DEMO DAY
Enterprise Data Management

2021
About Me

John Ebenezer  
VP - Head of Data and Analytics  
at Compunnel Digital

▷ 20+ years of professional experience leading Enterprise Data & Analytics Services.

▷ Industries like Banking & Finance, Insurance, Lifesciences & Healthcare, Education, Retail and Manufacturing working for Fortune 5000 clients.

▷ Specialize in Data and AI, including Enterprise Data Management (Data Governance, DQ, MDM), Data Warehousing, Data Analytics, Big Data Engineering, Data Architecture, Program Management, and Organizational transformation.
Agenda

The demonstration today showcases use-cases in Enterprise Data Management that brought transformational impact for customers:

- Enterprise data intelligence platform with automation and real time analytics while complying with HIPAA standards.
- Advanced Data Analytics solution incubating data quality and MDM that provides increased operational efficiency, risk management, and revenue growth.
- Unifying disparate systems with Data Management Solution to streamline data processes and IT & business alignment.
- Unified marketing analytics platform that drives digital ROI for you.
- Key Data Management Implementation Questions
- Modern Data Management Enrichments
The Use Case

Advanced Data Analytics Solution Incubating Data Quality & MDM
Business Problem

The Banking client was on roadmap to move all its branches to CBS (Core Banking Solution) enhancing customer convenience through Anywhere and Anytime Banking. It was also undertaking critical information delivery projects like Customer-relationship Management (CRM) and Enterprise-wide Data Warehousing (EDW).

The Roadblocks

- Absence of Data Standards and defined accountability for Masters
- No SLAs defined for the various stages of Master Data like approval, authorization, master data creation etc.
- Lack of standardized processes for creating and updating master data
- No standardized forms / templates for requests for creating and updating master data
Solution Approach

Identify
MDM Charter

Diagnose
Profiling and Gap Analysis

Design
MDM Solution Design

Develop & Deploy
Implementing the Process

Sustain
Support and Maintenance
Outcome

- Improved data accuracy that was regulatory compliant
- Consolidating the data from various source systems
- Enabled single version of truth
- Eliminating redundant and incorrect data
- Designed single common repository that enable data access more effectively

Technology Stack:

- [Informatica MDM](#)
- [Microsoft SQL Server](#)
- [Informatica PowerCenter](#)
Benefits

Building Trust in Data
Improve Customer satisfaction scores by 80% due to metadata, glossary and quality tool implementation.

Revenue uplift
Able to launch new Banking Products.

Productivity Gains
45-50% reduction in cost of servicing customers due to process validation & governance.

Role & profile-based Data Access
Entitlement defined for 14 business functions.

30% → 94%
Increase in overall Data Quality index within 6mths.
Business Problem

A leading Healthcare service provider known to treat traumatized and serious emotional problems among the teens and young people.

The client has continued to struggle across multiple divisions to access information as well as capture the information needed to meet industry standards with their funders and governing bodies. Because of this, they considered seeking a new EHR solution, but then learned their current system was likely not leveraged to its full capabilities. The organization, however, was unclear on how to optimize the solution and desperately needed guidance on appropriate next steps. It also lacked the internal methods to integrate system changes alongside their evolving organization and health care ecosystem.

- Their processes required an abundance of manual tasks and duplicative efforts and provided inadequate reporting and analytic capabilities. Additionally, the improper use of applications and use of external systems (including paper documentation) created inefficiencies, a lack of data integrity and a lack of visibility across the entire organization as there was no defined central view of the individuals served.
Solution Approach

Prioritise and Identify
- Prioritized the applications that had data quality issues.
- Identified the areas that had major business impact

Assess
- Performed assessment of the existing EHR systems.
- Identified the gaps in the operational workflows.

Enable DQ Processes
- Designed the end to end DQ process.
- Defined the business related DQ rules
- Performed DQ standardization, matching, enrichment and validation.
Outcome

- Improved data accuracy, consistency and accessibility.
- Identified the disconnected systems and multiple manual processes.
- The streamlined DQ process enabled the applicable use of intended purposes.
- Standardized the Critical Data Elements across the systems.

Technology Stack

- Informatica IDQ
- Oracle
- Unix
Benefits

- **Significant improvement in patient care response**
- **Increase by at least 50%** in effective and timely service to the patients.
- **40-50% increase** in effective administrative and clinical operations.
- **About 85% Reduction** in duplication of same patient data from different sources.
The Use Case

Unifying Disparate Systems with Data Management Solution
A US based university consisting of a group of colleges was facing wide discrepancies among departmental data, and which eventually led to submission of inaccurate data to institutional research. The university would struggle to steward the data they need to manage their business and future plans. Departments across every campus collected some kind of data, whether formally or in homegrown departmental shadow systems. This is where the problem lies: data is ubiquitous and plentiful, but it’s rarely organized, standardized, or leveraged for a common purpose.

They used to source data from separate departments using different tools which led to inconsistencies and inaccuracies. It was difficult to report on institution-wide activities, leading the leadership, administrative functions, and academic departments and colleges competing for limited funding opportunities.
Solution Approach

- Establish Executive Sponsorship
- Translate the Institutional goals and objectives to guiding principles
- Identify and Prioritize the key areas that has high business impact
- Build the Governance framework defining Processes and Policies
- Evaluate and decide on the right technology options
- Establish a Change Management Plan to enable Departments to opt in a new way of handling data.
- Effective communication across the organization
Solution: Enterprise Data Governance Framework [EDG]

Governance Strategy & Vision

Control (Efficiency)  Growth (Monetization)

Productivity  |  Brand Equity  |  Performance

Business As Usual & Program Management (PMO)

Data Stewardship
- Lineage
- Glossary
- MDM
- Metadata
- Data Quality
- Access & Ownership
- Security

Innovation

Change Management & User Adoption

DG Core Areas Block

DG Enabler Block

DG Foundation

People & Culture

Policy, Process & Operating Model

Tools & Technology

Data Architecture
- Acquisition
- Analysis
- Processing
- Storage
- Sharing
- Retention
- Disposal

Information Lifecycle Management

Structured
- Apps & DBs
- Master & Reference Data Hub
- EDW & DL

Unstructured
Outcome

Reduction in data fragmentation by bringing together data from multiple systems and ensuring security across all aspects.

More advanced and accurate data reports with the implied Data Governance strategy.

Consistent data quality through Data Stewardship group because of precleaned and standardized data.

Visibility across multiple data domains.

Technology Stack

Collibra
SAP HANA
SAP
Data Services
Benefits

Enhanced communication and joint decision-making.

Authority and control over the Institutional Data asset

60% Increase in data security and privacy compliance for data reliability

30% Increase in top line growth of Student Admissions and 40% reduction in operational cost that caused due to redundant efforts.
The Use Case

Unified Marketing Analytics Platform
Business Problem

Client is a large license health insurance provider to more than 3.8 million people in United States. They empower the people for their best health by leading with strength, integrity and innovation to generate substantial improvements in health care quality, affordability and member experience.

In Client’s current system, around 20% of the business correspondence was returned as there was no address standardization Source systems stored the addresses in comment field making it difficult for correspondence.

- Inability of system to manage and integrate heterogeneous and disparate data regardless of the source.
- Duplicate data across multiple systems leading to data divergence
- Lack of visibility across multiple data domains for members, providers, products etc.
- Inaccurate reporting & analytics from the Data Warehouse
Outcome

- Standardised address validation and address master
- Increase in the delivery of the addresses used
- Reduction in data fragmentation by bringing together data from multiple systems and ensuring security across all aspects.
- Consolidated Email Addresses for members

Technology Stack:
- Talend MDM
- DB2
- SQL Server
Benefits

- **Address Cleansing**
  - Increased by approximately **90%**
  - Includes mapping of city to zip

- **Almost 40% reduction**
  - in customer service cost

- **20-30% Reduction**
  - in administrative and operational cost

- **Unified view of the customer**
Key Data Management Implementation Questions

Demo Day 2021
What are the challenges in Kick starting a DM engagement?

Organizations continue to invest in data and analytics; however, many face challenges in translating DM investments to business outcomes and value.

Realizing business value and enabling business objectives should be the focus while deciding on DM investments.

A robust DM strategy and a clear, actionable roadmap of initiatives & supporting capabilities (people, process & technology) can help organizations stay on track in their pursuit of benefits from DM.

The approach to creating an effective DM roadmap starts with understanding the business objectives and key business questions.
What is the quick win that can return good ROI in the short term?

Focus on ‘speed to value’ and incremental realization of capabilities, outcomes and benefits help minimize the time stakeholders have to wait to see results and benefits.

Agility to respond to changes in priorities and competitive landscape is critical for DM roadmaps.

A consistent prioritization framework, aligned to business strategy and value, to evaluate various DM opportunities is important.

Digitalizing the prioritization and planning process can enhance effectiveness of the process by making it more simple, consistent, transparent and agile.
What are the new methodologies in implementing Governance?

- Collaborative Approach with Customizable Roles
- Advanced UX & UI with Messages and Alerts
- AI & ML for User Assistance and Support
- Proactive and Preventive DataGov for DataOps
- Automation of Technical Processes
- Metadata-Centric Multi-Cluster and Multi-Environment
- Open-Source Ecosystem, API-First and Cloud-First
- Scalable Architecture with Microservices
- Unlimited Users and no Vendor Loc-In
- Low Initial Investment and Pay-Per-Use Model
How to minimize the risk of DM Implementation failure?

Moving from...

- Unclear roles & responsibilities for data management
- Data in silos with inconsistent enterprise use
- Varying levels of data quality without formal mechanism to improve quality
- Lack of support, funding and alignment for data management
- Data management as a separate activity and group

Towards...

- A well-defined stewardship model based on clear roles and responsibilities with executive oversight
- Data organized and managed in enterprise domains with authoritative source and data mastered across the enterprise
- Robust end-to-end data quality program with clear roles, process and tools
- Strong executive support with sufficient funding and broader organizational alignment
- Data management embedded in the DNA and day-to-day of everyone working with data
What’s new in the Data Management space?

- Artificial Intelligence/Machine Learning for DM
- Augmented Data Management
- DataOps
- Data Governance Trends
- Data Fabric
Thank You!

103 Morgan Lane, Suite 102, Plainsboro, NJ 08536
609-606-9009
APPENDIX
Artificial Intelligence/Machine Learning for DM Automation

- AI plays an ever-increasing role in enterprise solutions. Unlike robotics, which automate manual tasks, AI automates computing tasks.

- That’s especially valuable given the large and diverse data sets most organizations use today.

- While the human role in enterprise solutions will never disappear, it’s foolish to argue against the advantage of AI-augmented humans.

- Gartner estimated organizations that offer a curated catalog of internal and external data will realize twice the business value from their data and analytics investments. It also predicted 80% of data lake projects will fail to deliver value due to challenges in inventory and curating data through next year. A big hurdle to this is the manual effort required to curate and manage the data catalogs. Gartner predicted 60% of data catalogs that do not use machine learning to assist in finding and inventorying data across a distributed environment will fail to be delivered on time.

- But for AI to reach its full potential, the data feeding its algorithms and models needs to be well-understood. Data lineage plays a vital role in understanding data – making it a foundational principle of AI.
“Augmented Data management: Metadata Is the New Black” ranks fifth among the Top Trends in the world of Data & Analytics.

Augmented Data Management technologies will free up to 20% of their time for collaborations, education and self-education, and for high-value DM tasks.

According to Gartner, by 2023 IT specialists will be less engaged in managing and preparing repetitive and low-impact data.

By 2023, organizations that dynamically automate, connect and optimize their DM processes via Active Metadata, Machine Learning, and Data Fabric will spend 30% less time on Data Integration. Today everybody needs to know what data are available, what their meaning within the organization is, how valuable and how reliable they are.
Making Sense of the widely available DATA

Huge amounts of data are often distributed in multiple sources, perhaps in various cloud systems. Making this information quickly available for different users is not as easy as pie. So, what can we do?

Step: 1
Boost up the metadata. These contain the organization’s data in the form of entities, their attributes and relationships between them. Consider a metadata management system as a corporate safe.

Step: 2
It is usually administered by the Chief Data Officer and neatly stores all that concerns the company’s data of interest.

Step: 3
The ability to use this information, i.e., to activate the metadata, allows the system to:

• Suggest new data quality rules; report the availability of new metadata; detect the presence of sensitive data for privacy purposes;
• Identify the use of the same data in different business processes and use cases;
• Determine the degree of relevance, and much more.
• Create a Data Fabric or an architecture designed as a structure of interwoven services, microservices and DM components.

Use Augmented Data Quality technologies to automatize data quality controls and resolve detected anomalies based on pre-established policies and rules. Use a Data Catalog to register all the company’s data assets and related entities. Advanced techniques (Artificial Intelligence/Machine Learning) allow to automatically collect and organize such metadata. Then it is easy to physically locate the data, understand their semantics and assess the quality. Besides, all parties concerned get smooth and controlled access and sharing.
Augmented Data Management

- Extensible Data Fabric
- Semantically Enriched Enterprise Graph
- Security, Compliance & Audit
- Refresh Data & Insights in Real Time
- Dynamic Clustering & Scalability
- Smart Data Discovery
- Auto-detection of Patterns

© 2021 Compunnel Digital, Inc. All rights reserved
Automation of Augmented Data Cataloging and Lineage

In the past several years, two key trends have emerged in data management. Data warehouses moved to SaaS, and now data pipeline and ETL tools are transitioning to SaaS as well.

Data operations moving to SaaS allows for auto-generation of data cataloging and lineage as data pipelines are being built. Inputs such as data pipeline metadata and the analysis of that metadata are created on the fly and will become a completely automated process.

Companies that can automate augmented data cataloging and lineage will leapfrog the rest of the market. This eliminates manual work and maintenance and opens data operations to personnel who are not comfortable with managing technical infrastructure.
Automation of Augmented Data Cataloging and Lineage
DataOps will reach the mainstream as data grows in volume and complexity, scaling a data operation without this holistic approach becomes much harder.

DataOps applies the principles of DevOps to data management. DataOps combines agile development, technologies, processes, and practices such as statistical process control to deliver data and analytics across a company.

Organizations that implement DataOps will accrue a number of benefits, including improved insights, cost reduction, higher efficiency, and most importantly for the long-term, rapid scalability.

Implementation is agile in nature that enables to realize the benefits or ROI quickly.
DataOps
(Data + Operations)

- Balance data governance
- Democratize data
- Automated where possible
- Go open source
- Communicate & collaborate
- Use end-to-end design thinking
Data Fabric
Data Fabric

A data fabric manages the collection, governance, integration, and sharing of data across a single unified architecture.

To construct data fabrics, companies are using capabilities such as graph technologies and semantic standards, along with solutions such as ETL, ELT, and augmented data management.

Data Fabrics will grow even more essential for companies undergoing digital transformation. As companies continue to migrate to the cloud, as data volumes and data types explode, and as data consumption turns toward on-demand channels, the need to seamlessly “weave” together data ecosystems will grow.