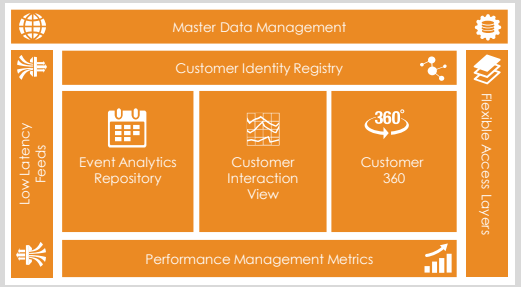
Connected Data – A Foundational Pillar of the Customer Journey Solution

TERADATA.

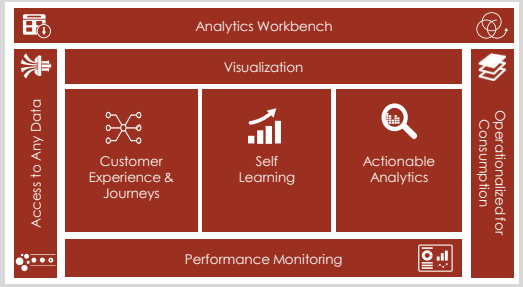
Customer Journey Solution:

Core Capabilities for Customer Journey Analytics & Execution

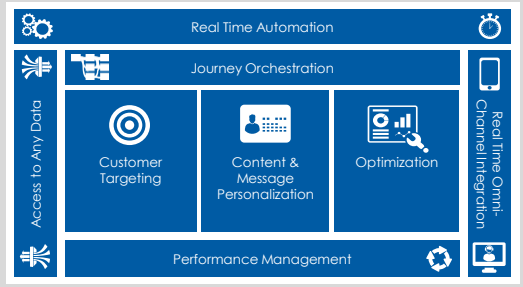
Connected Data



Connected Analytics



Connected Interactions



"GATHERING & CONNECTING ***VISUALISING*** ***ACTING"***

Consulting Services

Celebrus, MDM Analytic Data Platforms LDMs, Claraview, Consulting	Multi-Genre Analytics (Aster, R) Predictive Analytics (CIM/RTIM) Actionable Analytics (CIM/RTIM)	Customer Interaction Manager Real Time Interaction Manager

Technology Overview for Customer Journey Analytics, Gartner, August 2016

Customer Journey Solution

All of these elements are complementary

Connected Data

- Low Latency Feeds & Flexible Access Layers
- Event & customer interaction Views
- Customer Identify Registry
- Customer 360° Performance Metrics

Connected Analytics

- Actionable Analytics
- Customer Experience & Journeys
- Self Learning
- Visualization & Reporting

Connected Interactions

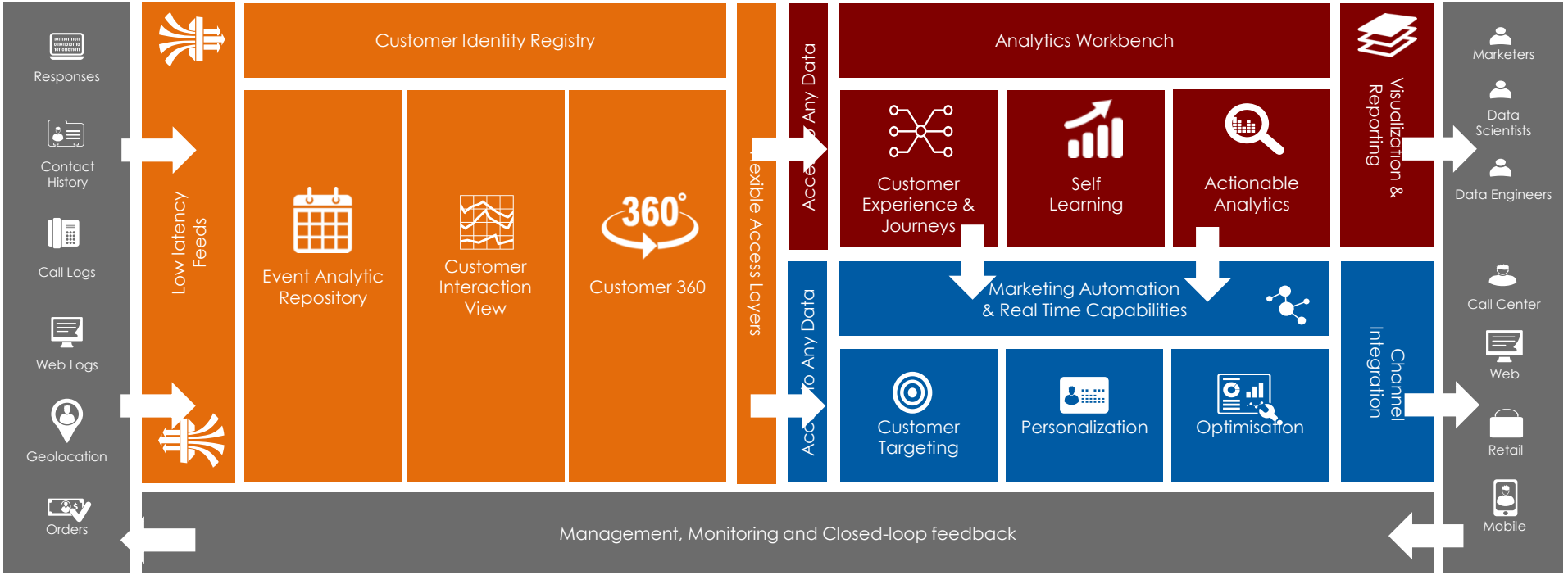
- Personalized Messaging
- Customer Journey Orchestration
- Omni Channel Choreography
- Contextual & Real Time Decisioning

Deployment Options: On premise, SaaS or Hosted

TERADATA.

Customer Journey Overview

A low latency architecture optimized for deploying operational analytics



TERADATA.

Customer Journey Insight Creation – Value Chain

Connected Data element focusses on the data engineering tasks

1. Reduced time to conduct data discovery

2. Reduced time to operationalize analytics

Prototype | Determine Value | Refine | Industrialize | Deploy to Production

Data Engineering

Data Science

Data Scientists

3. Business led data scientists are in control of more of the end to end value chain

4. Standardized definitions and common summaries reduce amount of domain expertise require to understand data

Domain Expertise

DevOps Engineering

5. Reduced reliance on specialist skills to productionize insight

TERADATA.

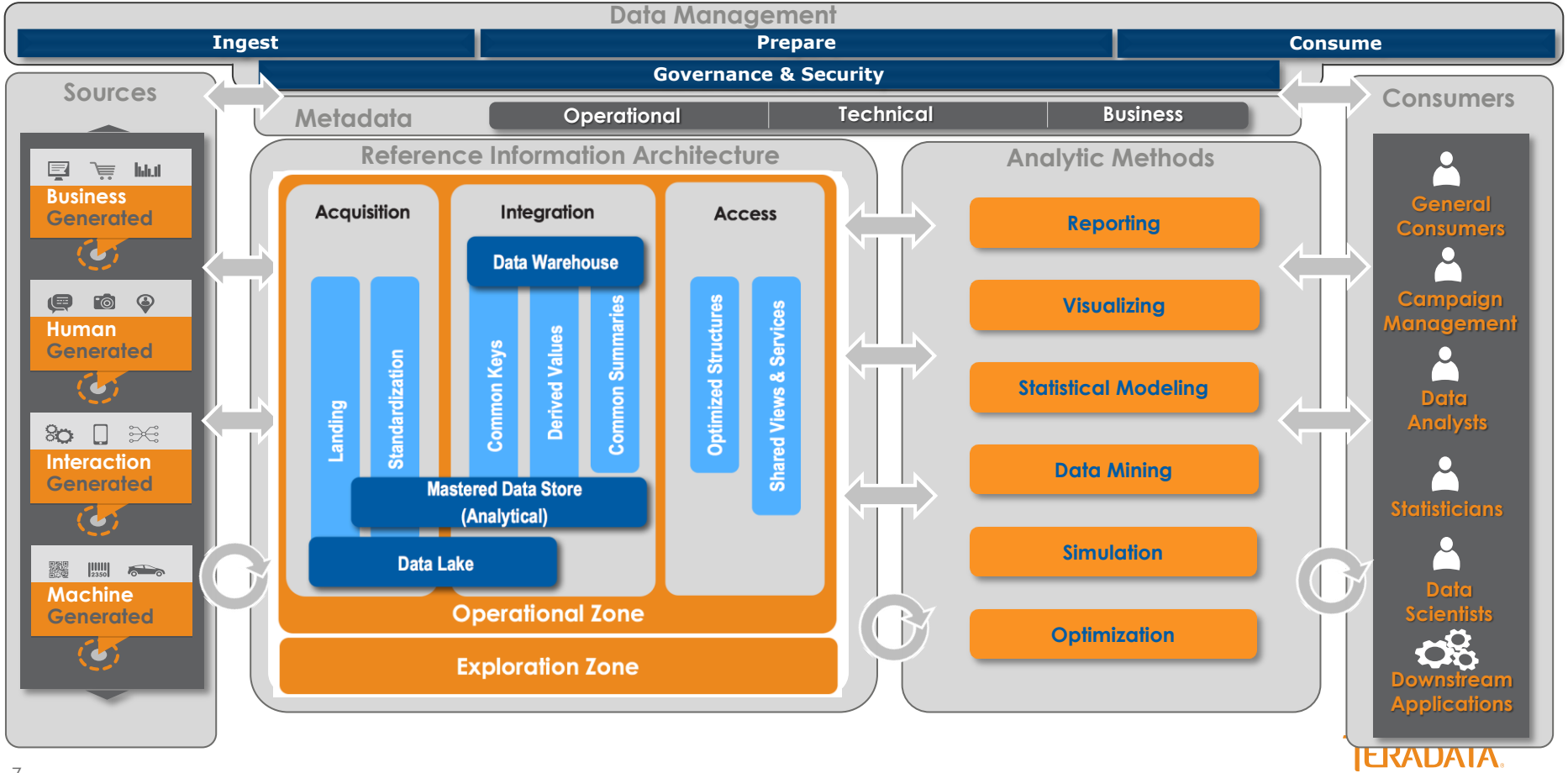
Customer journeys are complicated

- and this complexity continues to increase

–

3 key prerequisites are required	
Prerequisite #1	**Creating a complete view of interactions across channels provides context**
The explosion of digital channels and devices makes it increasingly difficult to understand customer journeys. Understanding Omni-channel journeys can be really difficult – it requires the integration of multiple data sources that are often developed independently of each other. This makes piecing together individual interactions tricky and time consuming but is vital to provide an accurate customer context.	
Prerequisite #2	**Interactions need to be linked to fully understand intent**
Customer journeys are no longer simple linear processes. Buying patterns for digital channels contain lots of separate activities that could be performed multiple times – both onsite and offsite. Timescales range from minutes for a complete online journey to hours / days / months – sometimes with large gaps between different activities. Having a complete journey enables the better identification of customer intent	
Prerequisite #3	**Immediacy has become a hygiene factor**
Being able to present the right information at the right time is critical to catch the customers when they are open to the influence of a company's brand. Google describes these as micro moments when customers want to perform a task such as I-want-to-know, I-want-to-go, I-want-to-buy or I-want-to-do. If organisations can orchestrate these moments with frictionless customer journeys – that are fast, simple and relevant they are more likely to get the mind share of the customer.	

Analytic Ecosystem



15++ years in the Big Data, Digital & eCommerce

shutterfly
Enterprise Data Architecture

NOKIA
Clickstream

GROUPON
EDW Strategy

charles SCHWAB
Enterprise Data Architecture

Expedia
Enterprise Data Architecture

overstock.com
Your Online Outlet™
Enterprise Data Architecture

facebook
EDW Strategy

ELECTRONIC ARTS™
Clickstream

ebay
Digital Marketing

PayPal
10 Year Active Badge

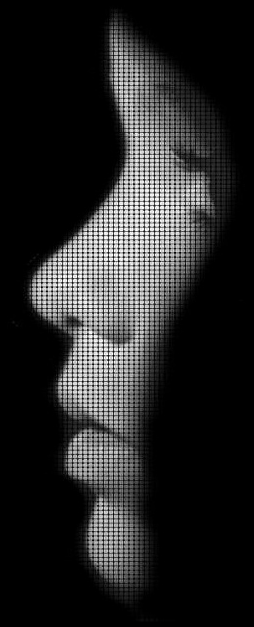
real
EDW Data Architecture

Rhapsody
Enterprise Data Architecture

T··Mobile
Big Data Migration

TESCO
Enterprise Data Architecture

meredith
Enterprise Data Architecture

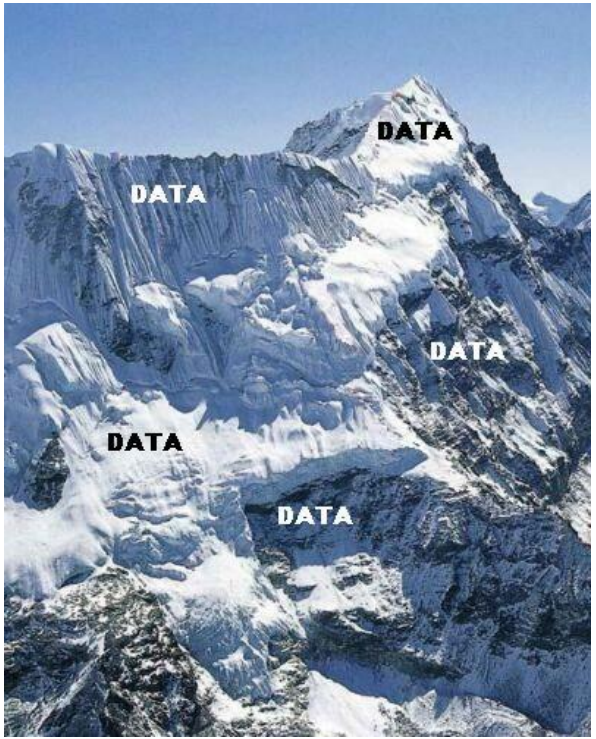


Data Architectures Broke About 4 or 5 Years Ago

- Business Started to Ask Different Questions
- New Technologies Emerged
- IoT: Data Volumes and Nature of Data Changed
- Funding began moving more fluidity between CapEx to OpEx

Total Amount of Data: Competitive Advantage

- **The amount of data is a competitive advantage**
  - Algorithms are good
  - Science is good
  - All the work you put into processing & managing the data is good
- **What makes a successful data platform is:**
  - Amount of data you can have in working memory
  - Amount of data you can join and can be a cohesive set (trusted data)
  - Amount of data you can use to do something with



TERADATA.

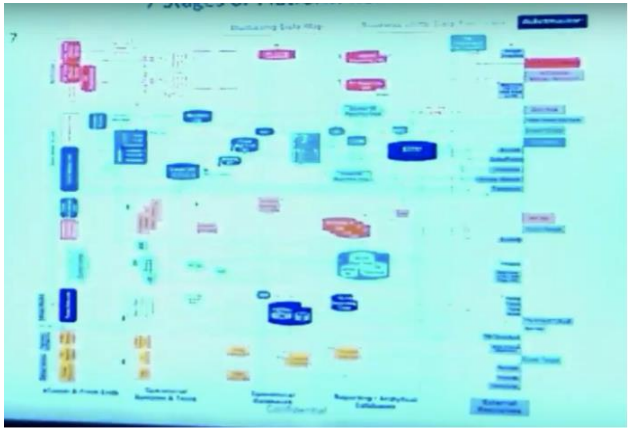
An Awesome Time for New Technology!

- Open source is like a playground for engineers
- Tends to skew the build vs. buy decisioning
  - Good for engineer's resume (CV)
- Many times technology decisions are not fully based on business needs

7 Stages of Platform Remorse

1. I can load my operational/transactions data e.g., from my eComm systems quickly into operational stores like MySQL, Cassandra & Redis); I can share it and I can scale it.
2. But then I get this log data that doesn't fit **into operational stores so I put it into Hadoop** and then put Hive on it
3. Then I get 3rd party data sets; some externally and some from within the company (Oracle legacy systems), Acxiom, Omniture, etc. and I have to put them someplace and ultimately join them with my other data
4. Now I am spending a lot **of time on ETL**; I have many developers probably tuning ETL as my sources grow and continually doing ETL to create new data sets. **I am expending a lot of resources on ETL**
5. I have my Hadoop, my operational store but I have interesting aggregates from 3rd parties. **Now I need something like a DW**; I do have my operational store but I need something away from my operational store **where I can a do discovery** and where I can add structure to my data and **connected to my Hadoop** system
6. Now my company has been successful, is growing and they need/want more applications. I have my DW, Hadoop, ingesting 3rd party data and **now executives want more applications built on-top of that data. So we hire more developers, build more ETL jobs, more processes** moving data to the front-end applications. Its becoming a mess
7. Now I **need BI, visualization** but this comes as an after **thoughts not architected** into the system

ticketmaster



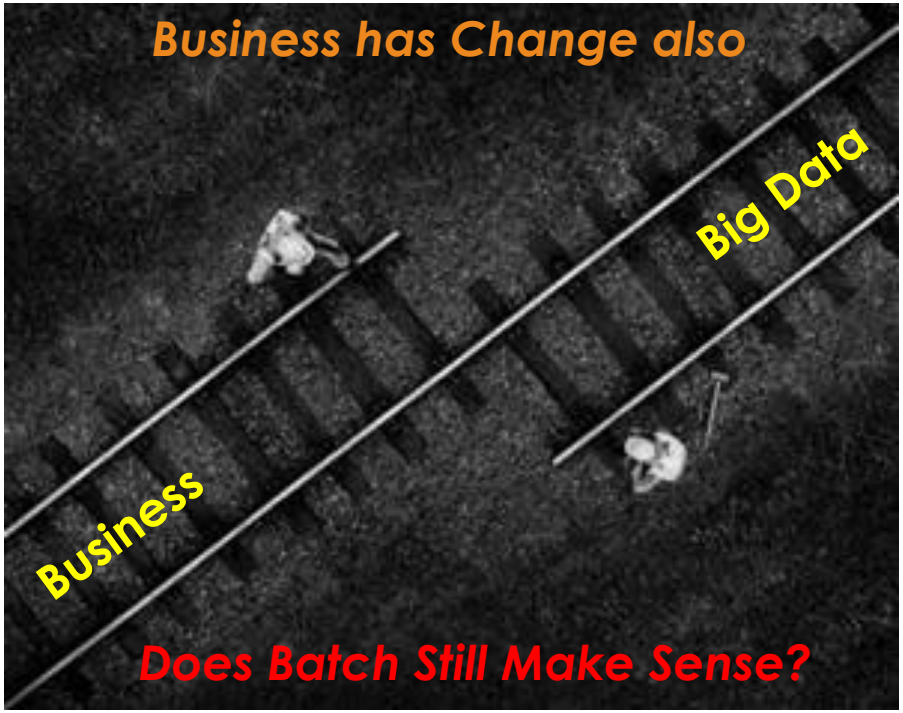
So this is what I end up with; **"A big Mish-mash of crap"** that is costly to support and doesn't actually support my going forward business needs.

John Carnahan: CTO

TERADATA.

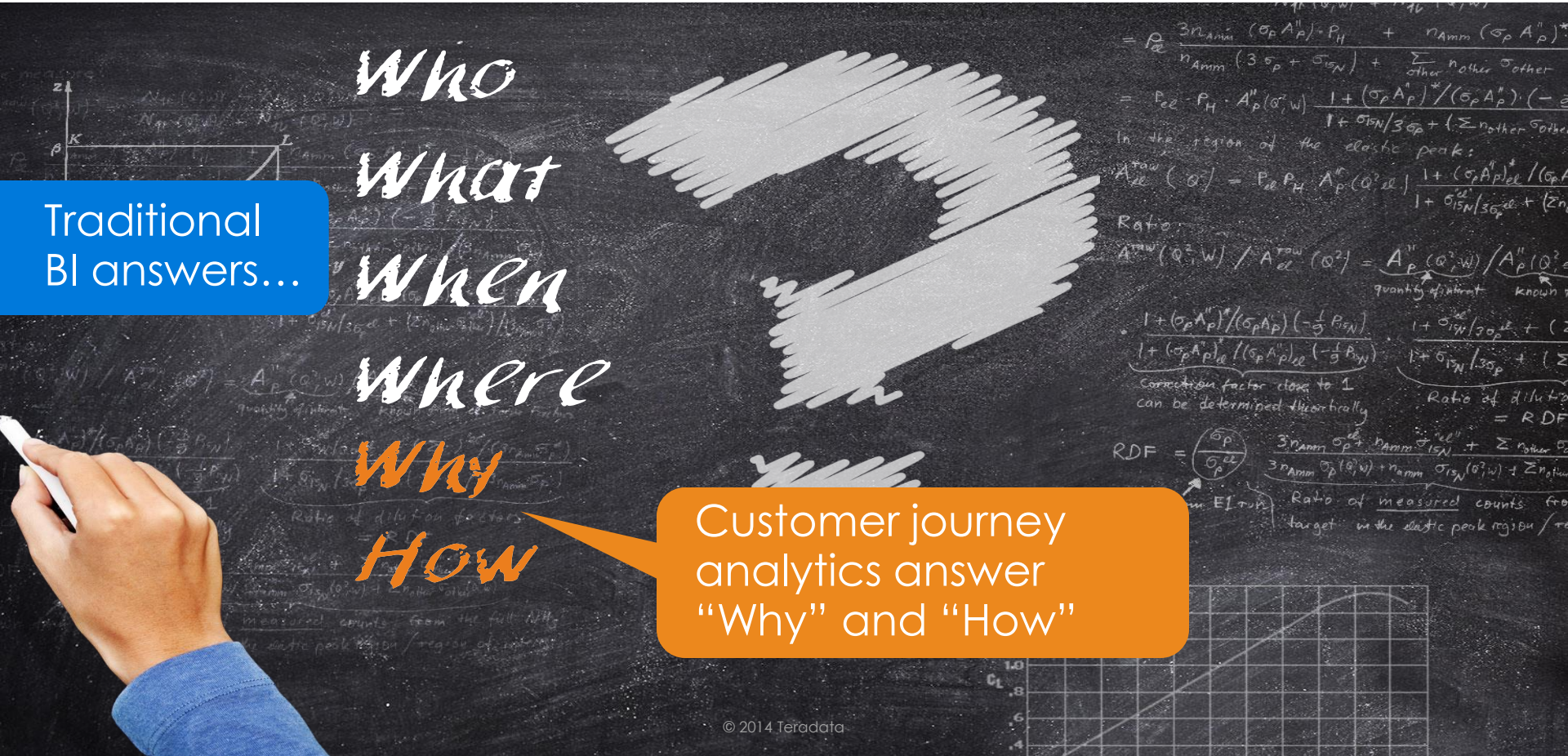
Big Data & Business are Not Aligned

- Continuous, 24/7/365
- Global
- Customer experience focused
- Product journey
- Requires experimentation
- Agile



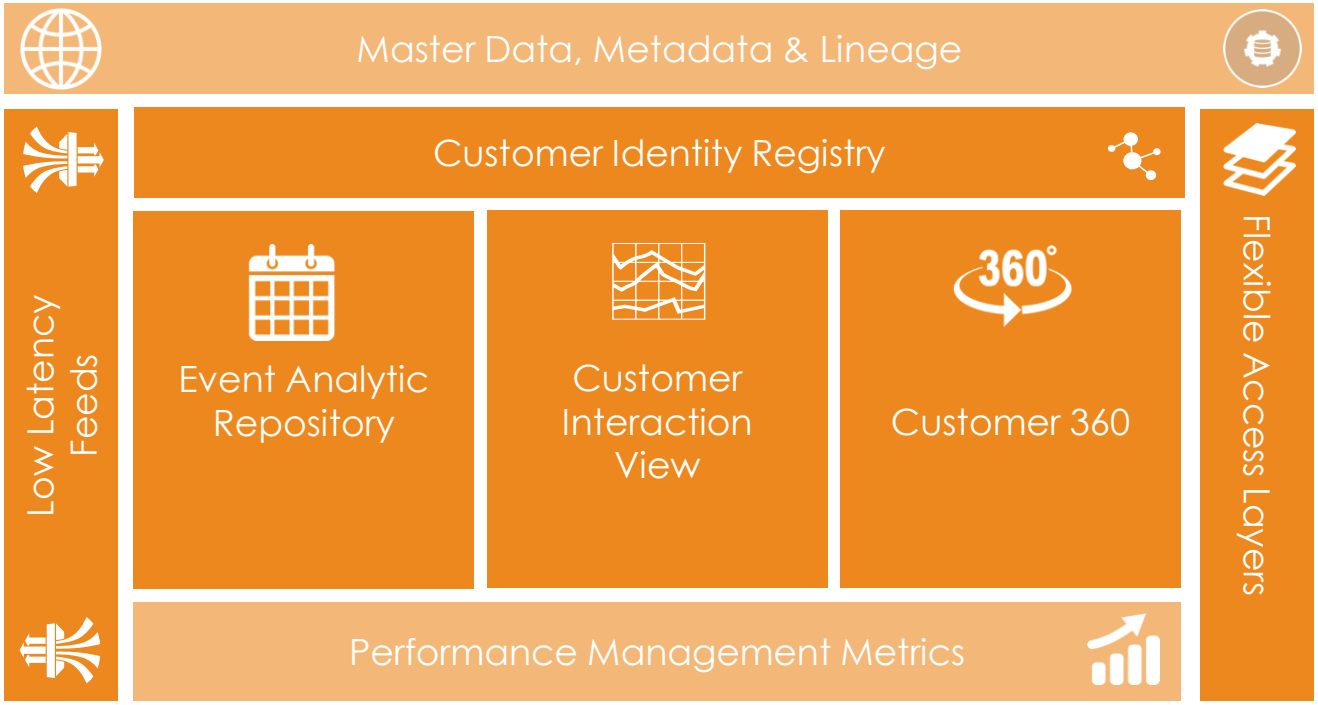
TERADATA.

The questions analytics are intended to answer…



Connected Data Capabilities

Scalable low latency architecture optimized for analytic & operational deployment



TERADATA.

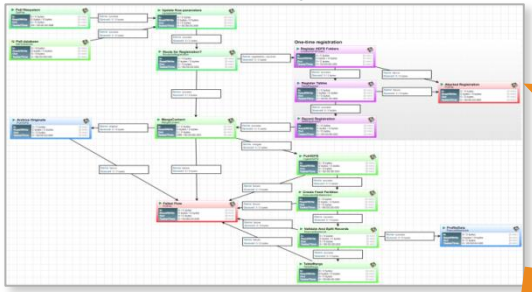
Low Latency Feeds

How we do it - Kylo

Moving from a batch paradigm to the the timely ingestion of many different types feeds – stored in their original fidelity as well as standardized into a defined event structure.

Granular event related data sources loaded in a timely manner via API's, streams and batch methods

Apache nifi

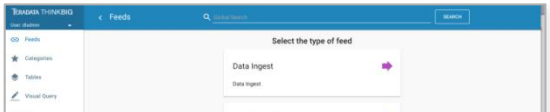


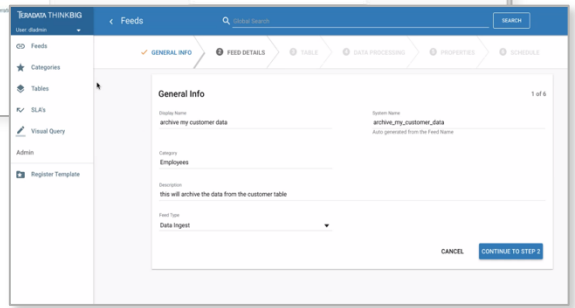
- IT Designers design models in NiFi
- Register into Think Big framework
- Integrated development process

register

deploy

Think Big Kylo





- Users configure new feeds in UI
- Feed deployed metadata + model
- Executed in NiFi

Feed types supported include API's, Streams & Micro Batch / Mini Batch / Batch workloads.

- Drag-and-drop pipeline design
- Dozens of data connectors
- 150+ pre-built transforms
- Data lineage
- Batch and Streaming
- Extensible

TERADATA.

Event Analytic Repository

How we do it

The Event Lake stores different types of events loaded from many different sources.

Events ingested and stored in their original fidelity from the source system – enables discovery analytics to be be performed on all the data if required.

Data can be re-processed if required to capture more detail from the original records

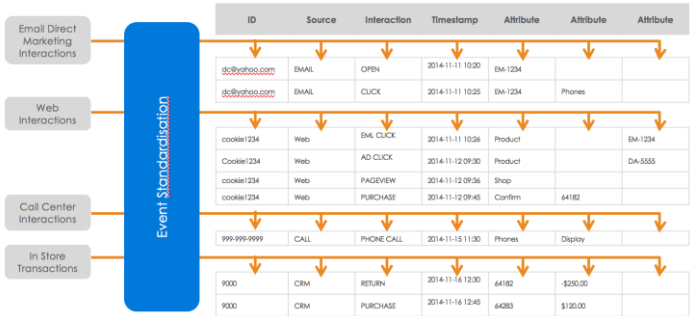
Event templates are created for many types of events – pre-filling the typical event attributes that should be captured for that sub-type

Flags denote sub-types for ease of filtering for downstream processes

3

The raw events are standardised to create a data structure optimised for rapid analytics and operational uses

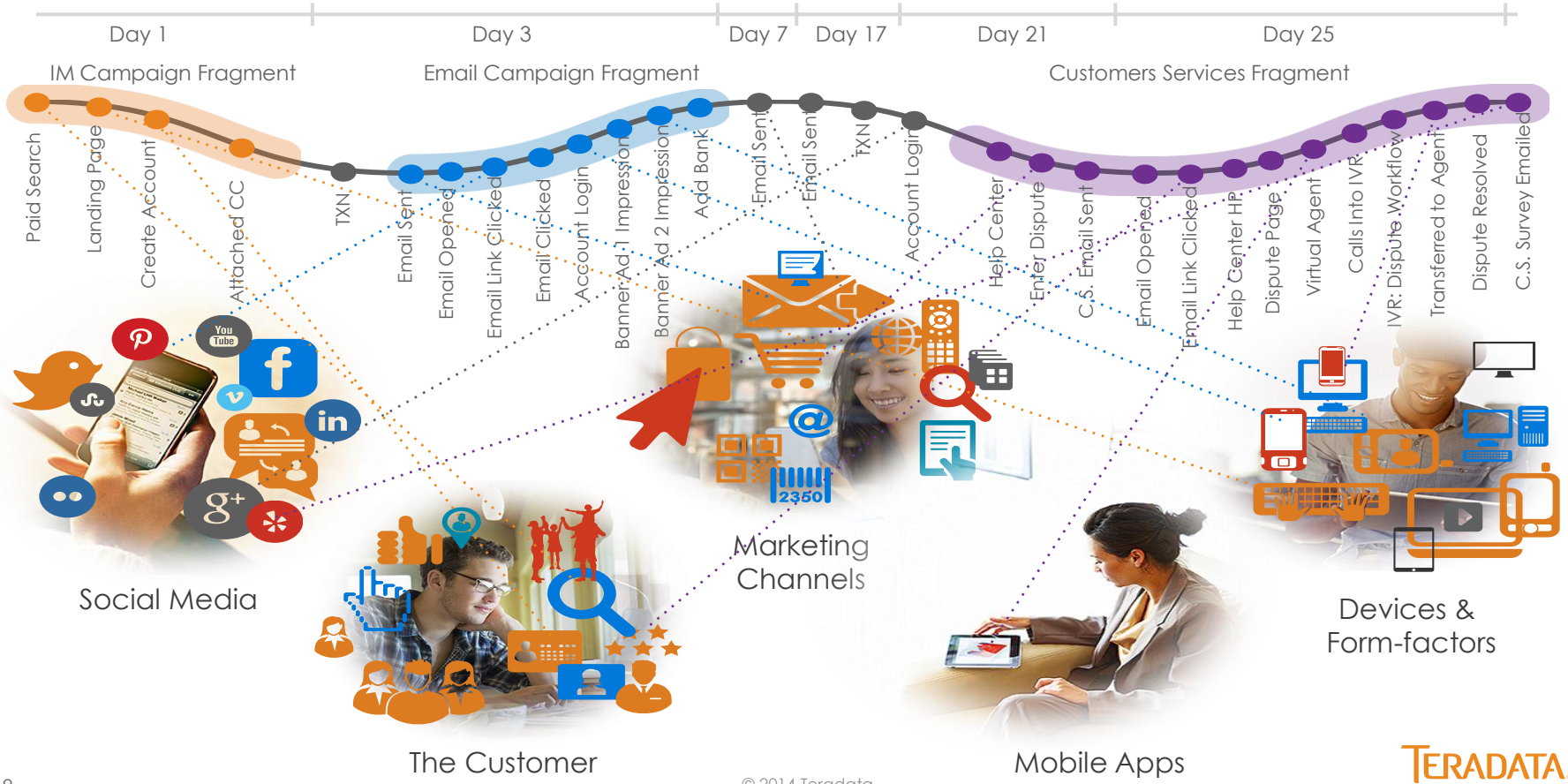
Processes such as session-isation will occur at this stage and key attributes will be captured - dependent on the event type



An event will always be linked to a master customer record created by the Expanded SCV

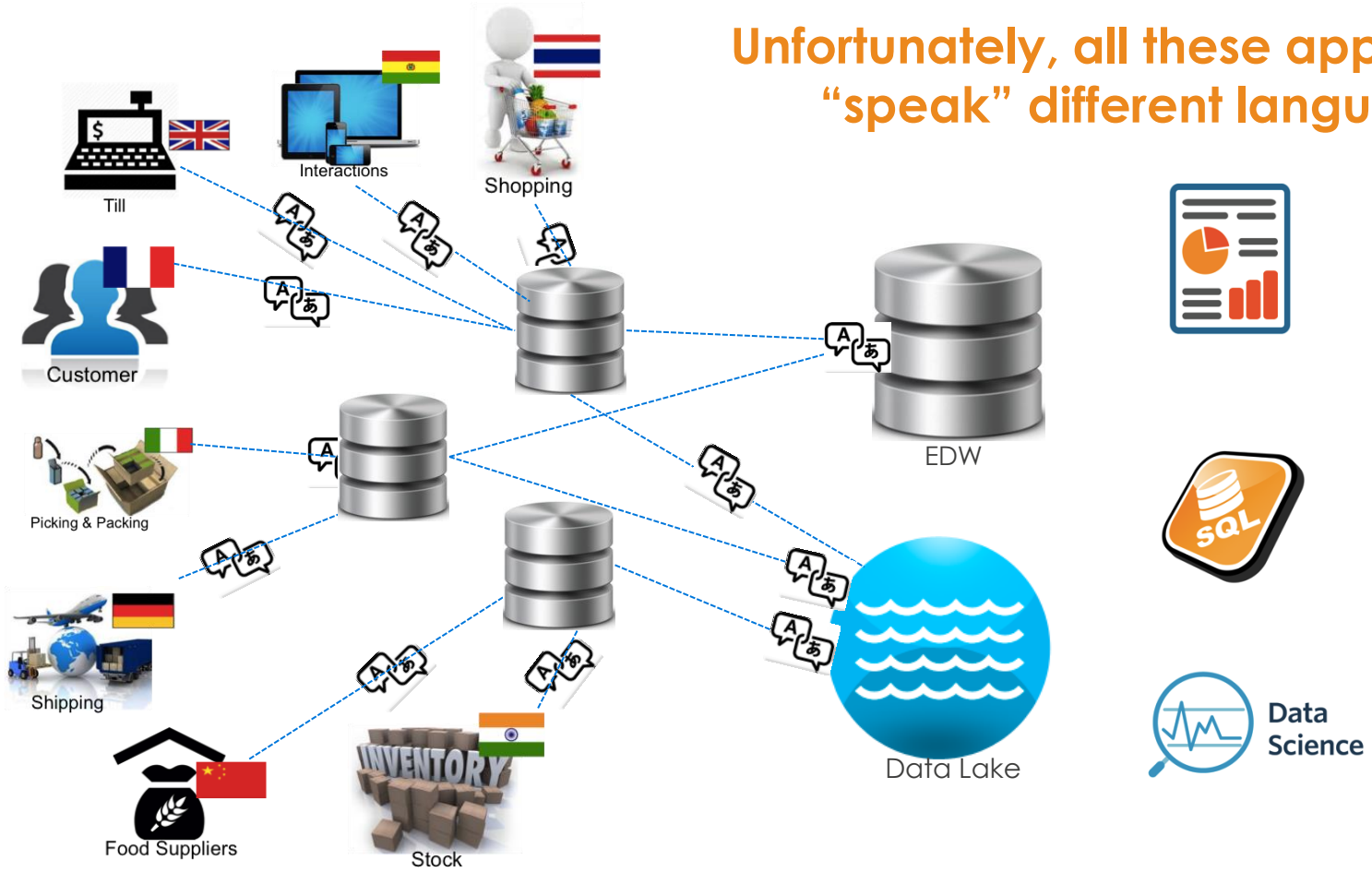
Repository to hold event data in original fidelity and standardised for operational / analytical usage

Customer Journeys span lots of “applications”



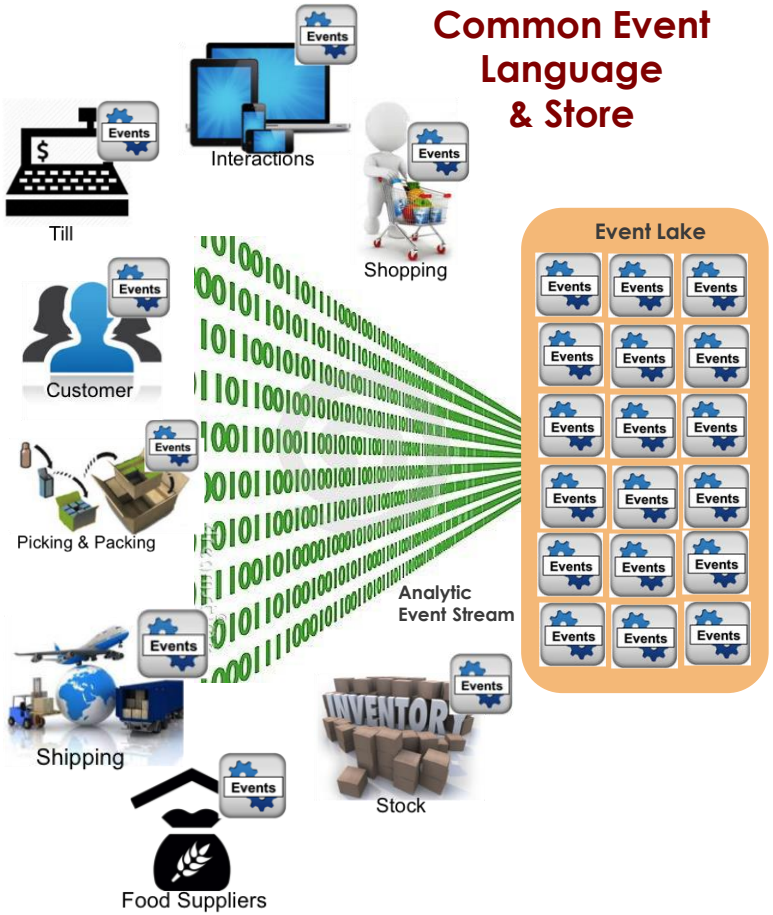
TERADATA

Unfortunately, all these applications "speak" different languages

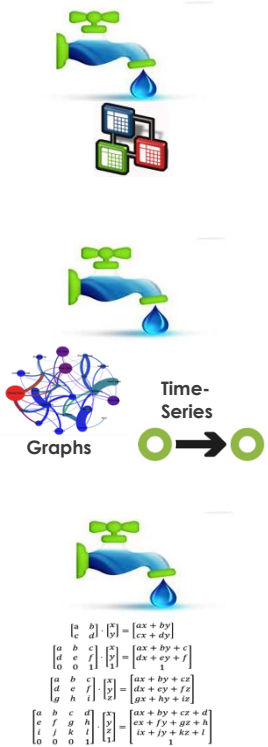


TERADATA.

Common Event Language & Store



Self-service Hydration



Analytics

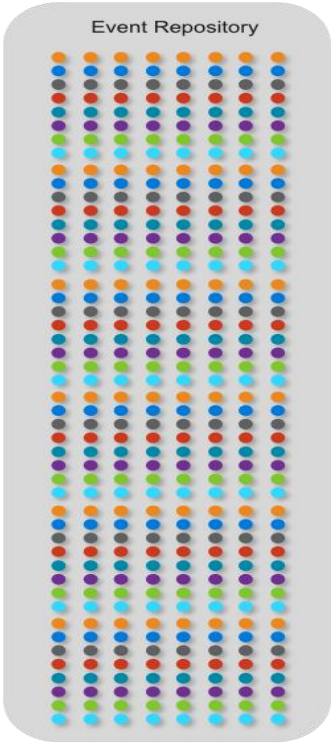
Data Science

TERADATA.

What's needed

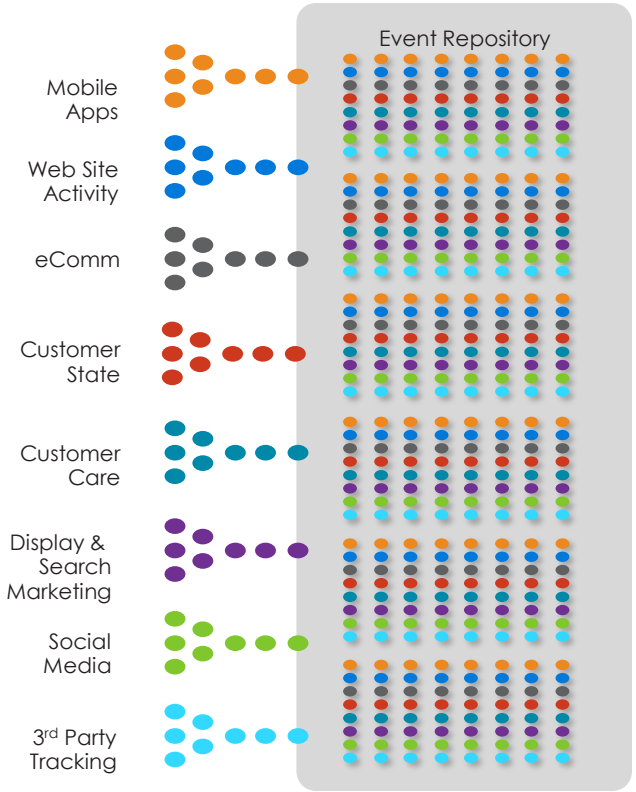
- Single source of **ALL** data, at atomic level
  - Full ingest, clean & store data in its original raw form
- Common event language for analytics
  - Data formed, at source, into events
- Event Lake
  - Fully indexed set of events w/ metadata repository; Searchable, Queryable
- Business access that is:
  - Stable, governed & trusted; low latency, self-service
- Greatly expand discovery & data science analytical capabilities
  - Allow analysts to form their own data
  - Expanded data structure types: NoSQL, graph, etc.
  - Expanded analytic tools: Machine Learning
- Governed
  - Standard, governed semantic layer for Company-wide metrics & KPI's

Event Lake



- Single source of **ALL** data, at atomic level
  - Full ingest, clean & store data in its original raw form
- Common event language for analytics
  - Data formed, at source, into events
- Event Lake
  - Fully indexed set of events w/ metadata repository; Searchable, Queryable
- Business access that is:
  - Stable, governed & trusted; low latency, self-service
- Greatly expand discovery & data science analytical capabilities
  - Allow analysts to form their own data
  - Expanded data structure types: NoSQL, graph, etc.
  - Expanded analytic tools: Machine Learning
- Governed
  - Standard, governed semantic layer for Company-wide metrics & KPI's

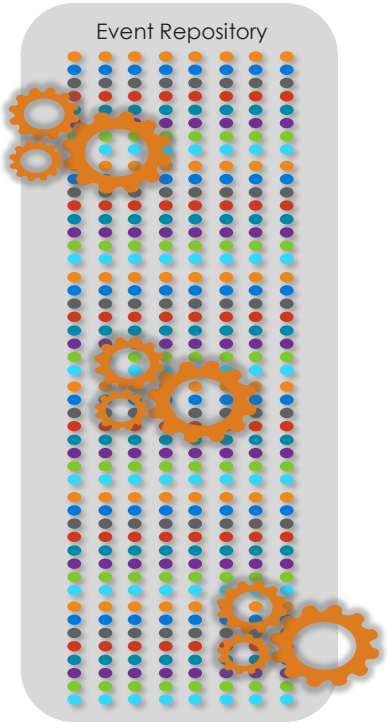
An Event Lake is the perfect landing zone for these structures



- XML, JSON and other semi-structured data formats
- Application reference data helps identify the same person across channels and devices
- Enables the "stitching together" of individuals usage across multiple different applications and data sources

TERADATA.

Real-time analytical routines enable interactions

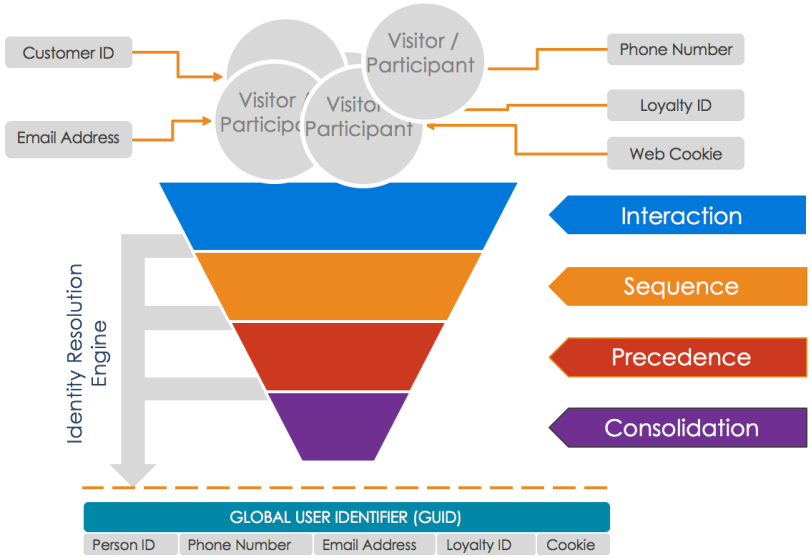


- Machine learning algorithms are used to:
  - Detect hidden patterns in data
  - Create useful predictions about unseen data
  - Decision making under uncertainty
- The Event Analytics Repository provides the universe of customer events; a trusted set of events
- Machine Learning algorithms can continuously search through the Event Repository looking for complex patterns of interesting behavior; triggering actions

Customer Identity Registry

How we do it

The Single Customer View is expanded (from the traditional internal known customer definition) using an Identity registry to match the identifiers that a customer is addressed by across different internal and external channels



A Global User ID is generated that masters all the internal and external identifiers – expanding the Single Customer View across an increased number of channels

Process

- **Every new event loaded represents an opportunity to expand the single customer view.**
- **An identity registry masters the customer view to ensure we always able to link the widest set of events and create the completest view of the customer**
- **Data sources are processed in sequence based on their reliability for associating interactions with a customer**
- **As soon as new information is captured it may be possible to create a new link and convert the status of the identity from unknown to known**

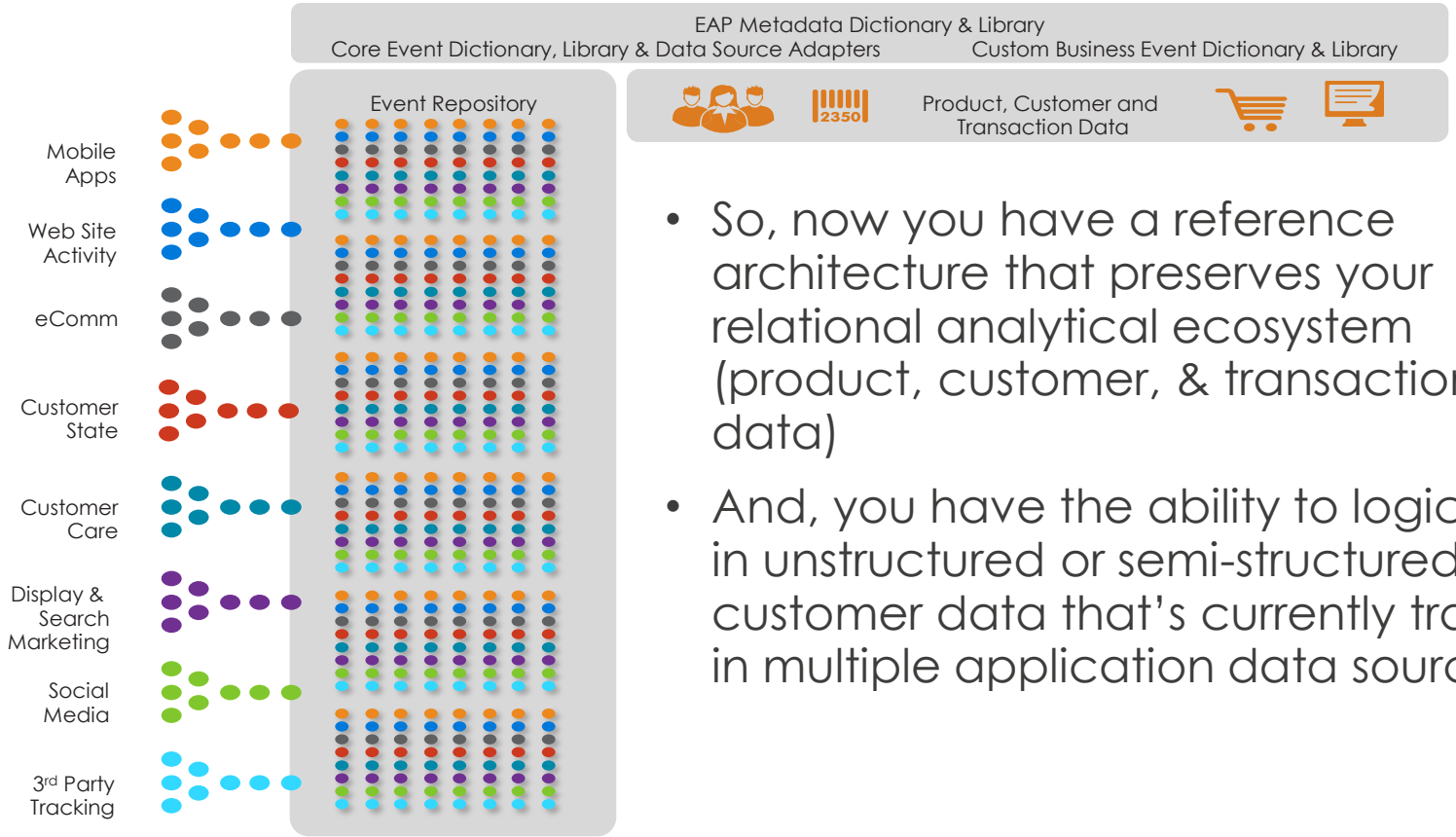
Expanded view of all the internal / external customer identifiers that an organisation is required to interact with

Different states of a customer identity

- **Prospect – known consumer identity using PII information – no existing relationship with organisation**
- **Applicant – known consumer identity using PII Information – customer in process for applying for first product**
- **Known Customer – aligning to internal definition of customer**
- **Anonymous Visitor – consumer interacting with digital channels – status of relationship unknown**

TERADATA

Coexistence with your relational data is critical



- So, now you have a reference architecture that preserves your relational analytical ecosystem (product, customer, & transactional data)
- And, you have the ability to logically link in unstructured or semi-structured customer data that's currently trapped in multiple application data sources

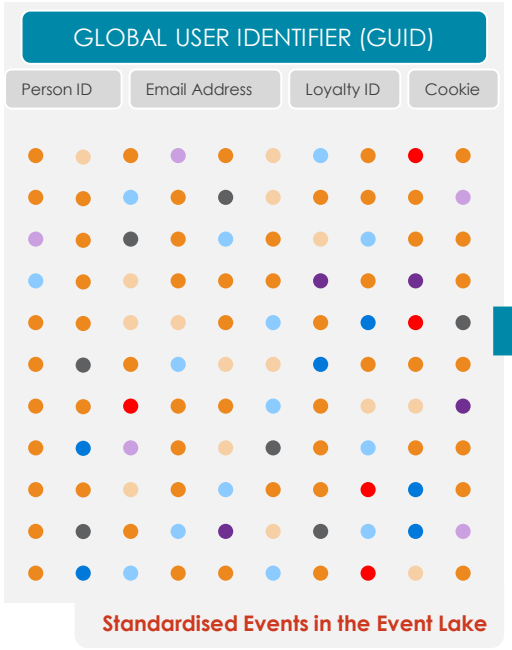
Customer Interaction View

How we do it

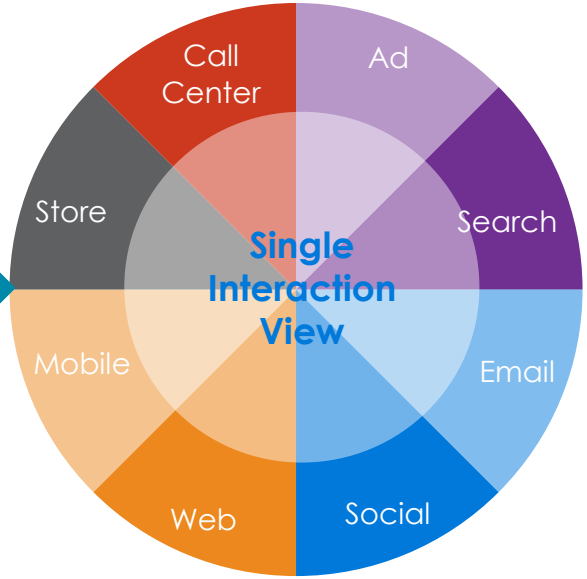
The customer interaction view integrates multiple sources of channel data that are typically developed and held independently of each other

Integrated view of customer interactions across all channels and types of contact

All events keyed against expanded single Customer View created using identity registry



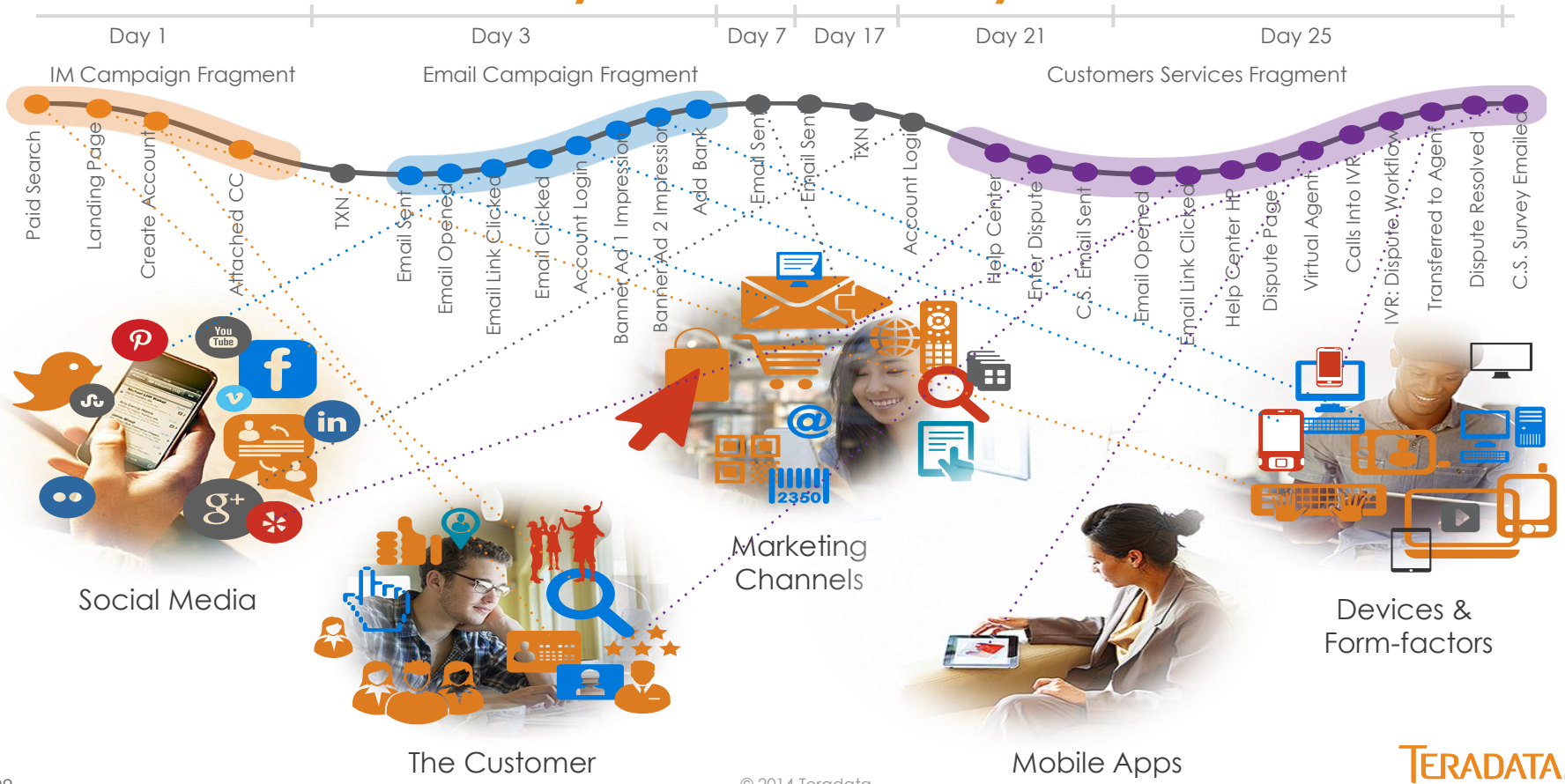
The Customer Interaction View creates a unified picture of how a customer interacts with the organisation across all channels and all types of contact



- **Creating a common summary across channels at a customer level**
- **Utilising the Identity Registry Global User ID to resolve different identities used by customers**
- **Integrating diverse digital media sources across paid, earned and owned media to allow more complete customer journeys to be analysed and optimised.**
- **Enabling more accurate measurement of the contribution of each channel through the use of marketing attribution techniques**

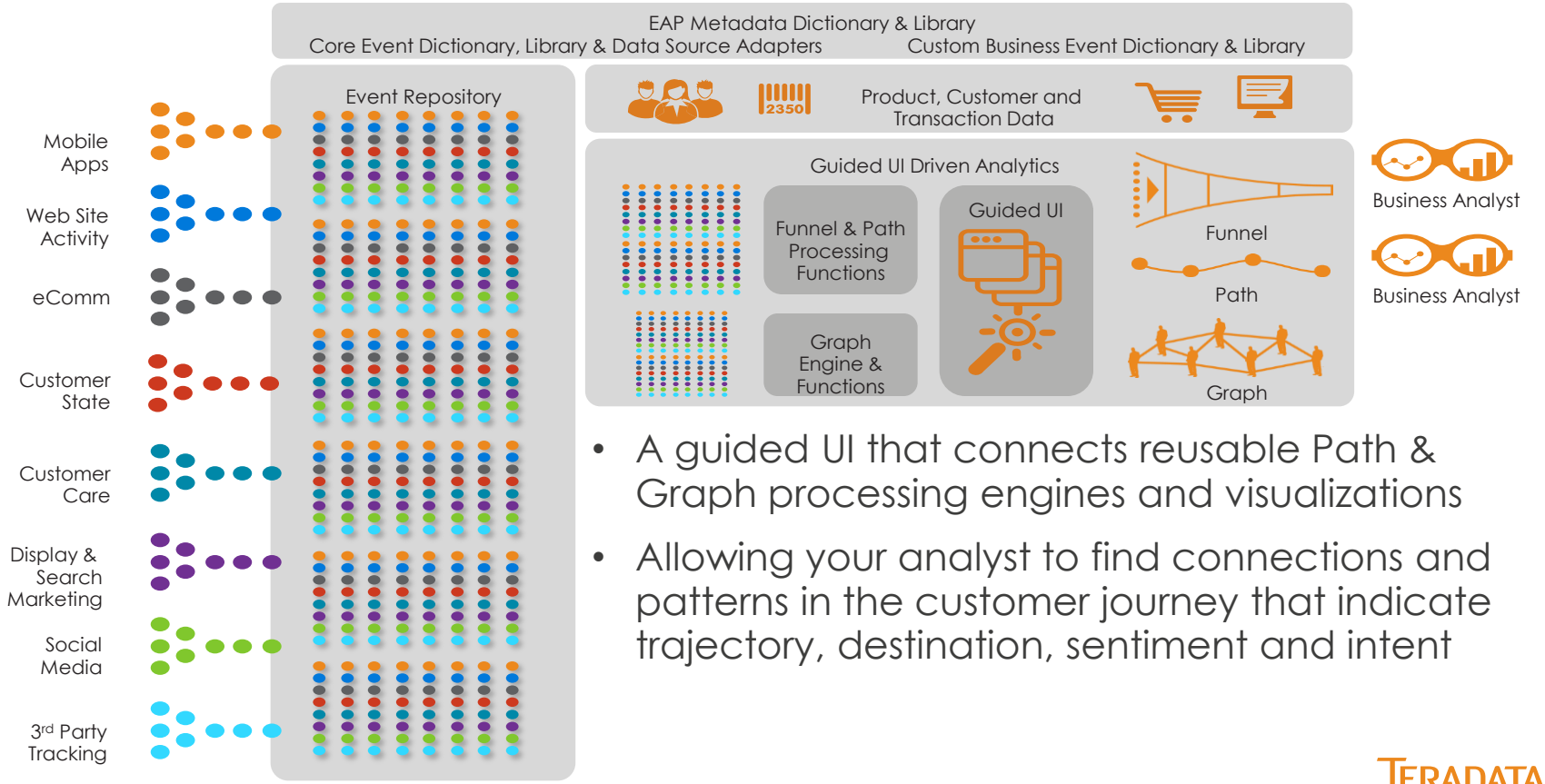
TERADATA.

The Customer Journey is connected by definition



TERADATA

Understanding why & how a customer experiences you



- A guided UI that connects reusable Path & Graph processing engines and visualizations
- Allowing your analyst to find connections and patterns in the customer journey that indicate trajectory, destination, sentiment and intent

Customer 360

How we do it

The 360 customer view becomes augmented to a new scale and level of granularity. The customer profile is able to expand to tens of thousands of attributes with the ability to specialise content by channel and use case.

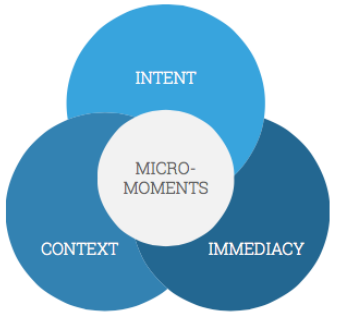
Store of derived event associated attributes for use by downstream analytical applications

The Customer 360 expands to store thousands of attributes that are derived from data in the event lake

- Data structure and format optimised for analytical consumption – de-normalised to minimise the amount of data wrangling required
- Attributes specialised to specific events or combination of events
- Rich metadata allows rapid search for candidate attributes
- Ability for data scientist to publish new attributes without need for IT intervention
- Automated delta refresh process

The flexible feature store provides the optimised input layer to some analytical processes

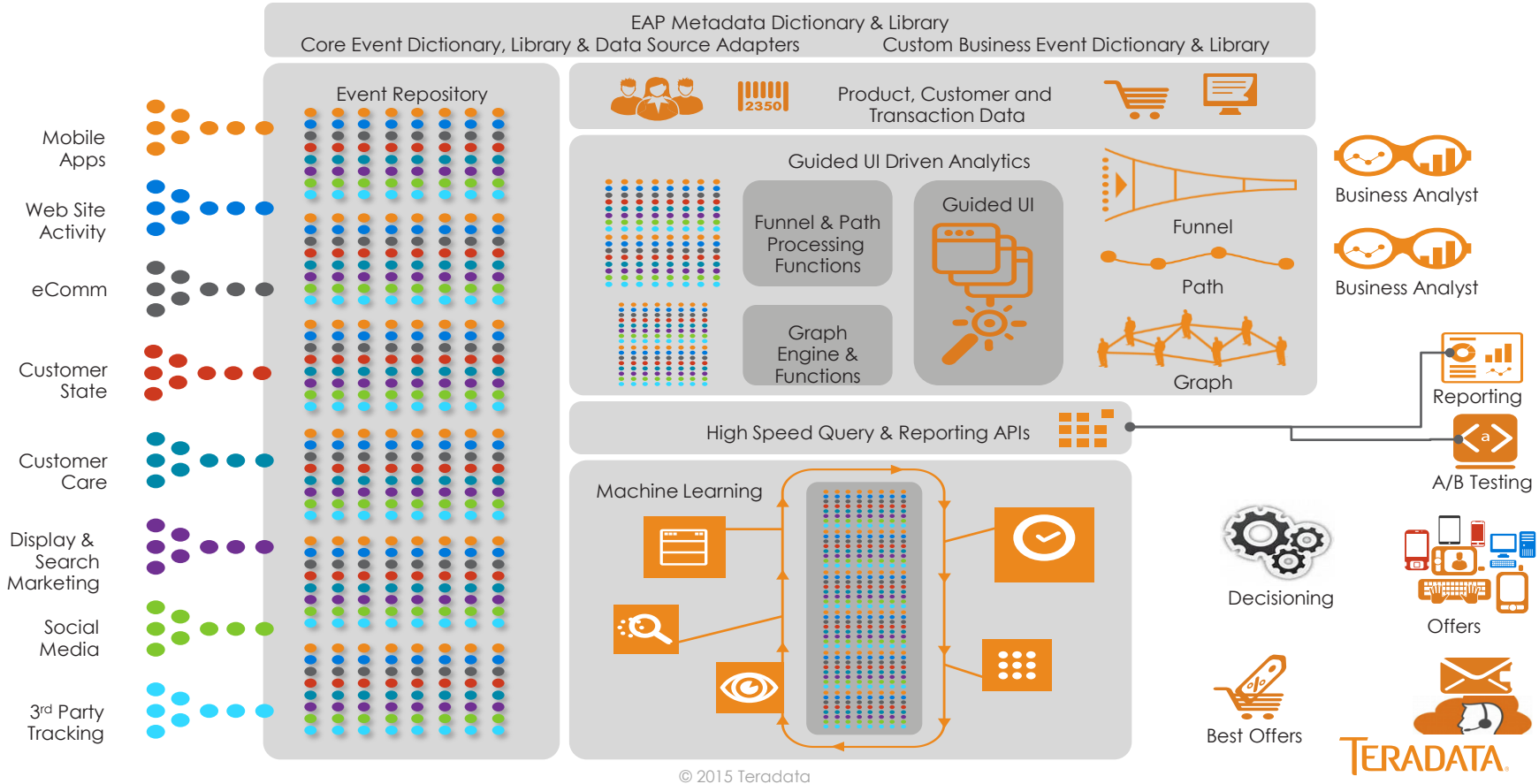
- Customer Experience & Journey Analytics
- Self Learning Analytics
- Actionable Analytics



The augmented customer 360 provides attributes that unlock new behavioural insights to add context, understand intent and improve relevance of interactions.

TERADATA.

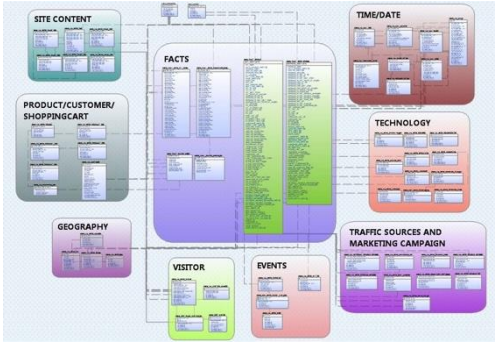
Machine Learning uses the same connected data



Performance Management Metrics

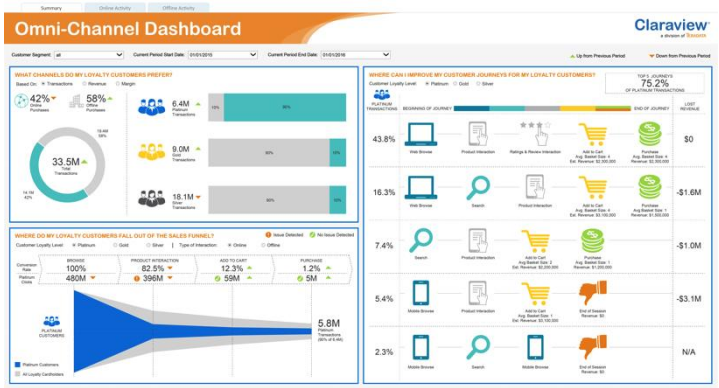
How we do it

A centralized store of KPI's that allows the rapid construction of Dashboards, Reporting & Interactive Management Information from raw and derived data building blocks contained in the Customer Journey Solution



- Data structures optimised for rapid build of performance management reports via off the shelf BI tools
- Shorten time to market for new types of reporting/KPIs
- Automates the refresh of all reports

- Uses standard definitions to ensure performance reporting outputs are consistent
- Utilises master data sources to provide consistent definition of organisation, channels & products
- Incorporates analytical outputs to understand complex patterns and forecast future trends



Store of business performance measures used for Customer Journey reporting – operational and business KPI's

Provide users self serve capability via interactive query tools

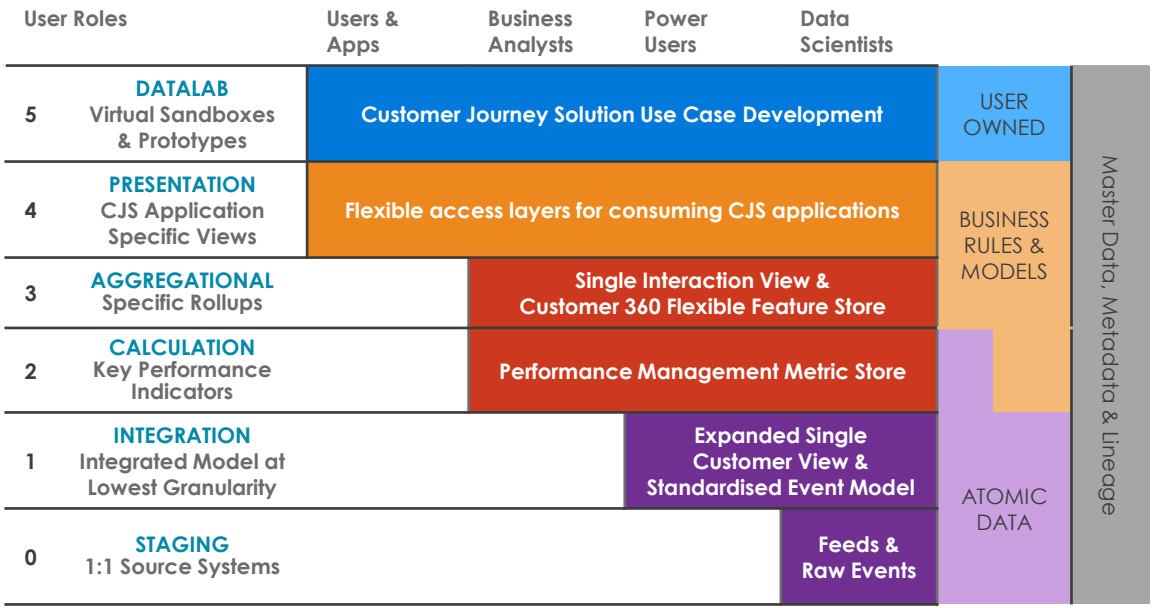
TERADATA

Flexible Access Layers

How we do it

The Customer Journey Solution employs an agile data platform approach - having discrete layers of access and permissions that play different roles in the ecosystem. Once new applications are developed in the Data lab layer, they can be rapidly productionised in the appropriate data layers.

Logical and physical views of customer journey data – optimised for downstream analytical consumption



Data layers exist cross platform and cross technology to serve data to the other connected elements of the customer journey solution

- Wrangled data to drive Connected Analytics
- Conformed dimensions to power Connected Interactions

Connected Data Ecosystem Technologies

Connected Data Element	Teradata	Enterprise Vendor	Open Source
Low Latency Feeds	Listener v2.0	CEP (TIBCO), Vendor ETL / ELT Celebrus, 02MC	Kylo
Event Analytic Repository	TD Warehouse Aster	Hadoop	Hadoop
Customer Identity Registry	TD Warehouse Master Data Management (MDM)	Informatica, Ab Initio	Kylo
Customer Interaction View	TD Warehouse Integrated CIM/RTIM Contact History	Hadoop	Hadoop
Customer 360	DS Generator, Warehouse Miner Teradata Analytic Calculator (TAC)		Kylo
Performance Management Metrics	Teradata LDM SMBB's	BI Vendors - Tableau / QLIK etc.	Dashboard Engine for Hadoop
Flexible Access Layers	QueryGrid 2.0 Restful API	HVR Software - replication	Presto
Master Data, Metadata & Lineage	Master Data Management (MDM)	Alation, AB Initio Graphs Informatica - Live Data Map/ Enterprise Information Catalog etc.	Wherehows Kylo – Feed Metadata / GCFR Kylo – JCR Metadata Repository

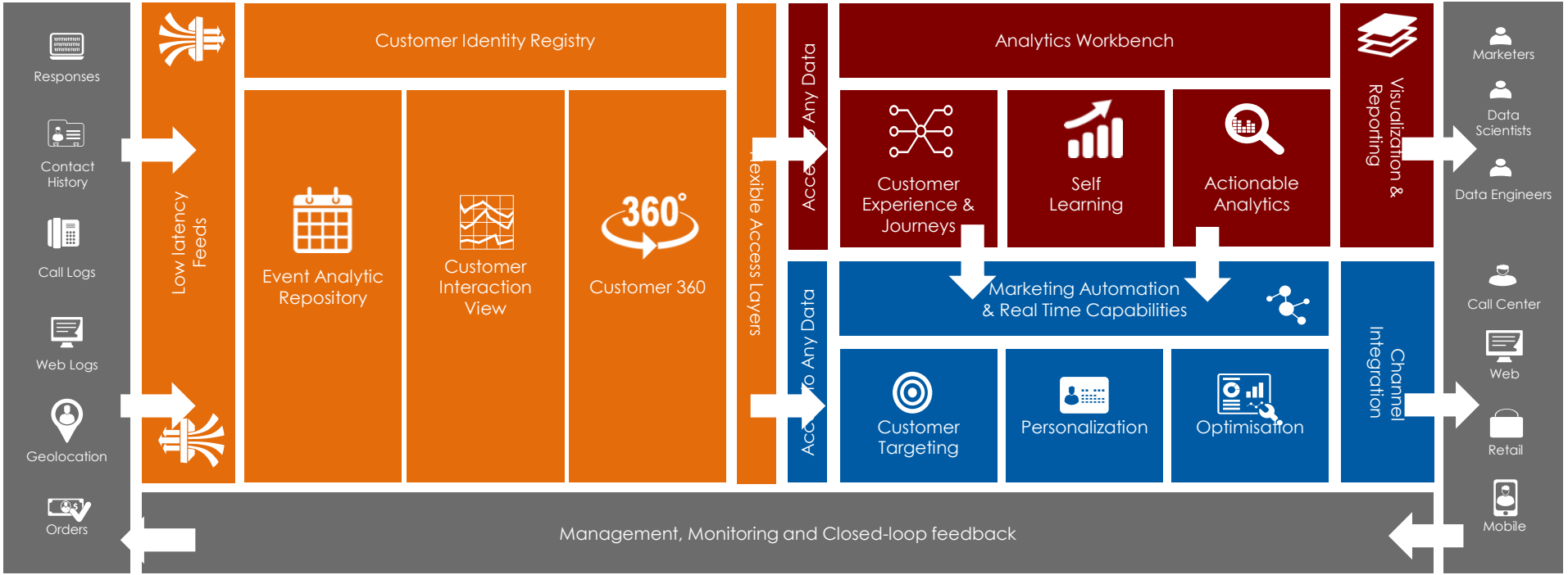
Connected Data Ecosystem Technologies

Most likely deployment options

Connected Data Element	Teradata	Enterprise Vendor	Open Source
Low Latency Feeds	Listener v2.0	CEP (TIBCO), Vendor ETL / ELT Celebrus, 02MC	Kylo
Event Analytic Repository	TD Warehouse Aster	Hadoop	Hadoop
Customer Identity Registry	TD Warehouse Master Data Management (MDM)	Informatica, Ab Initio	Kylo
Customer Interaction View	TD Warehouse Integrated CIM/RTIM Contact History	Hadoop	Hadoop
Customer 360	DS Generator, Warehouse Miner Teradata Analytic Calculator (TAC)		Kylo
Performance Management Metrics	Teradata LDM SMBB's	BI Vendors - Tableau / QLIK etc.	Dashboard Engine for Hadoop
Flexible Access Layers	QueryGrid 2.0 Restful API	HVR Software - replication	Presto
Master Data, Metadata & Lineage	Master Data Management (MDM)	Alation, AB Initio Graphs Informatica - Live Data Map/ Enterprise Information Catalog etc.	Wherehows Kylo – Feed Metadata / GCFR Kylo – JCR Metadata Repository

Customer Journey Overview

A low latency architecture optimized for deploying operational analytics



TERADATA.

Parting Thoughts

- CTO believes in dramatic changes to how data is collected, managed & used; a new approach to data
- Understands that the current data organization is not skilled or sized to support this vision
- Hadoop is not "operational-quality" today; but will be
- A need to find your new guiding principles for data
- Truly listen to business needs
- Don't re-build what is already working
- Question old frameworks & processes;
  - they tend to lead to incrementalism
- Commoditize reporting