Static view of transactions
- Static view of operations
- On demand reports/dashboards

Data Management

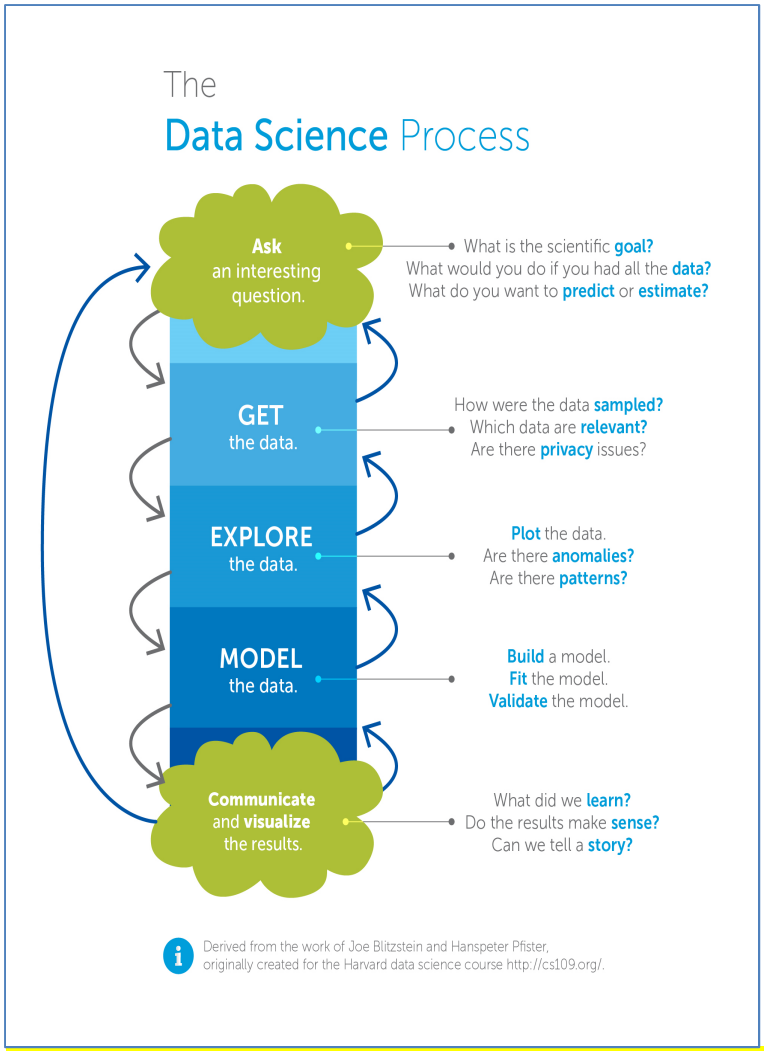
(Data Modeling, Data Architecture, Data Integration, Data Governance, Data Quality, Data Storage & Operations, Meta-data, Master & Reference Data, Data Warehouse & BI, Data Security)

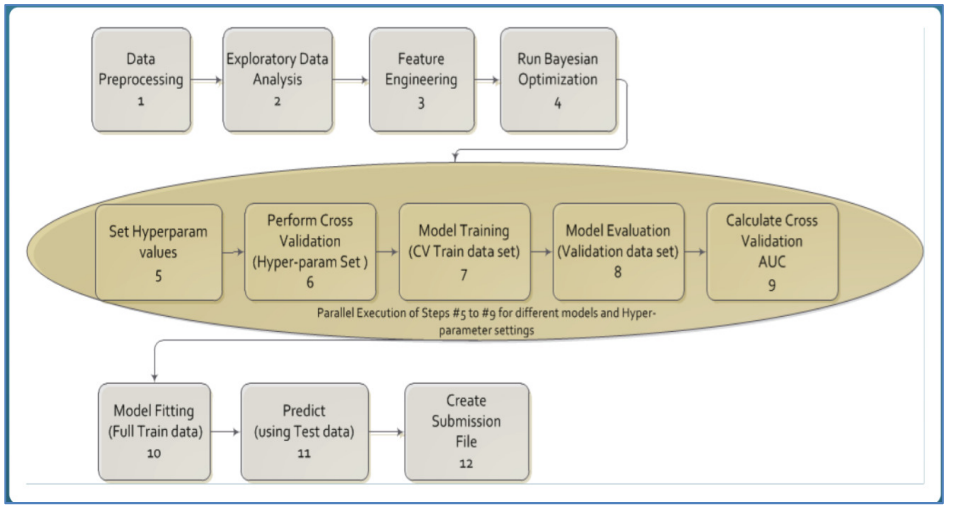
- Semantics
- Accountability
- Inventory
- Quality
- Process

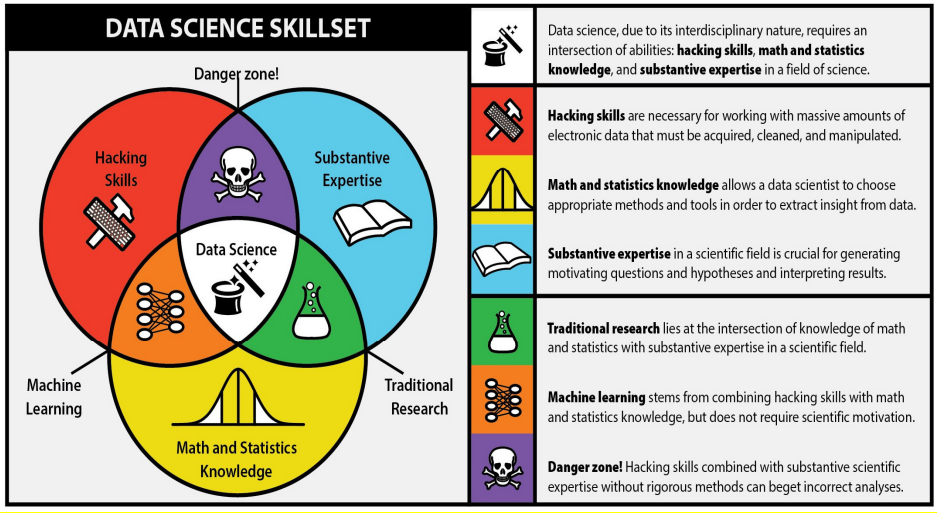
Data isn't valuable by itself, but it's the insights gleaned from it that are most important. Data Management is the **foundation** of the Data Value Pyramid.

Data Science Process

AlyData
Data Value Unleashed







Data science is an interdisciplinary field that applies scientific processes and systematic study to extract knowledge or insights from data in various forms, either structured or unstructured.

Creative Ways In Which Some Companies are Using Data

AlyData
Data Value Unleashed

NETFLIX

Using its data to help choose which content to license. It also uses its data to recommend the right content to the right users. But what's really noteworthy is Netflix's use of data to figure out which canceled shows to bet on. Its **recommendation algorithms** work to increase the audience by introducing viewers to shows.

Walmart

Timely analysis of real-time data is seen as key to driving business performance – as Walmart Senior Statistical Analyst Naveen Peddamail runs Walmart's Data Cafe and tells me: "If you can't get insights until you've analyzed your sales for a week or a month, then you've lost sales within that time. **Our goal is always to get information to our business partners as fast as we can, so they can take action and cut down the turnaround time. It is proactive and reactive analytics**."

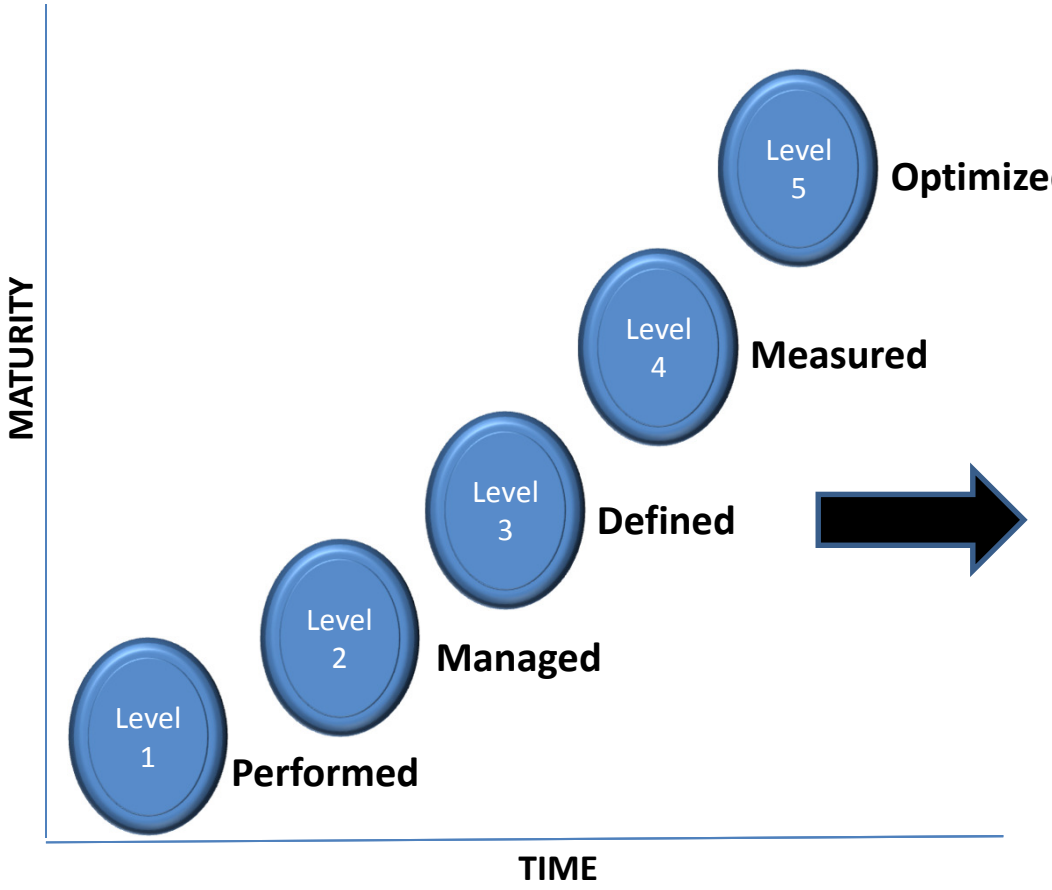
Rolls-Royce

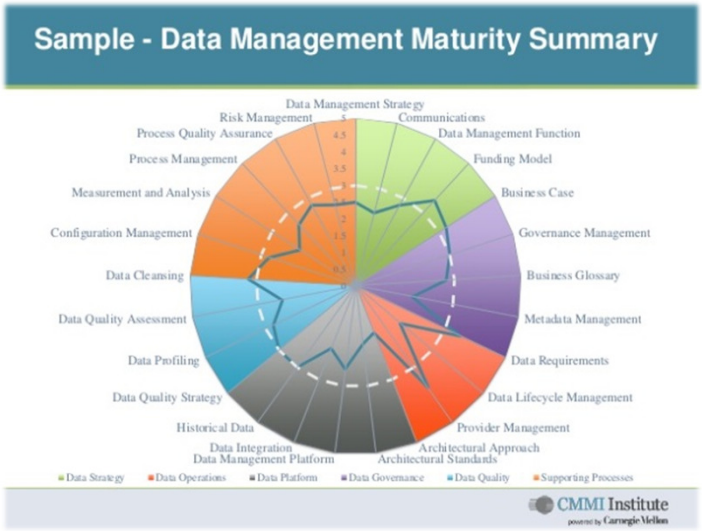
Big Data processes to use in three key areas of their operations: design, manufacture and after-sales support. Paul Stein, the company's chief scientific officer, says: "We have huge clusters of high-power computing which are used in the design process. We generate tens of terabytes of data on each simulation of one of our jet engines. We then have to use some pretty sophisticated computer techniques to look into that massive dataset and **visualize whether that particular product we've designed is good or bad**."

There are numerous examples of data-driven companies doing amazing things with data. Use these as inspirations and guides.

Data Management Maturity Model (DMM)

AlyData
Data Value
Unleashed



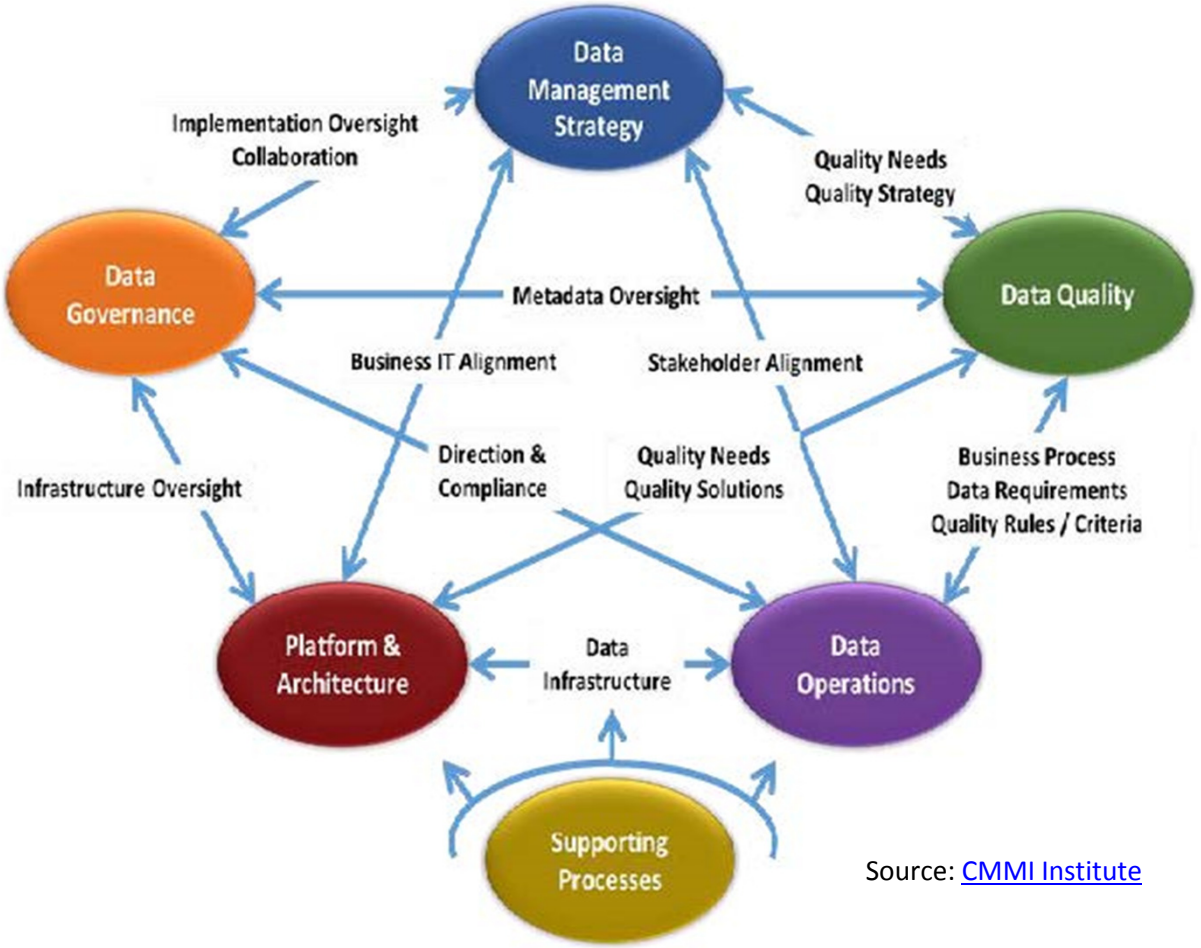


Source: CMMI Institute

DAMA Partnered with CMMI Institute To Develop the Data Management Maturity Model. DMM provides a methodology to measure the Maturity Of Data Management and supporting processes. Consumer Response is using DMM to assess its data management maturity and to proactively address gaps between the current and desired maturity levels.

DMM Structure and Data Management Strategy

AlyData
Data Value Unleashed



Source: CMMI Institute

Data Management Strategy
- Data Management Strategy
- Communications
- Data Management Function
- Business Case
- Funding

Data Governance
- Governance Management
- Business Glossary
- Metadata Management

Data Quality
- Data Quality Strategy
- Data Profiling
- Data Quality Assessment
- Data Cleansing

Data Operations
- Data Requirements Definition
- Data Lifecycle Management
- Provider Management

Platform & Architecture
- Architectural Approach
- Architectural Standards
- Data Management Platform
- Data Integration
- Historical Data, Archiving and Retention

Supporting Processes
- Measurement and Analysis
- Process Management
- Process Quality Assurance
- Risk Management
- Configuration Management

The DMM Structure Is Built On 5 Areas and Processes that Support an Organization's Data Management Strategy.

Implementation Pointers – Here's What You Can do

AlyData
Data Value Unleashed

Meet with business leaders to seek big problems they may be facing that have a major data component. Develop solutions and proposals to them.

Sense an opportunity based on your personal observations or what other data-driven companies are doing and propose them to your management team.

Imagine: Do some blue sky thinking to imagine what your company, customers, products, vendors could benefit from and make it happen.

Time-to-value: Data practitioners must be value-driven and agile in their approach, so that they can deliver solutions quickly.

Make things happen by being proactive.

Opportunities for Process Optimization and Cost Reduction

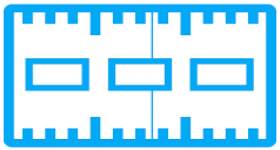
AlyData
Data Value Unleashed

Data Quality: At least 6% to 10% of IT operating budget wasted due to re-work and inefficient processing.

Metadata (context): Inability to find data, understand data semantics, and data related rules results in massive inefficiency.

Data Governance: Lacking data accountability and governance processes introduces compliance challenges.

Dark Data: 85% of data acquired isn't used for anything of value.



Data Wrangling: 70% to 80% of data processing time and cost is associated with data wrangling.

Data Lineage: Hard to define automatically, even though its critical from audit, compliance, and governance perspective incurs auditor wrath.

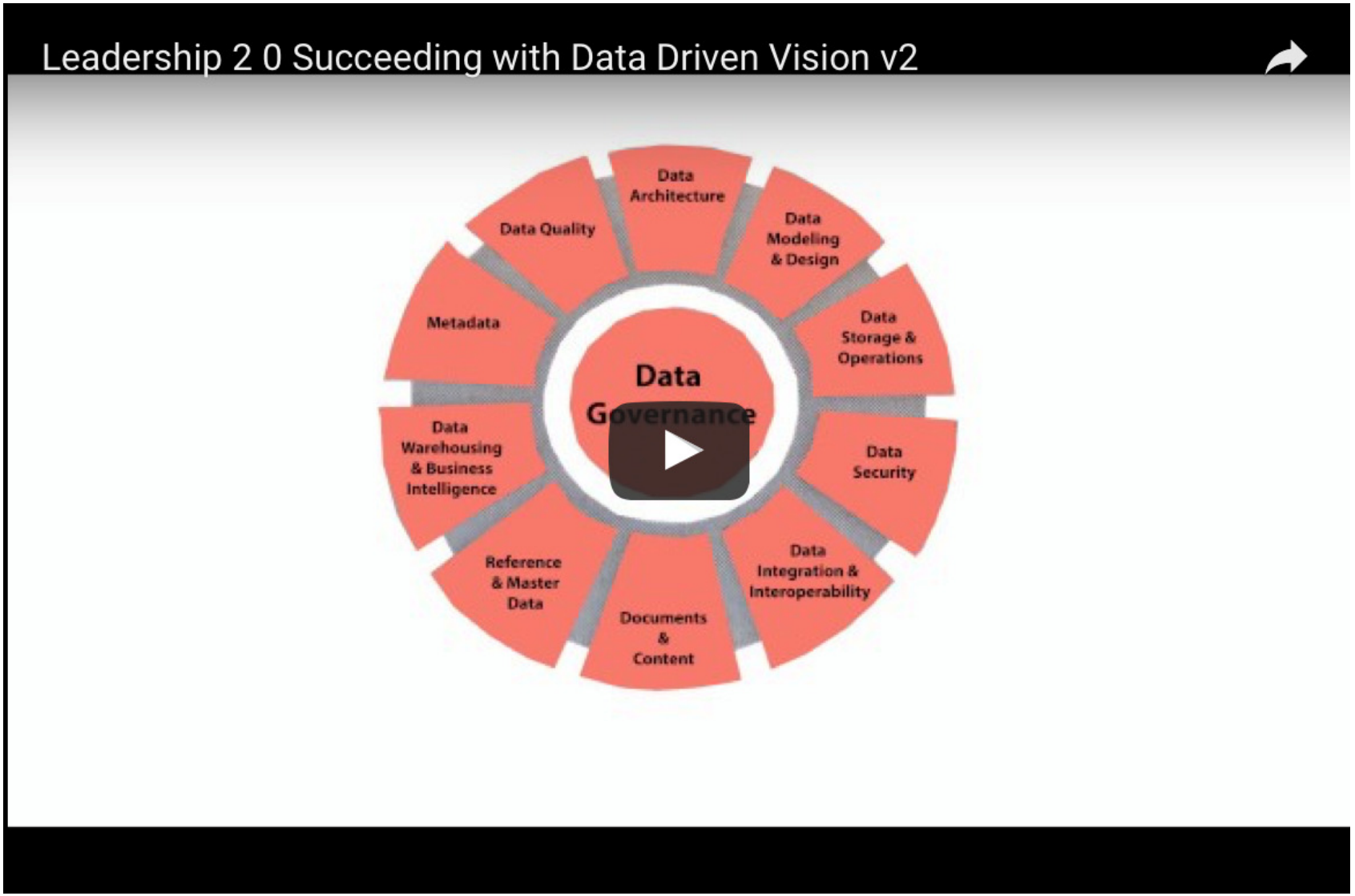
Data Security and Privacy: Access controls, tagging sensitive data, lack of oversight exposes company to risks.

Analytics and Deep Learning: Deep insights from data to uncover unknown/unknowns.

Companies can reduce operational costs by 6 to 10%, improve data preparation time by 30 to 50%, and improve staff morale by strategically investing in data management.

We Come Full Circle – Leadership 2.0

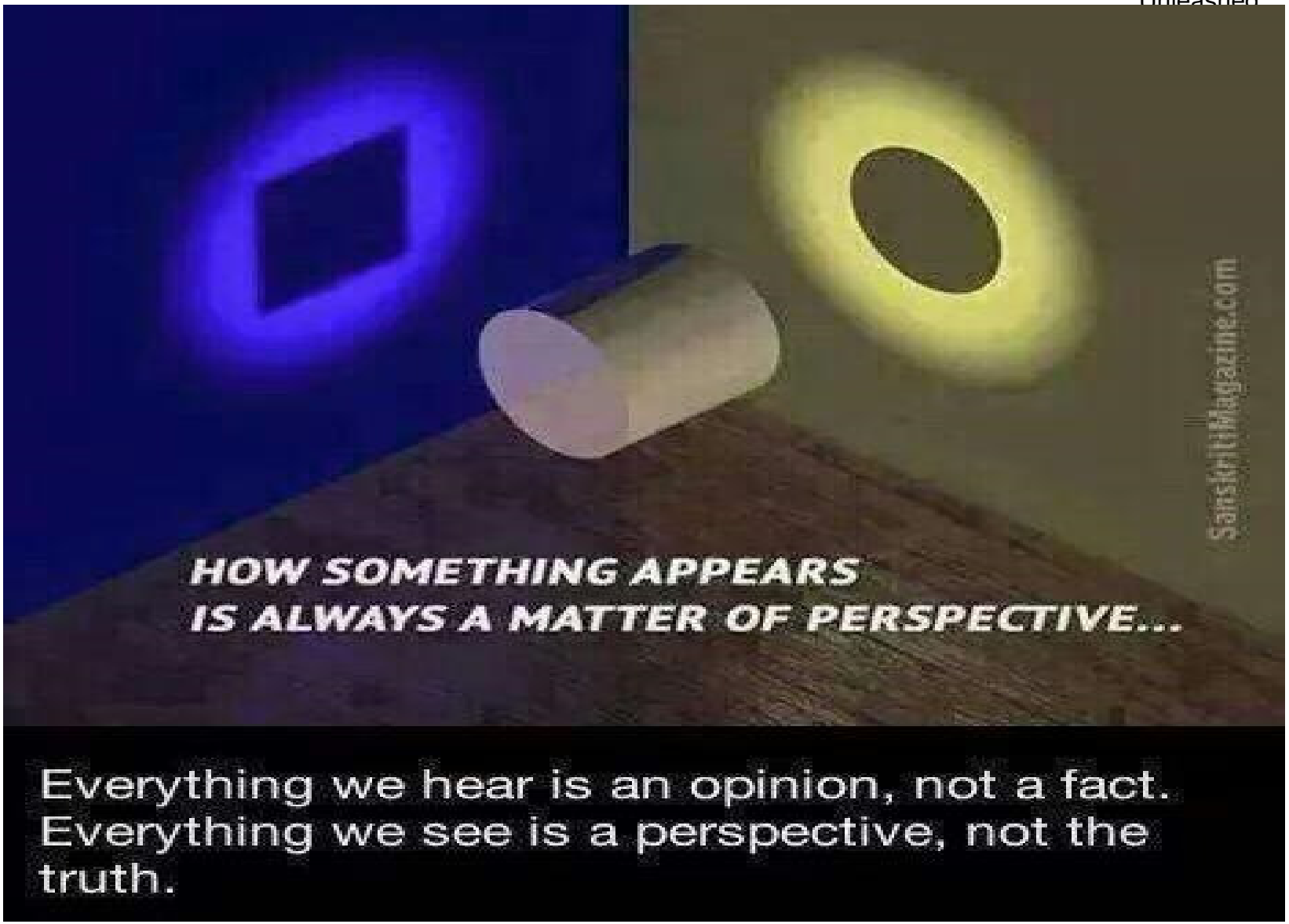
AlyData
Data Value
Unleashed

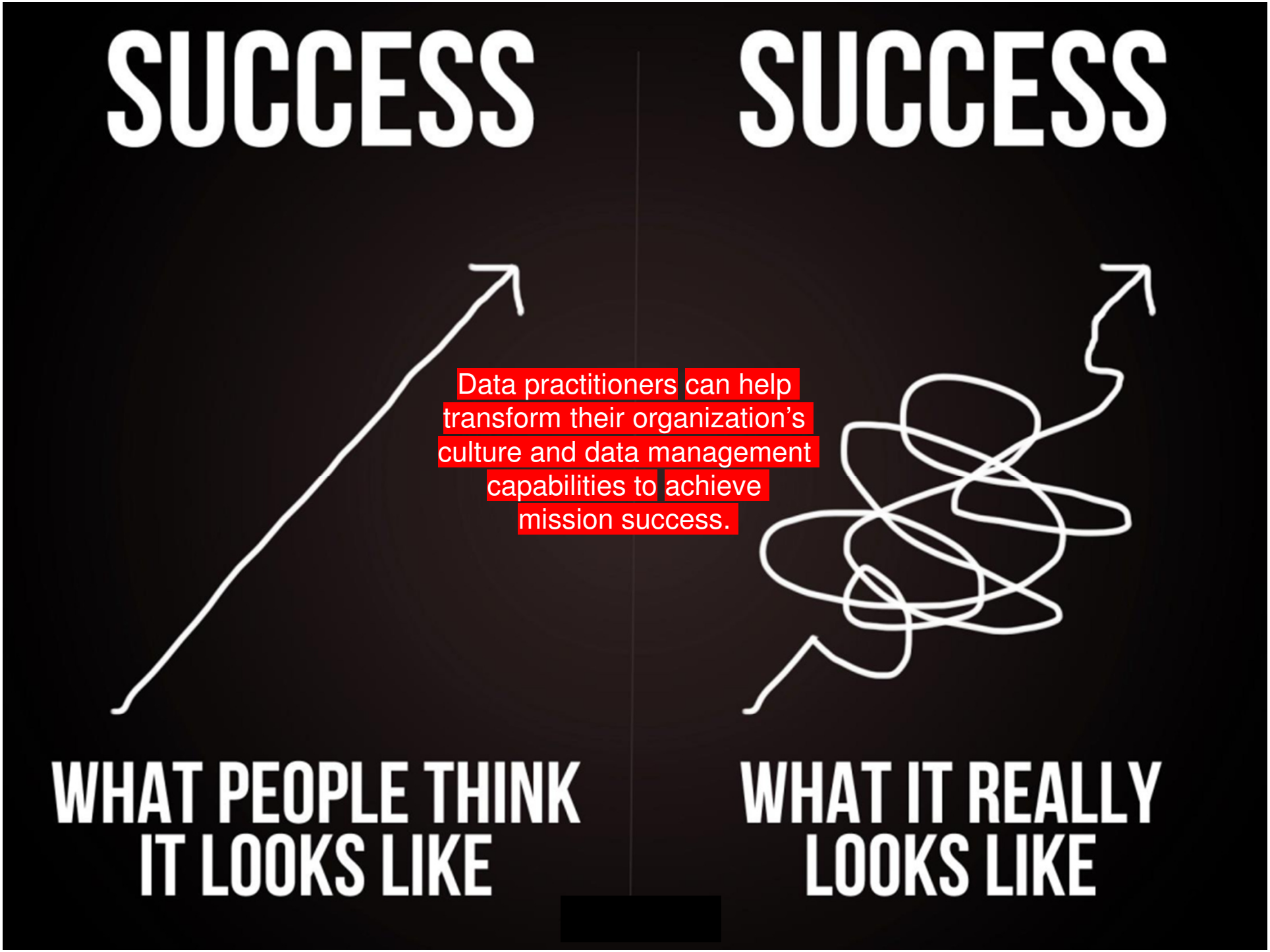


6 Key Takeaways

AlyData
Data Value Unleashed

1. **Mission Alignment:** Data practitioners must focus on value and risk - "what data does" and not on "what data is". They must identify opportunities for value creation and risk mitigation and address them.
2. **"Age of Data" – Leadership 2.0:** Due to the 3-D's of data a new Leadership 2.0 model has emerged and it is focused on People, Process, Technology, and Data.
3. **Data Culture**: We are living in the "Age of Data" and companies that wish to succeed must adopt a data culture. Data and Analytics are "strategic assets" for CEO's and support their six top priorities. There are 11 characteristics of highly mature data cultures and all organizations must strive to achieve them.
4. **Strengthen the Foundation**: Data management forms the foundation of the Data Value Pyramid. Therefore, organizations must mature it to ensure a strong foundation. A maturity assessment must be performed regularly to determine if the organization's data management areas are moving up the maturity curve.
5. **Apply Supply Chain Principles**: Data management fits a supply chain pattern. So, all supply chain concepts and methodologies are applicable. Practitioners must educate and raise awareness and implement solutions.
6. **Operational Efficiency**: Organizations can reduce operational budgets, optimize processes, and reduce risk by investing in data management and Analytics.

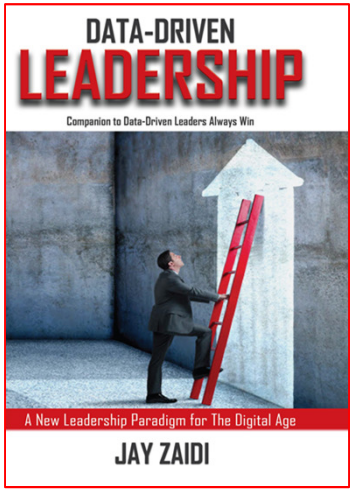

HOW SOMETHING APPEARS
IS ALWAYS A MATTER OF PERSPECTIVE...
SanskritiMagazine.com
Everything we hear is an opinion, not a fact.
Everything we see is a perspective, not the truth.

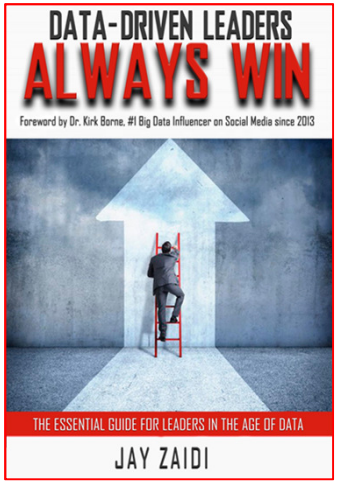

SUCCESS
SUCCESS
Data practitioners can help transform their organization's culture and data management capabilities to achieve mission success.
WHAT PEOPLE THINK IT LOOKS LIKE
WHAT IT REALLY LOOKS LIKE

About Me

AlyData
Data Value
Unleashed

- Founded AlyData in 2014. We specialize in Data Management and Data Science. Our mission is to transform organizations, by helping them innovate and gain a competitive advantage by unleashing value from their data and information assets.
- 13 Years at Fannie Mae. Last 5 years was a direct report to the CDO.
- Have authored 2 books and over 80 articles on data management on LinkedIn.
- Email - jayzaidi@alydata.com
- LinkedIn Profile - https://www.linkedin.com/in/javedzaidi
- Twitter - @jayzaidi





APPENDIX

5 Pillars of the New Business Model

AlyData
Data Value Unleashed

1. **Variety and Decentralization**: Social, Mobile, Analytics, and Cloud (SMAC) drive operations

2. **Better Insights**: Near real time insights for decision making, risk management, and to gain competitive advantage

3. **Agility**: Transformation of the operating model from SDLC to Agile and introduction of automated processes

4. **Transparency**: Sharing economy requires a sharing culture. Change in team dynamics to become more transparent and share data and algorithms.

5. **Innovation**: Innovate using data, people, algorithms, and process. New areas such as artificial intelligence (AI), deep learning, intelligent conversation engines, speech recognition, and image and pattern recognition. 9 The Intellectual Capital of this new world is Algorithms, Data, and People.

Businesses have to transform themselves and introduce a Data Culture.

Organizational Readiness Assessment

AlyData
Data Value
Unleashed

- Organizational readiness for change is a multi-level, multi-faceted construct. As an organization-level construct, readiness for change refers to organizational members' shared resolve to implement a change (change commitment) and shared belief in their collective capability to do so (change efficacy).

- We will measure the effectiveness (i.e. Lagging, Basic, Advanced, or Leading) across the following groups:
  - Staff buy-in
  - Data collector buy in
  - Leadership buy in
  - People resources
  - Data Use Policy
  - Intervenor buy in
  - Funder buy in

Developing a “data culture” within an organization requires data management training and proactive change management across departments. An Organizational Readiness Assessment (if conducted) helps develop organizational change management recommendations to ensure a data management program’s success.

References

AlyData
Data Value Unleashed

1. Gartner: Data Growth Is a Critical Enterprise Challenge, CIO Insight
2. The Big Data Universe Is Beginning To Explode, CSC Infographic, 2012 (1/3rd of data will live in or flow through the Cloud by 2020)
3. 8 Surprising Statistics About Insider Threats, Dark Reading, Aug. 2016
4. Big Data eBook Published by Information Corporation: https://www.informatica.com/content/.../big_data_government_ebook_2340.pdf
5. Data-driven Leaders Always Win: An Essential Guide For Leaders In The Age of Data by Jay Zaidi
6. DAMA – Data Management Book of Knowledge, First Edition, 2009
7. DAMA – The DAMA Dictionary of Data Management
8. DAMA – Certified Data Management Professional Designation
9. IBM Cognitive Colloquium Spotlights Uncovering Dark Data
10. Cleaning Big Data: Most Time-Consuming, Least Enjoyable Data Science Task, Forbes, March 2016
11. The Information Supply Chain: Achieving Business Objectives by Enhancing Critical Business Processes, George Marinos, Information Management, April 2005
12. Is Your Data Governance Program in Trouble?, Jay Zaidi, TDAN, April 2012
13. Data Governance Demystified – Lessons from the Trenches, Jay Zaidi, Dataversity, Jan. 2012
14. Why Data Should Be a Business Asset: The 1-10-100 Rule, Martin Doyle, SmartDataCollection, Aug. 2014
15. Siloed Data is Killing Your Business Analytics, Cyfe Blog, Dec. 2015
16. What is a Data Dictionary and What Role Does it Play in Semantic Interoperability?, Jason Wolfson, Aug. 2014