



# Data Quality Best Practices

Donna Burbank and Nigel Turner  
Global Data Strategy, Ltd.

August 28<sup>th</sup>, 2025



# Data Quality Best Practices

## **Gino Fortunato**

Solution Consultant, Reltio

[gino.fortunato@reltio.com](mailto:gino.fortunato@reltio.com)

<https://www.linkedin.com/in/ginof/>

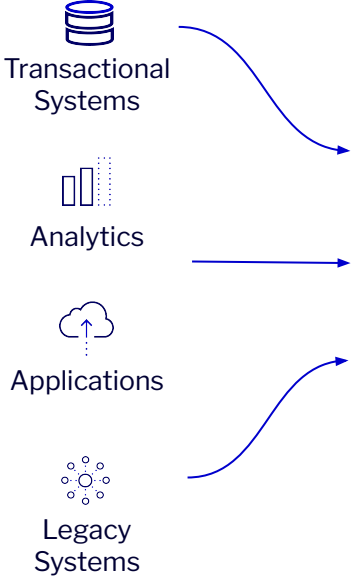
**RELTIO**

# Tackling Data Quality Problems: Five Principles for Sustainable Success

1. Tactical Projects Can't Solve Systemic Data Issues. You Need a Holistic, Proactive, and Continuous Approach
2. People and Governance Are Critical to Trust
3. Most Data Quality Problems Are Cross-Functional, Require Collaboration Under a Unified, Business-Oriented Understanding
4. Sustainable Improvements Require a Data Backbone with Automation
5. Data Quality Isn't Just About Risk—It's About Readiness

# Reltio's approach

## Data Sources



## Reltio Data Cloud™

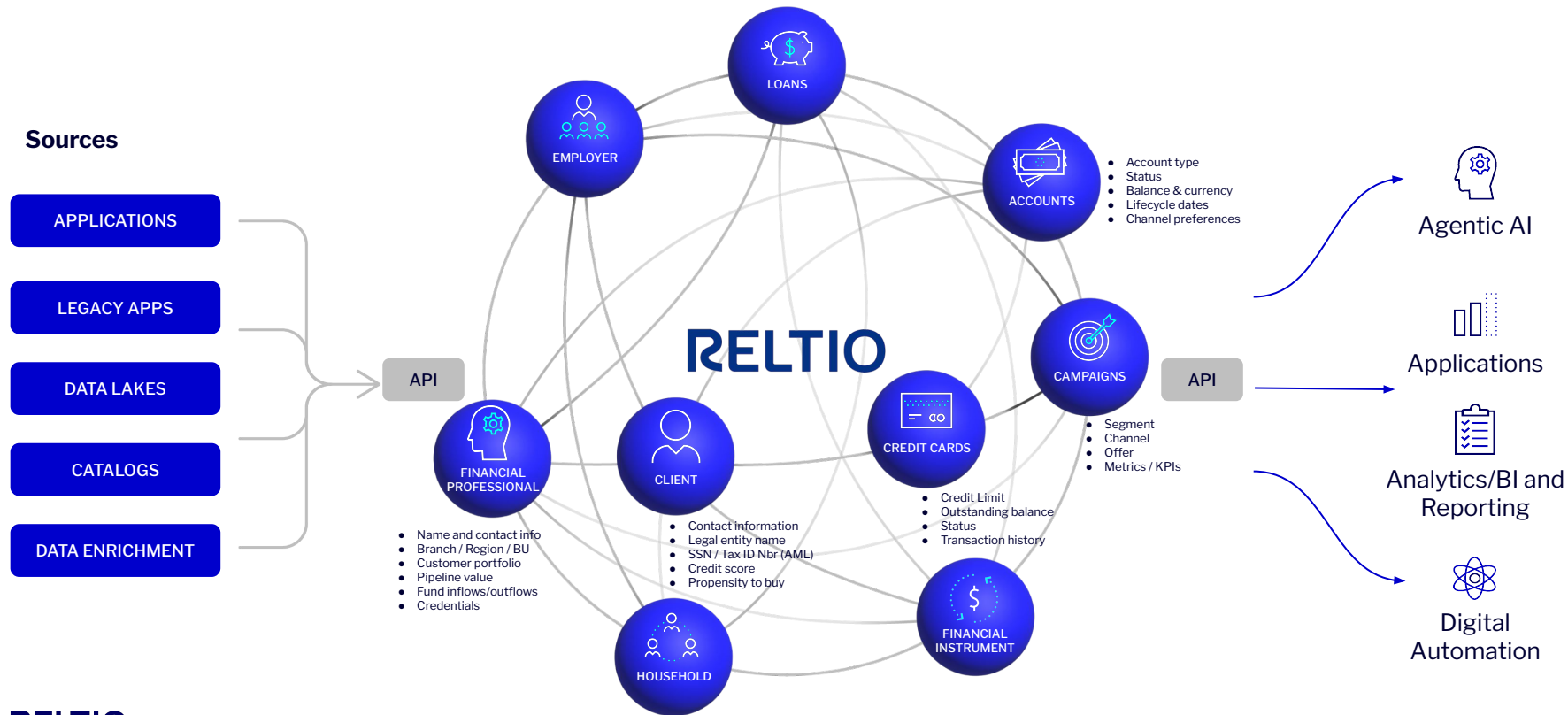
Context-rich 360° profiles—at agentic speed



## Data Consumers

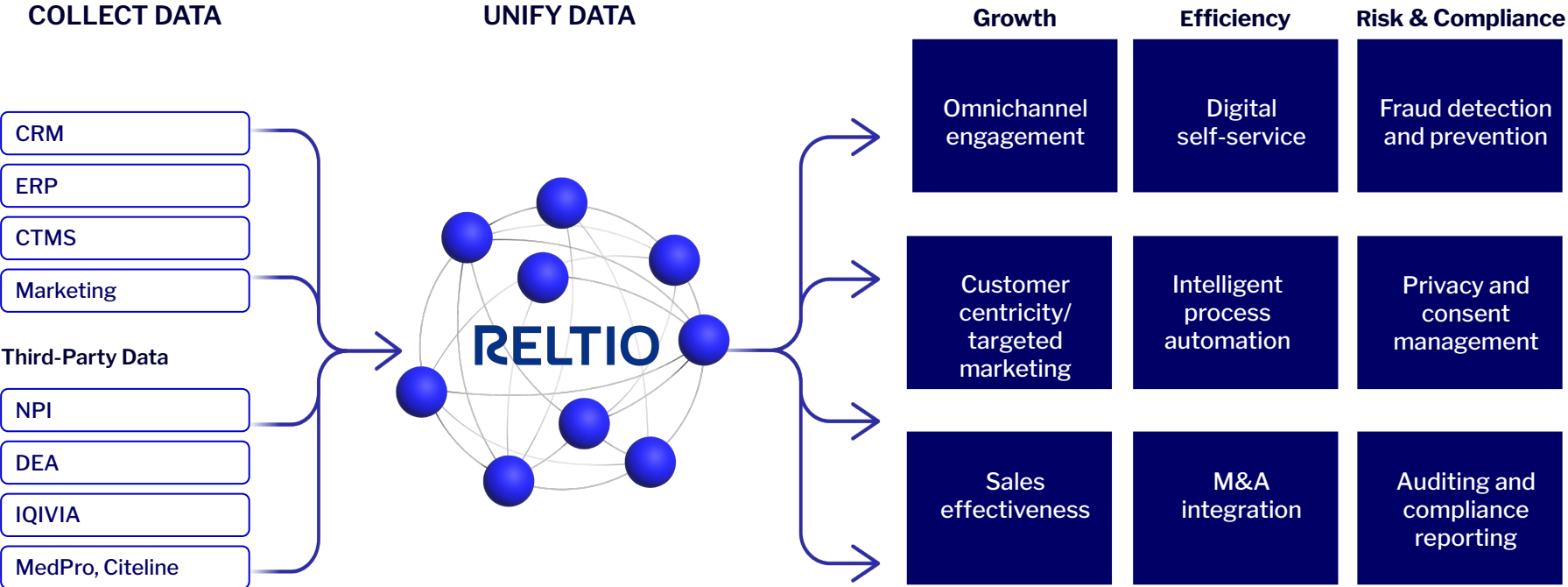


# Comprehensive, high-quality data for the business

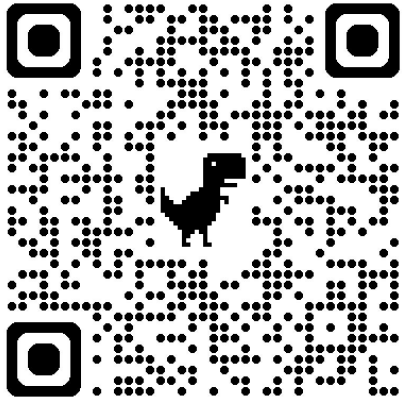


RELTIO

# Real-time, high-quality data for operational and analytical use cases



Learn More at [Reltio.com](https://relio.com)



RELTIO

Thank you!

# Donna Burbank



Donna is a recognized industry expert in data management with over 25 years of experience in data strategy, data governance, data modeling, metadata management, and enterprise architecture. Her background is multi-faceted across consulting, product development, product management, brand strategy, marketing, and business leadership.

She is currently the Managing Director at Global Data Strategy, Ltd., an international data management consulting company that specializes in the alignment of business drivers with data-centric technology.

In past roles, she has served in key brand strategy and product management roles for several of the leading data management products in the market.

As an active contributor to the data management community, she is a long time DAMA member, contributor to the DMBOK 2.0 and 3.0, Past President and Advisor to the DAMA Rocky Mountain chapter, and was awarded the Excellence in Data Management Award from DAMA International.

She has worked with dozens of Fortune 500 companies worldwide in the Americas, Europe, Asia, and Africa and speaks regularly at industry conferences. She has co-authored several books and is a regular contributor to industry publications. She can be reached at [donna.burbank@globaldatastrategy.com](mailto:donna.burbank@globaldatastrategy.com)  
Donna is based in Boulder, Colorado, US.





## Nigel Turner

- Worked in Information Management (IM) and related areas for over 30 years. Experience has embraced Data Governance, Information Strategy, Data Quality, Master Data Management & Business Intelligence.
- Principal Consultant, EMEA, for Global Data Strategy, Ltd.

- Spent much of his career in British Telecommunications Group (BT) where he led a series of enterprise-wide IM & data governance initiatives.
- Also been VP of Information Management Strategy at Harte Hanks Trillium Software, and Principal Consultant at FromHereOn and IPL.



- Nigel is very active in professional Data Management organizations and is an elected Data Management Association (DAMA) UK Committee member.
- He was the joint winner of DAMA International's 2015 Community Award for the work he initiated and led in setting up a mentoring scheme in the UK where experienced DAMA professionals coach and support newer data management professionals.
- Nigel is based in Cardiff, Wales, UK and can be reached at [nigel.turner@globaldatastrategy.com](mailto:nigel.turner@globaldatastrategy.com)

Follow on Twitter @NigelTurner8

# DATAVERSITY Data Architecture Strategies

## This Year's Lineup

- **January** Trends in Data Architecture
- **February** Building a Data Strategy - Practical Steps for Aligning with Business Goals
- **March** Building the Right Architecture for Analytics & Reporting
- **April** Data Architect vs. Data Engineer vs. Data Scientist – Making Sense of Roles in Today's Data-Centric Organization
- **May** Master Data Management - Aligning Data, Process, and Governance
- **June** Where Data Models Fit in Today's Modern Data Architecture
- **July** Data Architecture vs. Enterprise Architecture
- **August** Data Quality Best Practices (with guest Nigel Turner)
- **September** Modern Data Architecture: Practical Options for Today's Data-Driven Organization
- **October** Best Practices in Metadata Management
- **December** The Business Value of Data Modeling

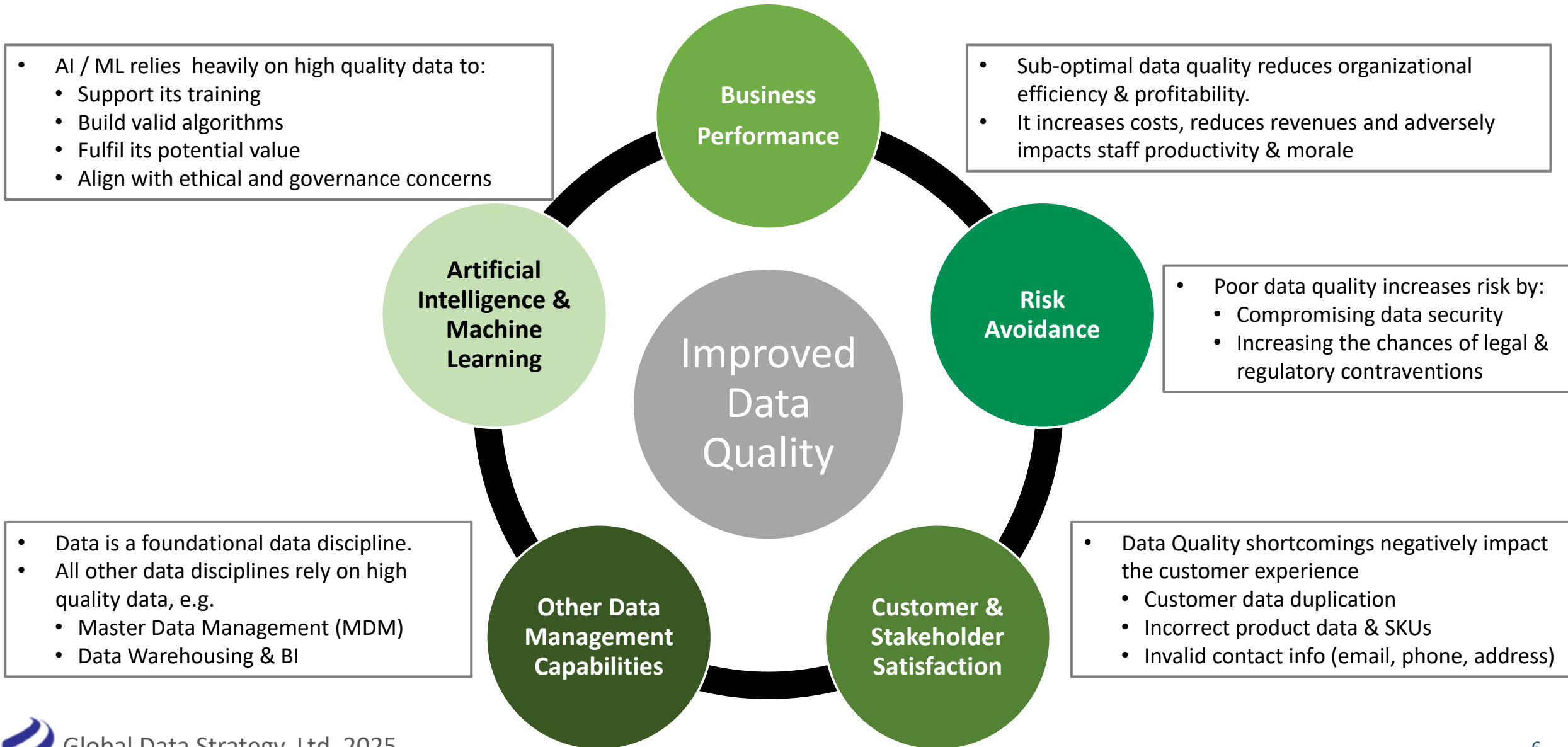


# What We'll Cover Today

- Data quality remains a major challenge for organizations, and has a **direct impact on business operations**
- Tackling Data Quality problems requires more than a series of tactical, one-off improvement projects.
- Addressing data quality issues **requires a holistic architectural approach combining people, process, and technology.**
- Ethical AI depends on a strong data quality foundation.
- This webinar provides practical ways to tackle data quality to support analytic & operational activities.

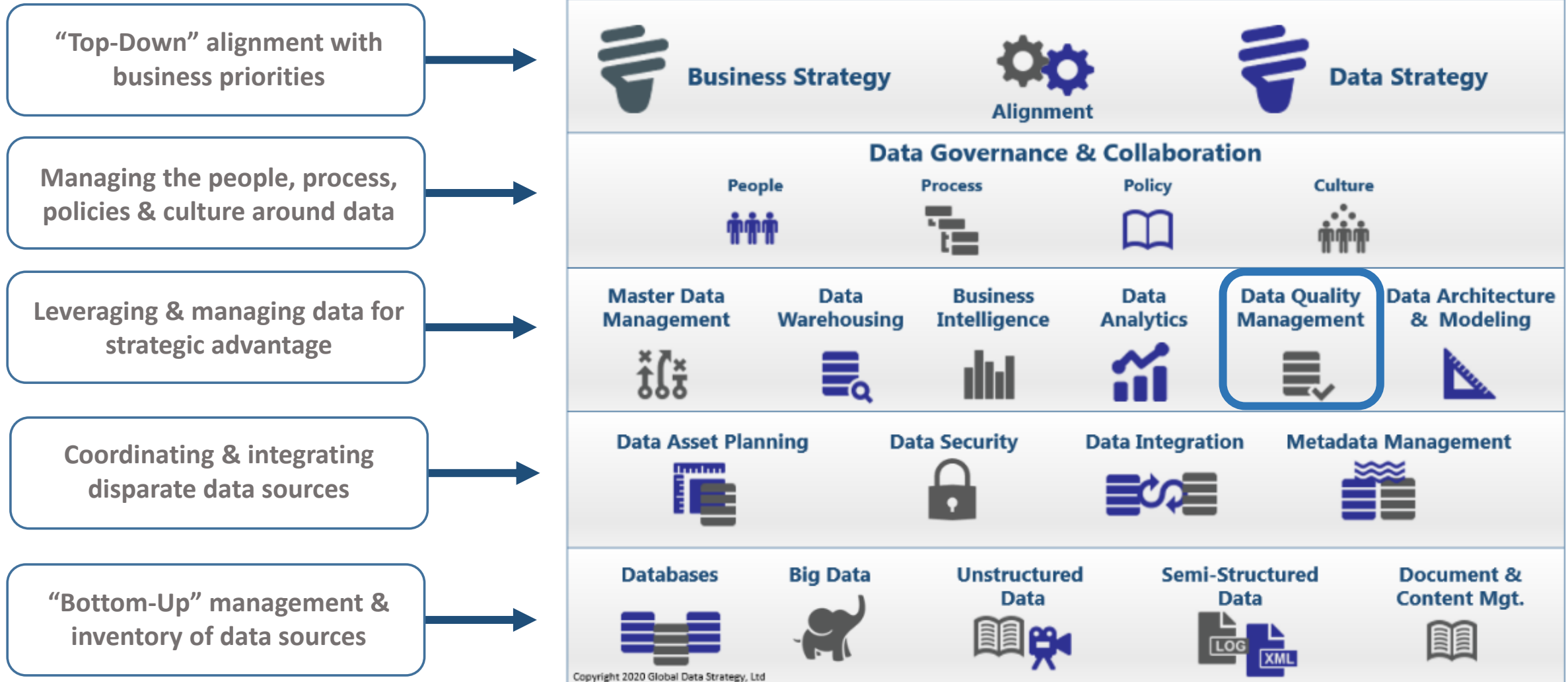


# Drivers For Improved Data Quality

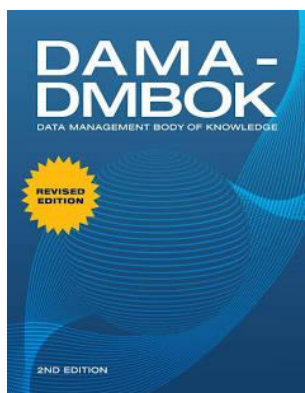
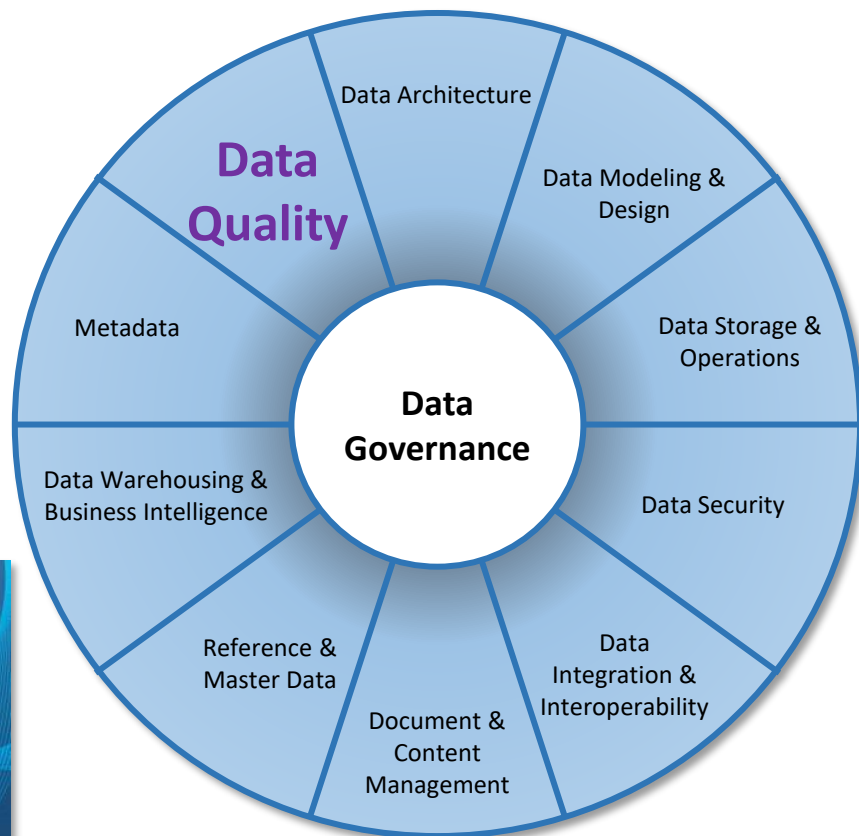


# Data Quality is Part of a Wider Data Strategy

A Successful Data Strategy links Business Goals with Technology Solutions



# Data Quality: a Foundational Data Discipline



Discipline	Example Relationships
<b>Data Quality</b>	Data Quality requires Data Governance to drive & sustain improvement so it becomes BaU
<b>Data Architecture</b>	Designs the structural framework for the management of data across platforms & systems
<b>Data Modeling &amp; Design</b>	Identifies business definitions, entities & attributes to focus Data Governance activities
<b>Data Storage &amp; Operations</b>	Lack of business accountability for data can impact Data Storage & Operations efficiency & reliability
<b>Data Security</b>	Lack of business accountability makes data less secure & more open to fraud
<b>Data Integration &amp; Interoperability</b>	Depends on defined & consistent data formats & content
<b>Documents &amp; Content</b>	Good Data Governance practices support Documents & Content, e.g. content ownership, version control, tagging, taxonomies
<b>Reference &amp; Master Data</b>	Manages widely shared, business critical data, ensuring single truth data. Governance of the data is critical for success.
<b>Data Warehousing &amp; Business Intelligence</b>	Data Governance is the foundation of effective DW&BI (e.g. business definitions for KPIs etc.). Also garbage in, garbage out is as true as ever.
<b>Metadata</b>	Metadata provides context & meaning to data and so supports effective Data Governance

# Data Quality – A Simple Definition

Data that is **demonstrably** fit for purpose

**Demonstrably:** Implies that data quality & improvement can be measured, and business impact demonstrated

**Fit for Purpose:** Data quality must meet the needs of the organization and its stakeholders

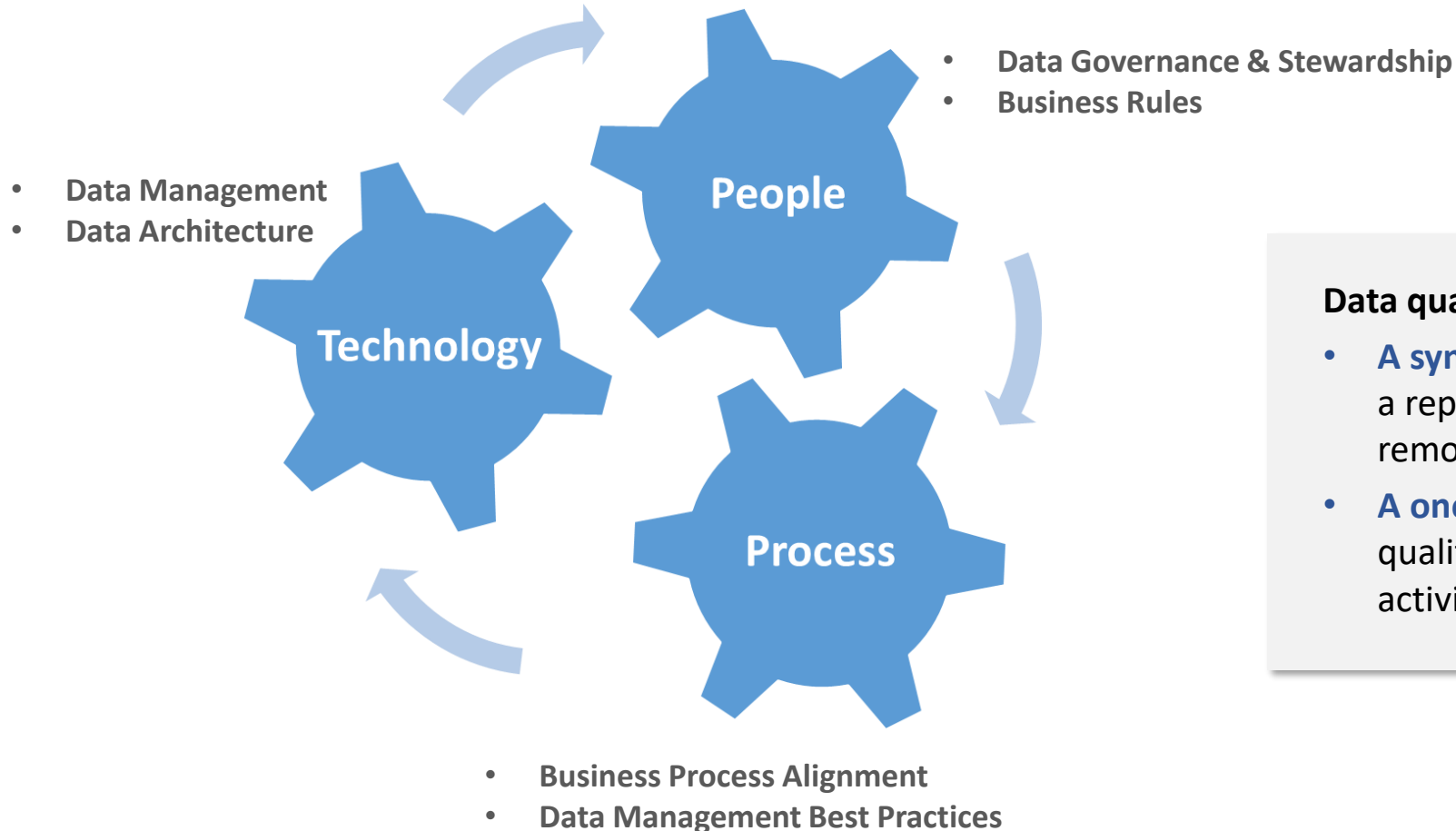
# Fit for Purpose Data

- “Fit for purpose” implies that 100% data quality is not always required or necessary for AI or other purposes
- What is “fit for purpose” depends on the specific needs of the organization and the application
- It is important to understand the specific business context and use of the data
- When data is used for several different purposes, the definition of fit for purpose will vary from user to user
  - Fit for purpose data ideally needs to meet the needs of all its users, including AI
  - Where not possible, a compromise should be sought



# Addressing Data Quality: the Need for a Holistic Approach

Improving Data Quality requires a combination of People, Process, and Technology

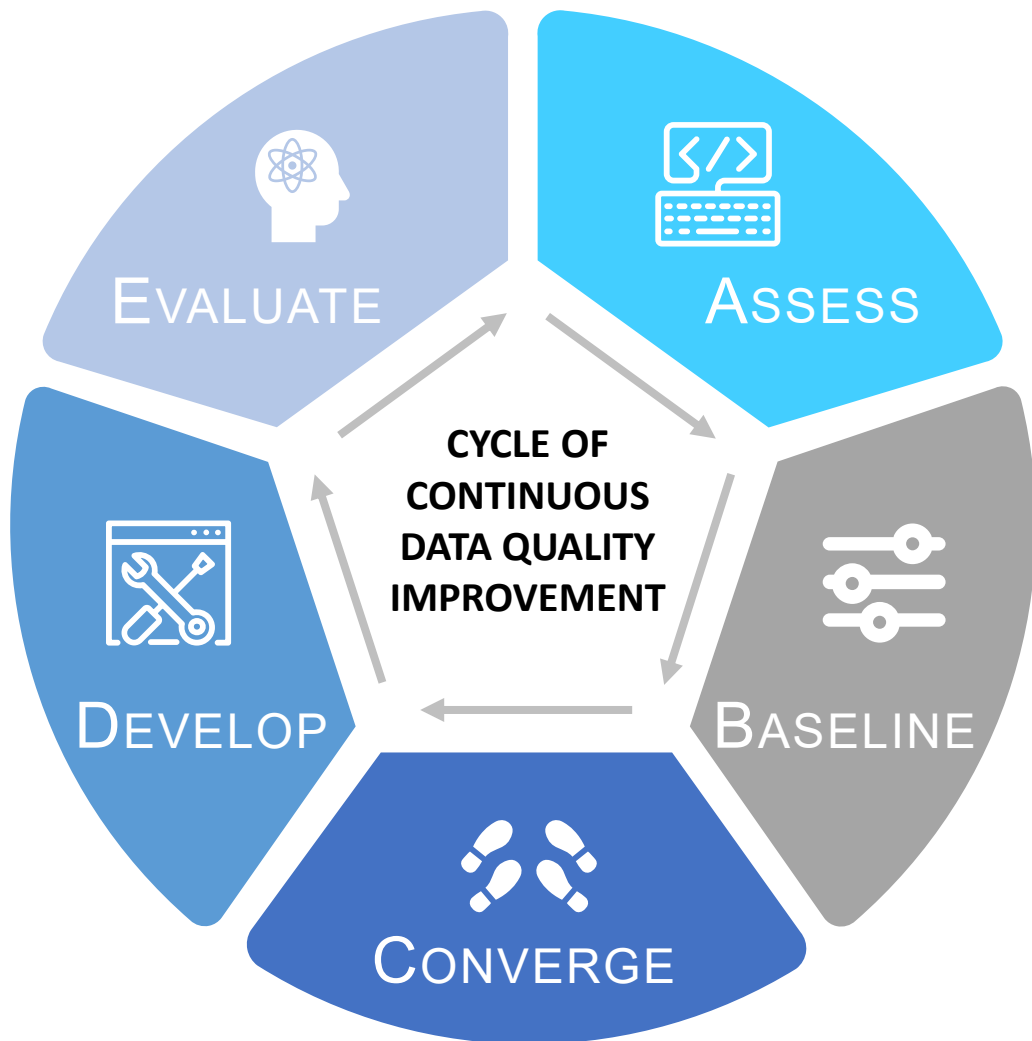


## Data quality is not:

- **A synonym for data cleansing.** Data cleansing is a repeated cost of failure and usually does not remove the root causes of poor data quality
- **A one-time effort.** Data is volatile and so data quality has to be a perpetual business as usual activity enabled through data governance

# Tackling Data Quality Requires a Consistent Method

## The A2E Approach



### Assess Business Usage

Understand what data exists and how it is used within the organization

### Baseline Data Sources

Baseline the current quality of the data and assess how well it is meeting business needs

### Converge on Business-Critical Areas

Focus priorities to optimize early business benefits and set 'fit for purpose' quality targets to guide improvement activities

### Develop Improvements

Design & deploy improvement initiatives (encompassing people, process, and technology) and measure the impact against targets

### Evaluate Benefits and ROI

Regularly measure the data and continue to improve it so that it continues to meet current and future business needs



## A2E Step 1: Assess

### ASSESS THE BUSINESS LANDSCAPE

- Understand the business & its primary objectives
  - Data Governance Prioritization Log & Use Cases
- Identify the primary data stakeholders (Business, IT, External parties (e.g. customers, suppliers, partners))
- Evaluate scale and impact of current data quality and establish:
  - What's working well ?
  - What needs to be improved ?
  - What are the business benefits of better data quality?
- Identify potential ROI benefits if Data Quality issues were addressed

### OUTPUTS & TOOLS

- **Highlight:**
  - Business impact
  - Main data creators and consumers
  - Current problems and issues with the data
  - Opportunities & potential benefits
  - Accountability for the data
- **Outputs include:**
  - High Level Business Case
  - Data Quality Issues (& Opportunities) Log
  - RACI Stakeholder Matrix
  - Business Data Model / Conceptual Data Model
  - Business Process Models



## A2E Step 2: Baseline - The Value of Data Profiling

### Quantifying Data Issues

- Data profiling assesses the quality of data sources
- The benefits of data profiling include:
  - Enables fact-based discussion of the causes and impacts of data problems
  - Excellent starting point in data quality workshops and business engagement

Results Browser

Job: US Customer Data Profiling

Input Field	Total Number	Minimum Length	Maximum Length	Minimum Value	Maximum Value
ID	5438	9	9	AAC434152	ZZZ642455
Name	5438	11	39	Anne Mullen	de Chana, Sergio Marques
Street	5438	2	41	# 3 Riverdrive Rd. East	Wilson & Kirk Road
City	5438	3	20	ABERDEEN	waterloo
State	5438	2	2	AB	WY
ZIP	5438	4	10	01801-6202	n2j4a9
Country	5438	1	13		United States
Phone	5438	1	25	(113) 072 3578	x
Cell	5438	4	14	(113) 575 3765	9978 158
Work	5438	4	28	(113) 007 6029	x7562
eMail	5438	16	35	Aaron.A.Koontz@thu.com	zoi.gibso@snomail.com
DoB	5438	19	19	Jan 1, 1900 12:00:00 AM	Mar 29, 2007 12:00:00 AM
Gender	5438	1	1	F	U
Active	5438	1	1	0	Y
CreditLimit	5438	1	5	0	32800
StartDate	5438	19	19	Apr 1, 2006 12:00:00 AM	Apr 1, 2009 12:00:00 AM
EndDate	5438	19	19	Apr 1, 2008 12:00:00 AM	Apr 1, 2014 12:00:00 AM

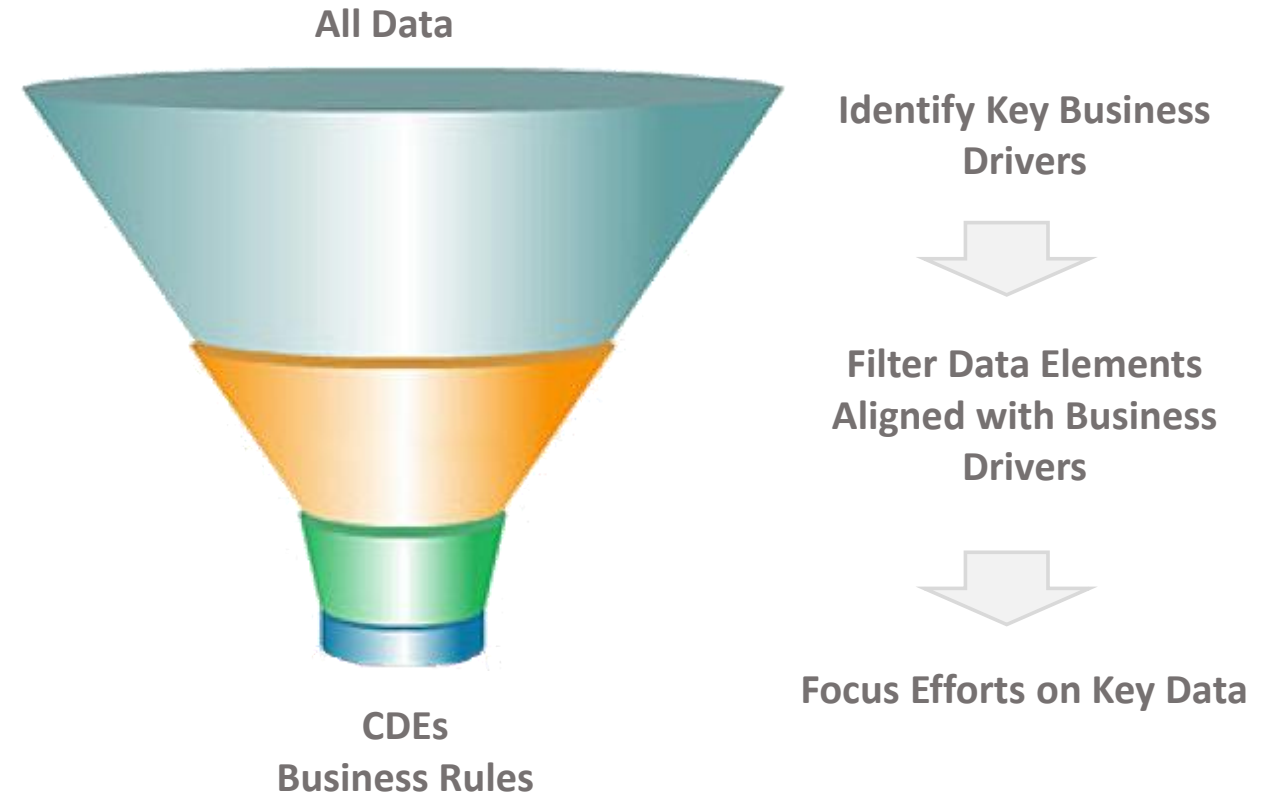
Min and Max Profile | Data

Example partial Data Profiling report



## A2E Step 3: Converge on Business-Critical Areas

- **Identify Critical Data Elements:** Focus on the data that provides high value to the business (e.g., Data quality rules for Fax number format is likely less relevant than the accuracy of a location).
- **Create Business Rules** so that data quality can be tested and measured against your organization's unique requirements.





# A2E Step 3: Determining Critical Data Elements (CDEs)

## What is a Critical Data Element (CDE)?

A Critical Data Element (CDE) refers to data that is deemed essential for an organization's operations, decision-making processes, regulatory compliance, and overall strategic objectives.

## Why identify CDEs?

- Not all data is equally important - focus on the data that matters most
- May be master or reference data that is widely used across the organization
- Enables Data Steward to focus on a limited subset of data to maximize the business benefits of improvement
- CDEs often have higher data quality thresholds than non-critical data

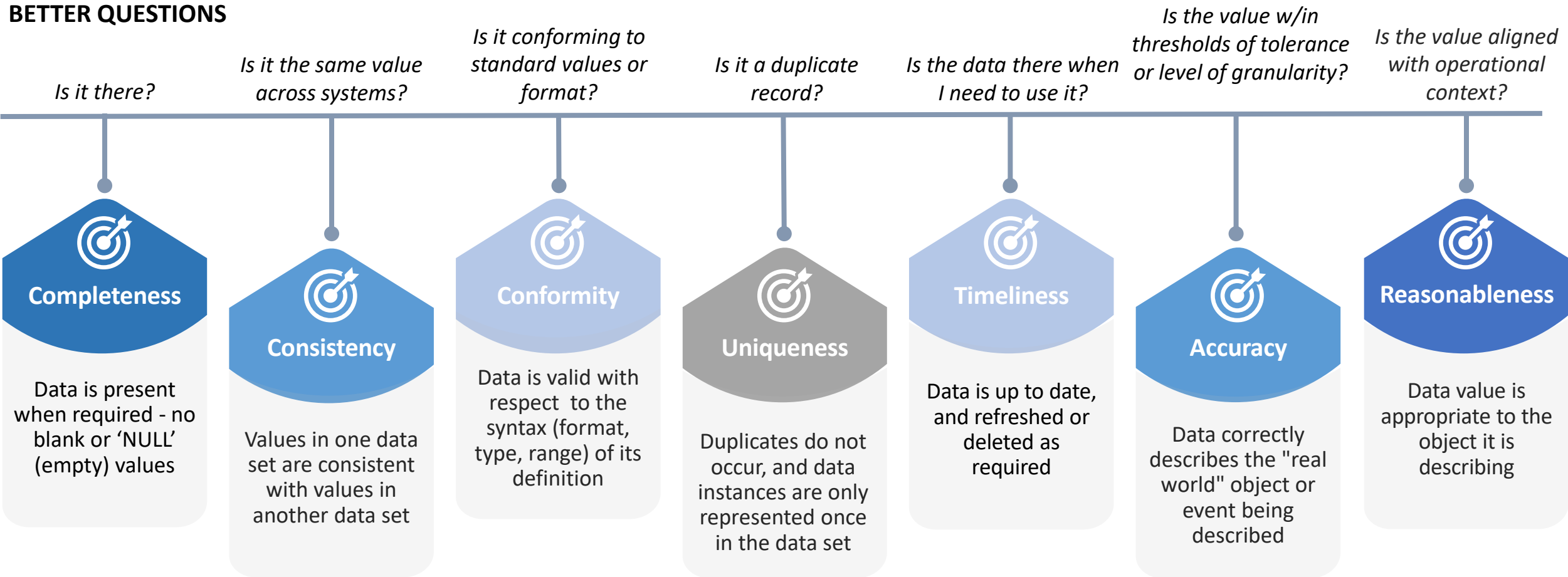
## How do you identify a CDE?

- Criticality to an operational or business process
- Importance in delivering a key business change driver
- Sensitivity due to PII or other privacy / security concerns
- Legal or regulatory requirement

Employee Attributes	CDE?
Employee ID	CDE
First Name	CDE
Family Name	CDE
Date of Birth	CDE
Mobile Number	CDE
Fax Number	
Favorite Ice Cream	

# Business Rules – It's More Than "Is It Right?"

## BETTER QUESTIONS



## DIMENSIONS OF DATA QUALITY



## A2E: Step 4: Develop Improvements The Importance of Quick Wins

**Quick Win Project:** A “Quick Win” Project is a project that shows early value while at the same time building towards a long-term goal. A successful “quick win”:

- Aligns with business objectives and **solves a high-value business problem** *or*
- Creates a **proof of concept for a high-value business opportunity.**
- Sets a **solid foundation for future efforts**
- Acts as a “**light bulb moment**” for **key stakeholders** to understand the value of data-centric activities.

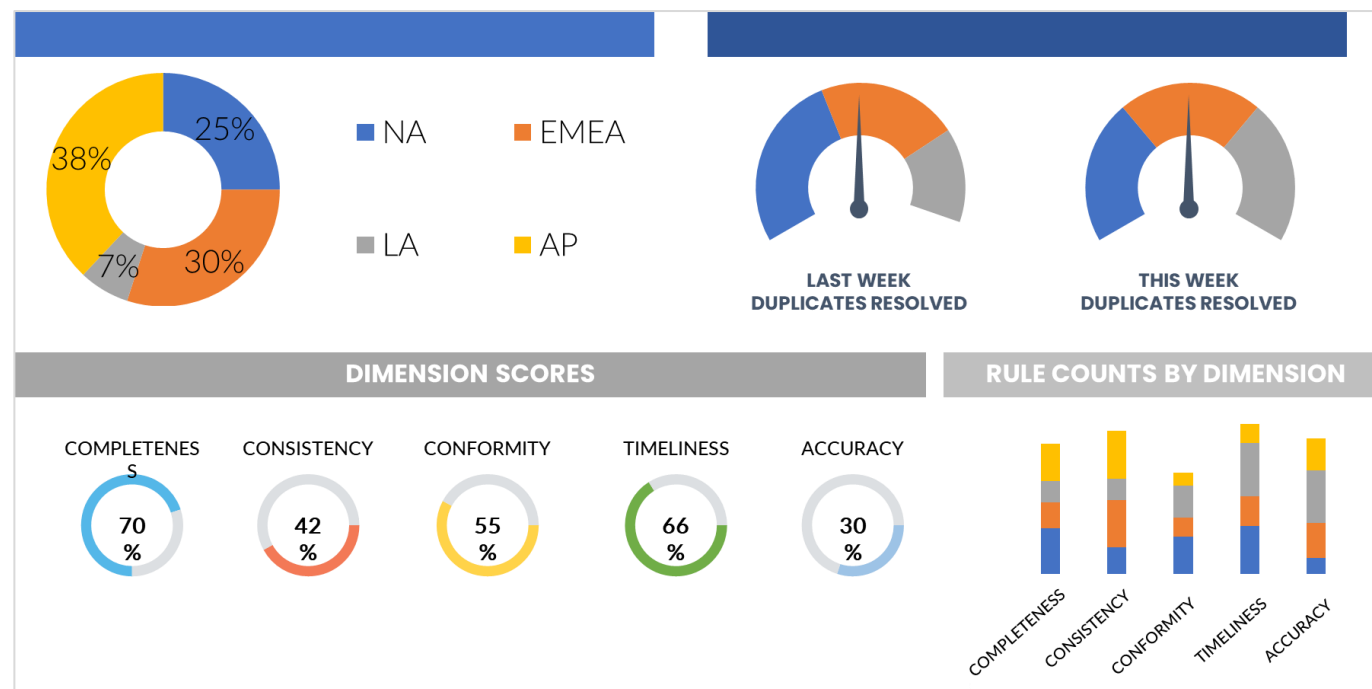




## A2E Step 4: Develop & Implement Improvements

- Identify related operational KPIs that would be impacted with improvements in data quality
- Establish KPI Thresholds and Targets
- Build Data Quality Dashboard(s)
- Create processes that enable the business to monitor and remediate data quality issues more proactively

Example Foundational Data Quality Dashboard





## A2E Step 5: Evaluate

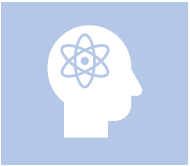
### EVALUATE & SUSTAIN GAINS

- Embed Data Quality improvement as a business as usual activity
- Evolve Data Quality improvement teams into wider Data Governance structure
- Track Data Quality improvements via Data Quality Dashboards
- Monitor financial and business benefits over time
- Evangelizing benefits – part of your job is marketing!

### OUTPUTS & TOOLS

- Evolving & incremental Data Improvement Plans
- Formal Data Governance Roles in place for targeted data areas
- Regular Data Quality Dashboard updates and analysis
- Business Process Change
- Continued ROI and financial benefits
- Communication Plan and Organizational Change Efforts





# A2E Step 5: Monitor & Report Business Rule Adherence

- When Business Rules are implemented they can be used to:
  - Check continued adherence of existing data
  - Enforce the rules on new data to prevent new problems
- Monitor via Data Quality Dashboards
  - Regular reports on adherence of data to Business Rules
  - KPIs drive continuous data improvement
  - Data quality trends
  - Highlight areas where corrective action required
  - Indicate where Business Rules may need to be amended to meet changing business needs
- Relate Data Quality to Business Outcomes
  - Address the “so what” objection
  - Puts a financial or other benefit on continued data quality improvement



Data Quality Dashboards

How Is Data Quality Relevant in the Emerging World of Artificial Intelligence (AI) and Machine Learning (ML)?

# The AI / ML Revolution Is Truly With Us...But...

The generative AI global market, valued at \$43.8 billion in 2023, will grow to \$67.18 billion in 2024 and \$967.6 billion by 2032

*Source: Fortune Business Insights*

## The Opportunity

Approx 18,000 new GenAI start ups were created globally in 2024

*Source: Informatica*



## The Risks

Over 30% of GenAI projects will fail in 2025. The primary cause of failure will be poor data quality, exacerbated by inadequate control of risk, uncontrolled costs, and unclear business value

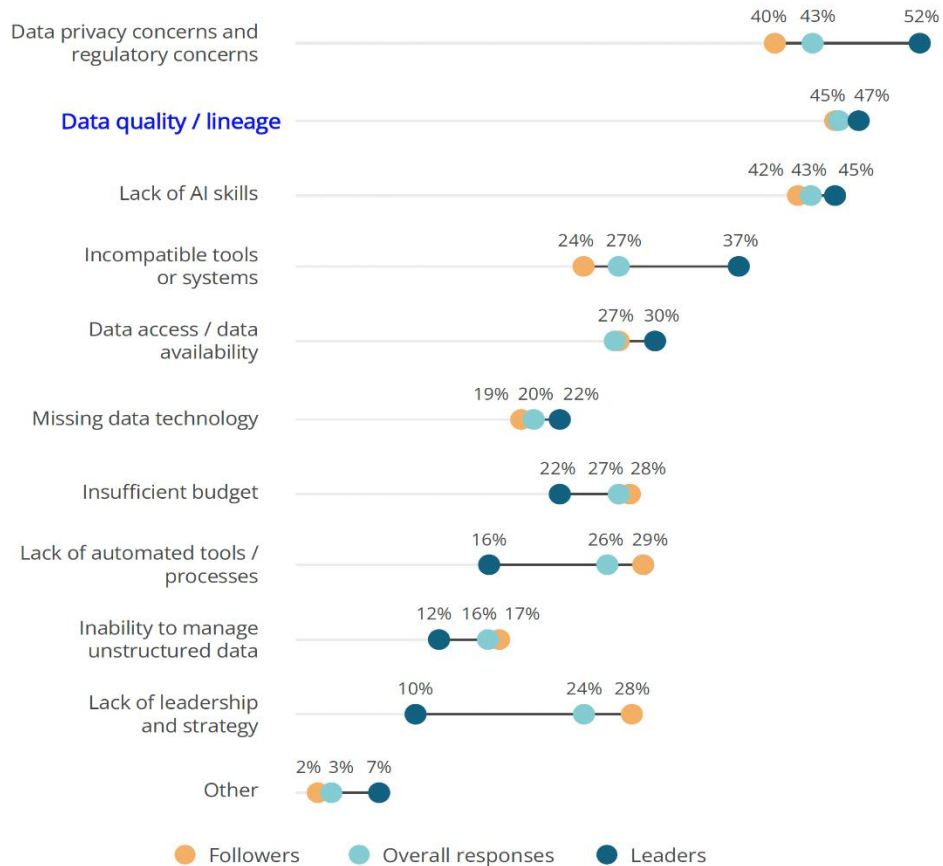
*Source: Gartner*

“AI is just a new tool, one that can be used for good or bad purposes... Although machine learning has huge potential, data sets with ingrained biases will produce biased results – garbage in, garbage out”

*Source: Sarah Jeong, US Tech Journalist*

# Data Quality: The Top Challenge in Managing AI Data

## Challenges and Obstacles for Managing Data for AI



### Viewpoint



Data is the foundation that AI is built on. Without data, AI is not able to function. Access to quality data drives the quality of AI, improves context and ensures accuracy of AI outputs. For many companies AI is a forcing factor for new investments in data systems to confirm that these systems are operating at the highest level.

Respondents to this research have clearly identified data quality as the top challenge to managing data for AI. Almost half (45%) of respondents ranked it as the top challenge or obstacle to managing data for AI. Close behind data quality are lack of AI skills and data privacy/regulatory concerns, both ranked at 43% by all respondents. The significant gap (16%) between these three challenges and all others shows clearly where companies are focused and encountering challenges.

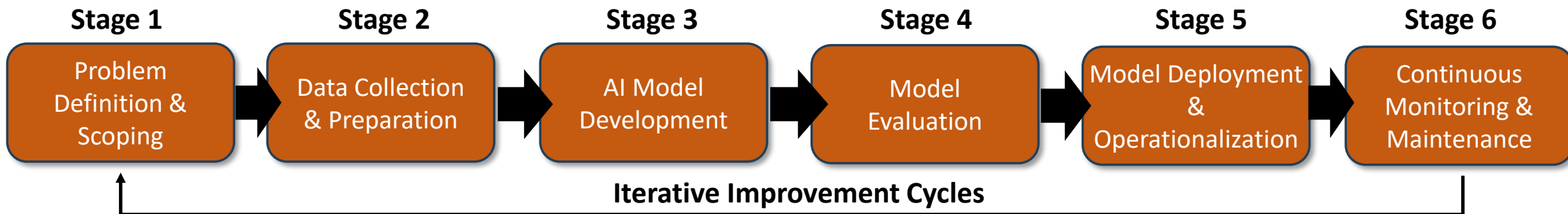
AI Leaders rank data privacy and regulatory concerns as their top challenge, with a 52% ranking versus 40% for Followers. Data quality is ranked second among the leaders. The data also illustrates that followers struggle more with lack of automated tools / processes than leaders (29% versus 16%).

Cross-industry  
global survey  
of 370  
Companies

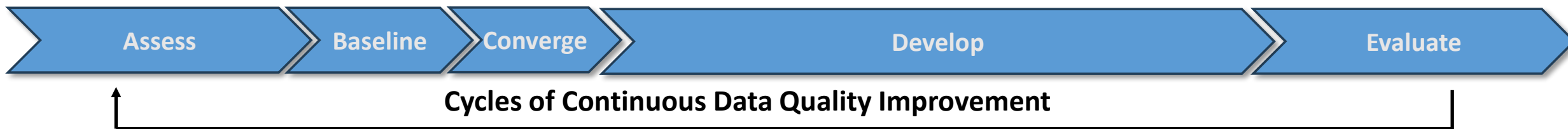
Source: 'Preparing & Delivering Data for AI: Adoption Trends, Requirements and Best Practices', BARC Research Study, March 2025

# Key Stages of an AI Project: The Relationship With A2E

## Typical AI Project Development Stages



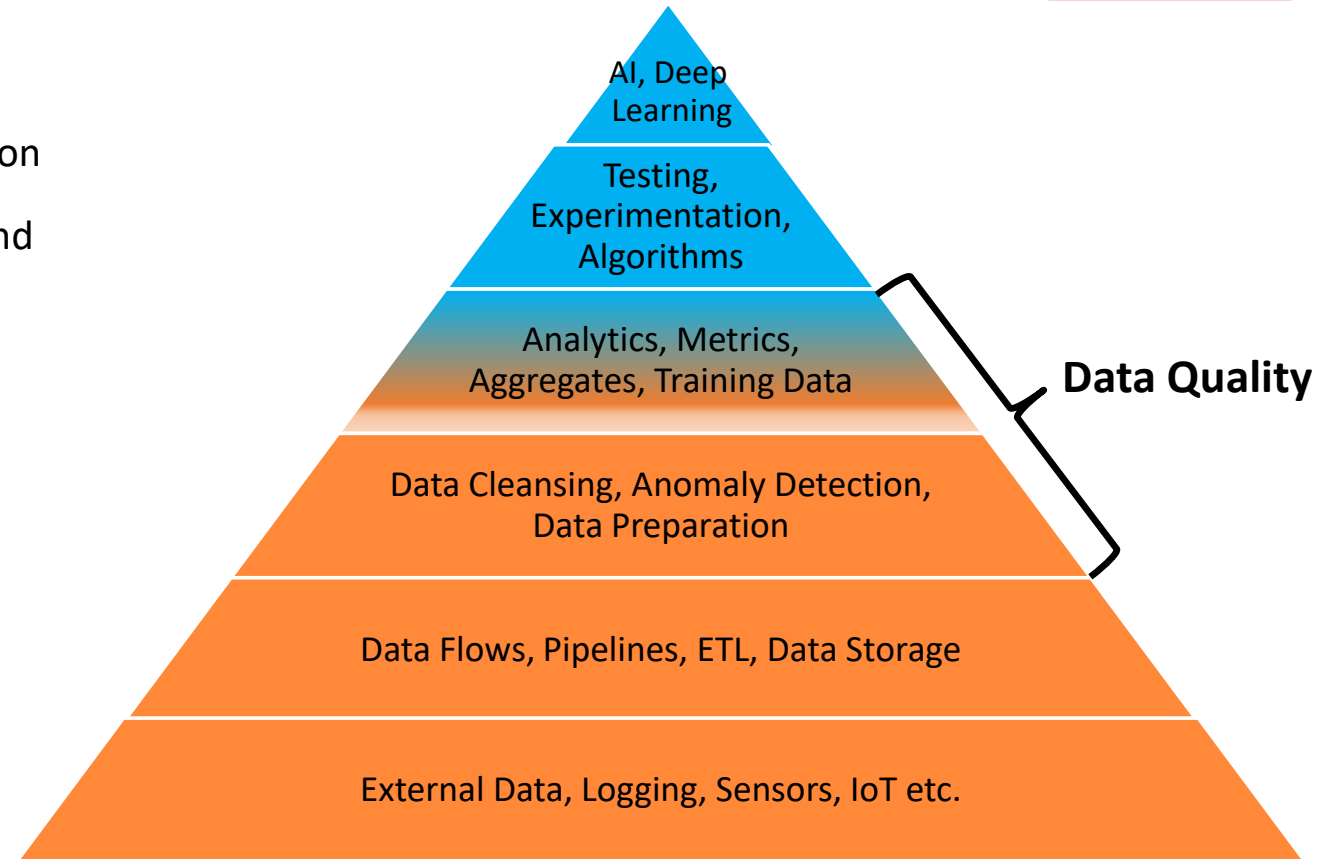
## A2E Data Quality Stages






- **Assess:** Helps to define the data required for an AI use case, data definitions and business data usage
- **Baseline:** Quantifies the current state of the required data and highlights incomplete & inaccurate data
- **Converge:** Scopes the use case to ensure it identifies CDEs and delivers potential business benefits
- **Develop:** Helps to collect and prepare the data for model training and deployment
- **Evaluate:** Continuously monitors data quality and identifies further improvements required for use case refinement

# Data and AI/ML : The Critical Importance of Data Quality

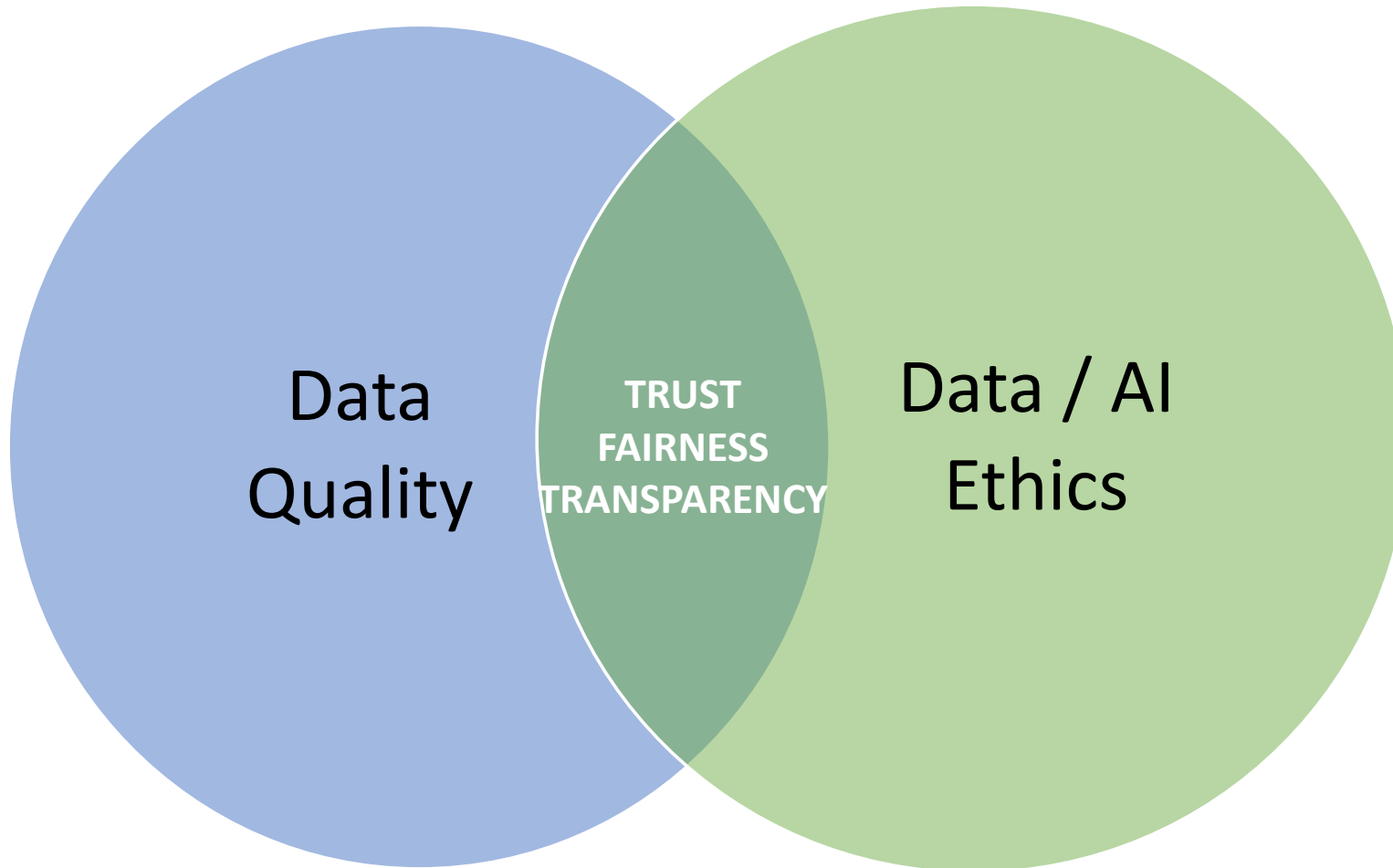
- AI heavily depends on a sound data management foundation
- Data management capabilities generate key data inputs, and enable data movement, integration and storage
- Poor data quality will skew AI decisions and results
- Data quality is a key AI component, encompassing:
  - Data baselining
  - Data omissions and inaccuracies
  - Data cleansing
  - Data preparation
- Building AI applications without including data quality improvement as an integral activity is a recipe for failure



## AI Pyramid

-  Data Management
-  Data & AI Management
-  AI Management

# The Intersection of Data Quality & Data / AI Ethics



## TRUST

- Poor data quality undermines trust in data sets used to train and source AI.
- AI decisions and outcomes cannot be relied on.

## FAIRNESS

- Biases in data sets and AI outcomes resulting from poor data quality can skew decisions and outcomes (e.g. credit scoring,
- This may adversely affect individuals and organizations who are impacted by the bias.

## TRANSPARENCY

- Where AI is used in decision making it is critical that the decision rationale is clear and explicable to those impacted.
- If this cannot be done an organization will have dissatisfied customers and may be open to legal challenges.

# The Future: Data Quality & AI as Mutual Friends

## Data Capture & Validation

- AI can help to automate data capture and so avoid human error
- Where human data entry is necessary AI can validate data input based on defined business rules and prevent the input of bad data

## Predictive Analysis

- AI can be used to predict missing or unknown values in a dataset based on patterns in the data
- This can improve the completeness and accuracy of the data

## Anomaly Detection

- Machine Learning Models (ML) can be used to identify data that could indicate anomalies
- For instance, a ML model can be trained to identify transactions that fall outside of the normal range for a business, which could indicate data quality errors or fraudulent activity

## Automated Data Cleansing

- AI can automate the process of data cleaning, including identifying and removing duplicates, filling in missing values and correcting inconsistent or inaccurate data
- Machine learning algorithms can be trained to recognize errors that are often hard to spot with manual review



## Proactive Data Quality Management

- Proactively monitoring data quality metrics and applying predictive analytics can detect potential data quality issues before the business and customers are impacted
- ML algorithms can analyze large scale, historical data to detect patterns, issues and provide recommendations to enable Data Stewards to implement preventative measures

## Data Standardization

- AI can automate the process of transforming data into a standard format
- This can help to ensure consistency and improve the accuracy of data

# Summary

- **Data quality management remains a challenge** for most organizations
- **To fix data quality requires a holistic approach** involving People, Process, Technology & Data
- **Build a solid data quality foundation** as a prerequisite of success for AI use cases and to ensure AI is applied ethically to ensure trust, fairness and transparency
- **Use a repeatable and common approach such as A2E** to increase the chances of delivering and sustaining a high data quality culture, and ensure AI and business success



# DATAVERSITY Data Architecture Strategies

## This Year's Lineup

- **January** Trends in Data Architecture
- **February** Building a Data Strategy - Practical Steps for Aligning with Business Goals
- **March** Building the Right Architecture for Analytics & Reporting
- **April** Data Architect vs. Data Engineer vs. Data Scientist – Making Sense of Roles in Today's Data-Centric Organization
- **May** Master Data Management - Aligning Data, Process, and Governance
- **June** Where Data Models Fit in Today's Modern Data Architecture
- **July** Data Architecture vs. Enterprise Architecture
- **August** Data Quality Best Practices (with guest Nigel Turner)
- **September** Modern Data Architecture: Practical Options for Today's Data-Driven Organization
- **October** Best Practices in Metadata Management
- **December** The Business Value of Data Modeling



# Who We Are: Business-Focused Data Strategy

Maximize the Organizational Value of Your Data Investment



In today's business environment, showing **rapid time to value** for any technical investment is critical.

But technology and data can be complex. At Global Data Strategy, **we help demystify technical complexity** to help you:

- Demonstrate the ROI and **business value of data**
- Build a data strategy **at your pace to match your unique culture** and organizational style.
- Create an **actionable roadmap for “quick wins”**, which building towards a long-term scalable architecture.

Global Data Strategy shares experience from some of the largest international organizations scaled to the pace of your unique team.

Global Data Strategy has worked with organizations globally in the following industries:

Finance · Retail · Social Services · Health Care · Education · Manufacturing  
· Government · Public Utilities · Construction · Media & Entertainment ·  
Insurance .... and more



Thoughts? Ideas?  
**Questions?**