Database Answers

Holy Trinity Church, Claygate, UK

The Power of Data Platforms May 10th

Barry Williams
barryw@databaseanswers.org

1. Why did I write this book ?

My purpose is to summarise my experience over 20 years in a format that anybody can use, refer to and contribute to.

Therefore, over a period of time it can be used as a reference that will evolve to always reflect our experience.

If you would like to add comments based on your experience and your opinions I would be very pleased to hear from you.

You can reach me on my email address given above.

In the meantime, I hope you enjoy this book and find it helpful.

2. Theory

2.1 Introduction

Platforms contain Layers.

Lower level s are basic data and higher levels are more user-friendly

Techopedia has a useful definition here :-

- https://www.techopedia.com/definition/3411/platform

which says (in part)

"A platform is a group of technologies that are used as a base upon which other applications, processes or technologies are developed."

From this we start to see the emergence of Layers as part of a Platform.

Wikipedia has a good definition for a Computing Platform on this page :-

- https://en.wikipedia.org/wiki/Computing_platform

which defines a Platform to be :-
> "a framework on which applications may be run"

I would add clarification by saying :-
> "a framework with clearly-defined Layers on which applications may be run
> by end-users with Analytics and User-profiling"

I have implemented an approach that I call 'The Barry's Secret Sauce'.

It contains generic ingredients that can be provided in different flavours by different Suppliers, such as IBM, Oracle, Salesforce and Teradata.

In this White Paper, I present my approach to Data Platforms.

My Platforms incorporate 'Barry's Secret Sauce', which includes these ingredients :-

1   Canonical Data Model
2   CRUD Services (Create, Read, Update, Delete)
3   Customer Data Platform
4   Design Patterns
5   Events
6   Functional Layers
7   Generic Data Platforms
8   Industry-specific Platforms (Banking, etc)
9   Inheritance
10   Layered Data Architectures
11   Mapping (eg Passengers to Customers)
12   Model-View-Controller (MVC)
13   Publish-and-Subscribe
14   Reference Data Architecture
15   Reporting Toolkit (Patterns of Generic Reports and KPIs)
16   Rules Engine
17   SOA Facilities
18   Transaction/Master and Reference Data
19   Triggers (eg for KPIs)
20   User-Defined Hierarchies

2.2 What is a Platform ?

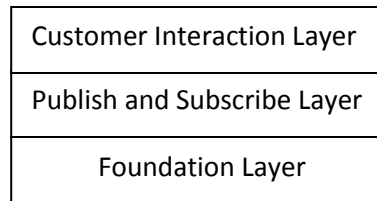The term 'Platform' is commonly used but without an exact definition of its meaning.

My definition is that a 'Data Platform ' is made up of Layers of Data which are logically related and become more 'user-friendly' at the top.

In this White Paper, I will discuss three examples of Platforms as Design Patterns :-
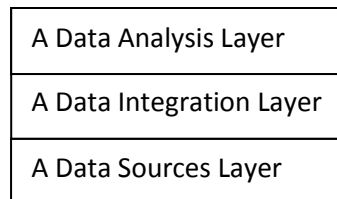1.   A Customer Data Platform
2.   A Functional Data Platform
3.   A Data Platform
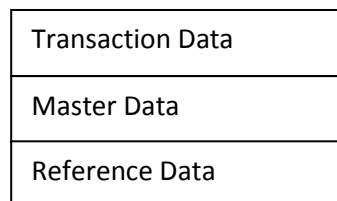
The Layers in each Platform are shown below.

2.2.1 A Customer Data Platform

Customer Interaction Layer
Publish and Subscribe Layer
Foundation Layer

2.2.2 A Functional Data Platform

A Data Analysis Layer
A Data Integration Layer
A Data Sources Layer

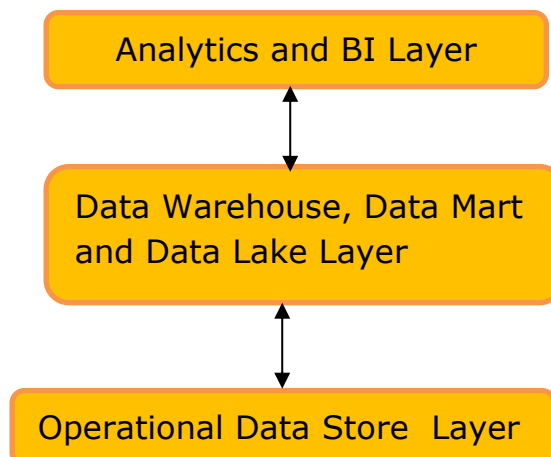### 2.2.3 An Operational Data Platform

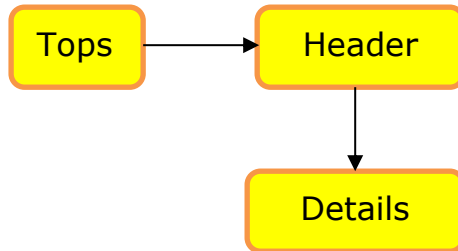| |
|---|
| Transaction Data |
| Master Data |
| Reference Data |

2.3 Our Generic Platforms

2.3.1 Data Platform

In our Generic Data Platform there are three Layers in our top-to-bottom view which hold quite different types of Data :-
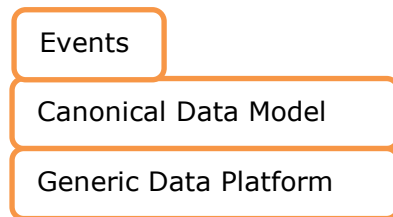
2.3.2 Publication Platform

In our Generic Publication Platform there are two Layers in our top-to-bottom view which hold quite different types of Data :-
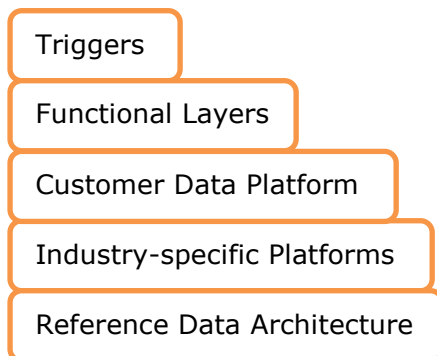
```
┌──────────┐      ┌──────────┐
│   Tops   │─────▶│  Header  │
└──────────┘      └──────────┘
                       │
                       ▼
                  ┌──────────┐
                  │ Details  │
                  └──────────┘
```

2.4 Implementation Plan

There are three Phases in the Implementation Plan.

2.4.1 Beginning Phase 1

- Events
- Canonical Data Model
- Generic Data Platform

2.4.2 Intermediate Phase 2

- Triggers
- Functional Layers
- Customer Data Platform
- Industry-specific Platforms
- Reference Data Architecture

2.4.3 Advanced Phase 2

- Event   Rules Engine
- Publish-and-Subscribe
- Trans/Master/Ref Data
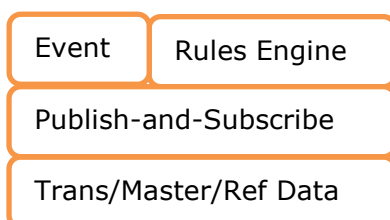
2.5 Our Industry-specific Platforms

This shows details of Platforms and their role.

Industry-specific   Platforms

Banking Platform	Logistics Platform	Retail Platform
Banking Model	Logistics Model	Retail Model

Generic   Platform

Canonical Data Model
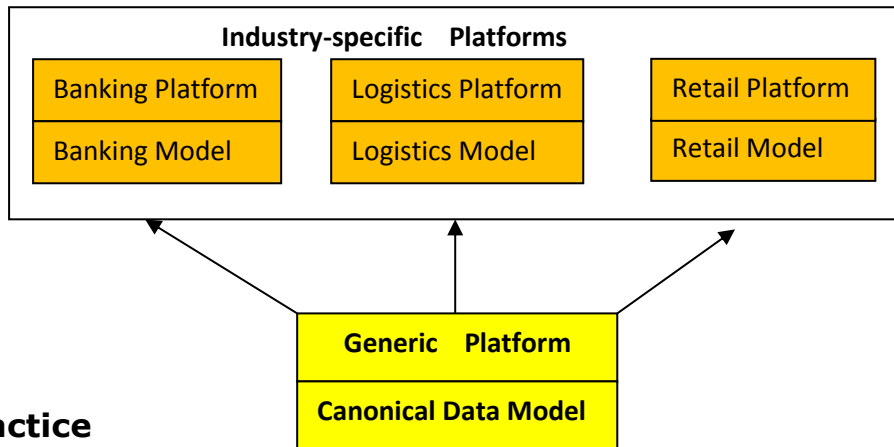
3. Practice

3.1 Anatomy of a Platform

3.1.1 Layers

A Platform is defined by a series of Data Layers.

LAYER	Data Types	Banking	Customer-Services	Logistics	Retail
Transaction		Payments			Purchases
		Withdrawals			
CRUD					
TOI	Customers	Yes	Yes	Yes	Yes
Master Data	Customers			Yes	
	Products			Yes	
	Schedules			Yes	
	Services	Yes	Yes	Yes	Yes
Ref Data	Account Types	Yes			
	Customer Types	Yes	Yes	Yes	Yes
	Payment Methods	Yes	Yes	Yes	Yes

3.2 The importance of Events

Events are a major part of our Canonical Data Model that we discuss in the next Section.

In 2003, Gartner Research introduced a new approach called Event-Driven Architecture which is sometimes called an 'Event-Driven SOA'.

It is featured in this Wikipidia entry :-

- https://en.wikipedia.org/wiki/Event-driven_SOA

Here is a diagram of an Event-driven SOA from the Oracle Web Site :-

- http://www.oracle.com/technetwork/articles/soa/schmutz-soa-eda-405955.html
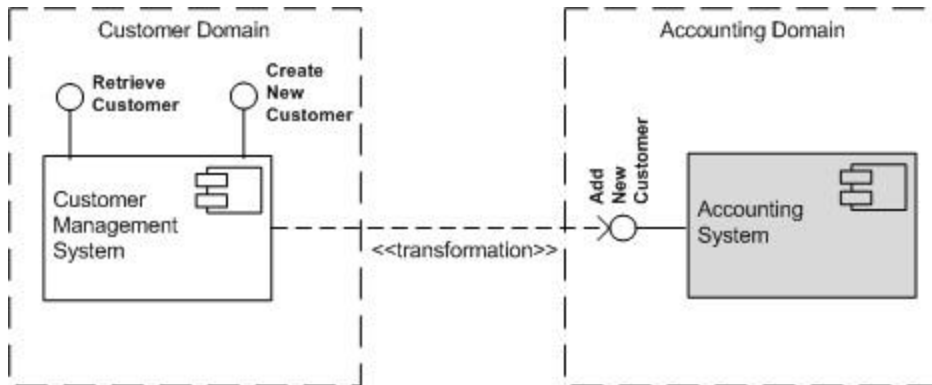
We discuss this, with Business Activity Monitoring ('BAM'), on my Database Answers Web Site :-
- http://www.databaseanswers.org/data_models/bam_event_driven_architecture/index.htm
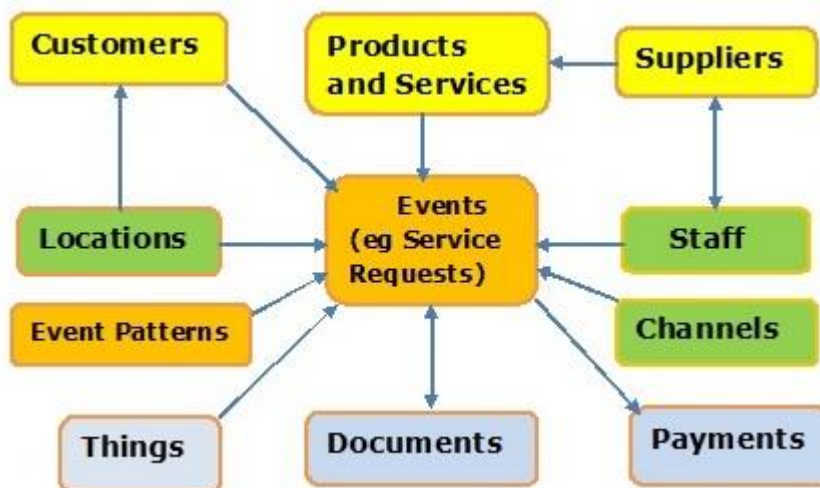
3.3 Step 1 – The Canonical Data Model

I start with the Canonical Data Model.

This is discussed in this page of my Database Answers Web Site :-

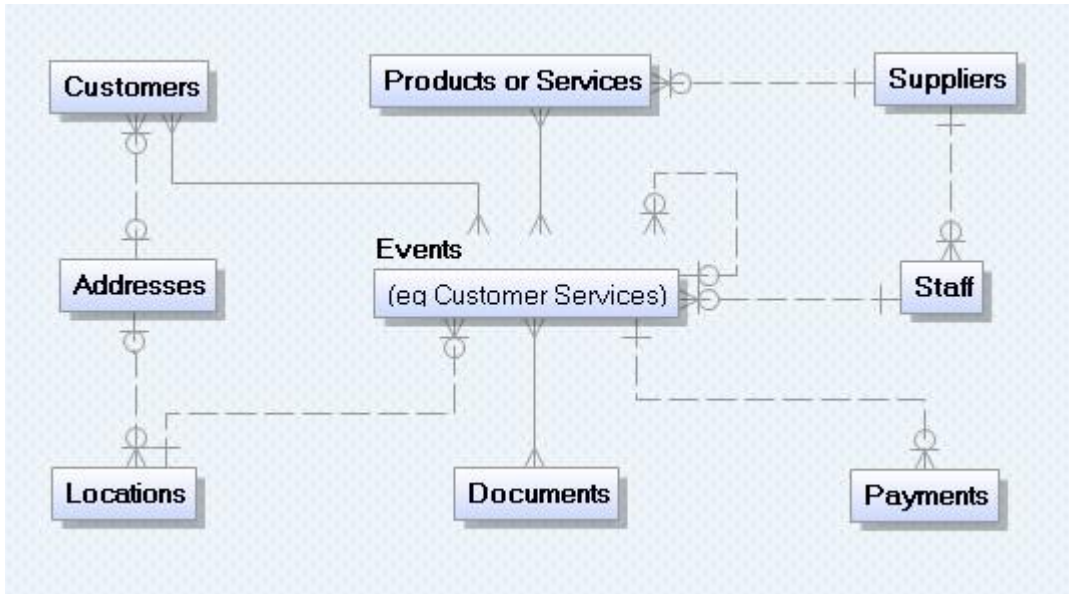- http://www.databaseanswers.org/data_models/canonical_data_models/index.htm

It takes different forms :-

- A Conceptual Model intended for business users :

- Here is a Logical Model intended for Data Analysts and DBAs :

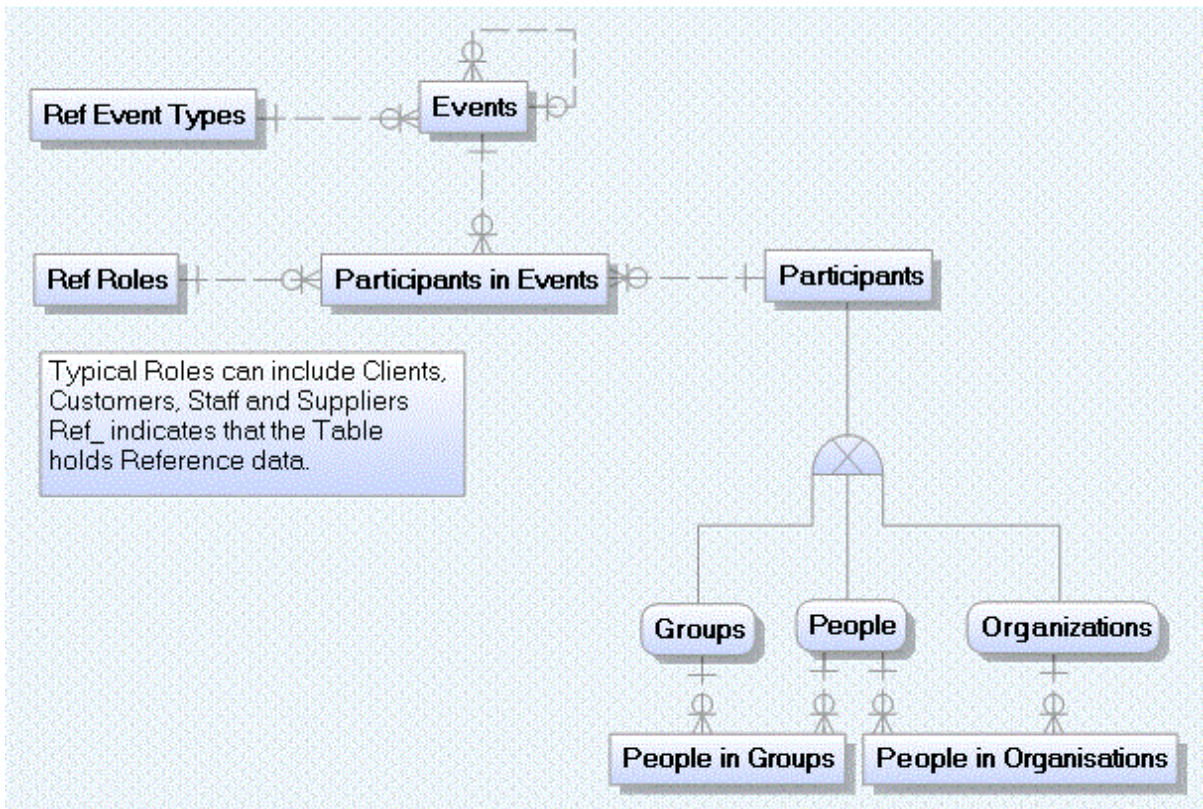At this level, we also show a way for handing Parties, (that we call Participants) that professional Data Modellers favour over Customers, Suppliers and so on.

It is shown on the page of our Web Site

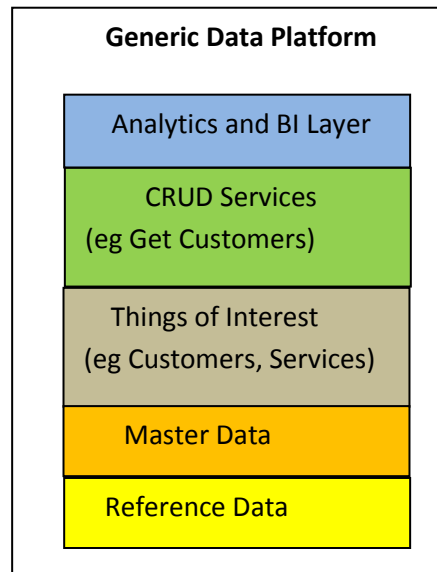- http://www.databaseanswers.org/data_models/parties_roles_and_customers/index.htm

and looks like this :-

3.4 Step 2 – A Universal Data Platform

Next we add separate Data Layers to create the Generic Data Platform

Generic Data Platform

Analytics and BI Layer

CRUD Services
(eg Get Customers)

Things of Interest
(eg Customers, Services)

Master Data

Reference Data

3.5 Steps 2 and 3 –Combined Platform and Model

Now we combine the Generic Data Platform and Canonical Data Model :-

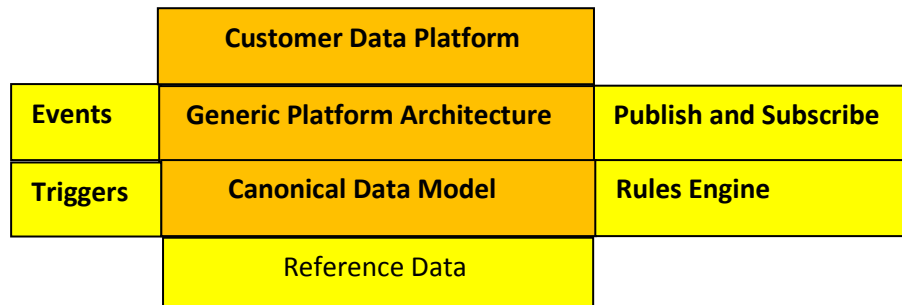**Generic Platform Architecture**

Canonical Data Model
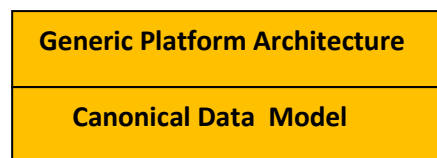
3.6 Step 4 Add Barry's Secret Sauce

My Platforms contain the contents of my 'Barry's Secret Sauce', which include :-

- Canonical Data Model
- Customer Data Platform
- Events
- Functional Layers
- Generic Data Platforms
- Industry-Specific Platforms
- Publish-and-Subscribe
- Reference Data Architecture
- Rules Engine
- Transaction
- Triggers

At this Step, we add five elements from the Sauce :-

- Customer Data Platform
- Events
- Publish and Subscribe
- Rules
- Triggers

	Customer Data Platform	
Events	**Generic Platform Architecture**	**Publish and Subscribe**
Triggers	**Canonical Data Model**	**Rules Engine**
	Reference Data	

4. Proof-of-Concept

4.1 Planning

My first planning activity is related to the Architecture for the Generic Data Platform.
I like the idea of incorporating the well-established Model-View-Controller Architecture.
 I can start by building on this Data Model :-
- http://www.databaseanswers.org/data_models/mvc_model_view_controller/index.htm

Then I will adopt a Microservices approach for Web Services for CRUD operations on data in my Data Platforms.

4.2 Role of SQL

So far as possible I will leverage the potential of SQL to implement a wide range of my Secret Sauce.
Then I will add vendor-specific features which add significant value.
For example, Oracle's APEX 'Data-as-a-Service'.
Details are included for each vendor in Chapter D.

4.3 Implement the POC

We start with three elements from Barry's Secret Sauce :-
- Customer Data Platform
- Events
- Publish and Subscribe

Events	**Customer Data Platform**	**Publish and Subscribe**

4.4 POC – Part 2

Next we add two more elements from Barry's Secret Sauce to define what we call a 'Generic Platform Architecture':-

- Rules
- Triggers

An essential part of our unique approach is to use the power of Data Models.

This allows us to design Data Models and use them to create Database Tables and create and load modest volumes of demonstration data or produce high vo9lumes of data in spreadsheets.

This Model was created using ERwin and uses ERwin's ability to show Many-to-Many Relationships.

It supports our Event-Oriented approach which is very powerful.

with this Approach we have adopted a generic framework which exploits Inheritance in the design of Customer and Service Models.

The combination of these two features provides an economical and powerful and design to support a wide range of Applications.

We have converted this to a Universal Data Model which is a very useful step towards the

In this White Paper we discuss three Applications :-
- Banking, Logistics and Retail

Triggers	Generic Platform Architecture	Rules Engine

5. Analytical Platform

Techopedia provides a definition of an Analytics Platform on this page :-

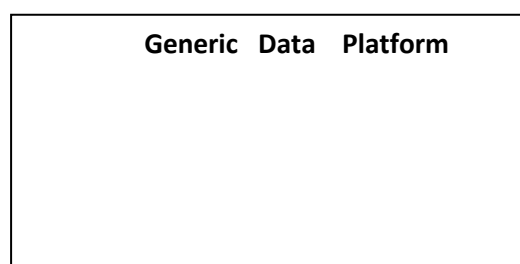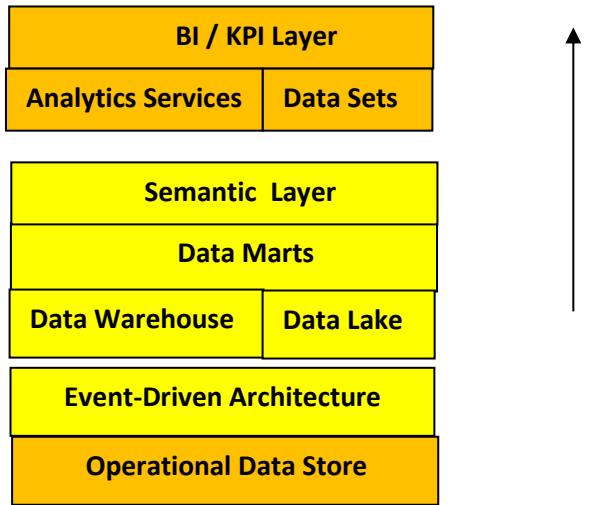- https://www.techopedia.com/definition/29493/analytics-platform

Here we show the Layers that can exist in an Analytics Platform.

We have included a Semantic Layer which provides a very powerful facility to translate technical, (data) term into business-oriented terms.

For example, 'Party' is a term commonly used by data professionals to refer to suppliers and customers.

A Semantic Layer would then be used to translate 'Party' to Customer or Supplier, as appropriate.

Generic   Data   Platform

10/05/2017 18:19

BI / KPI Layer	
Analytics Services	Data Sets

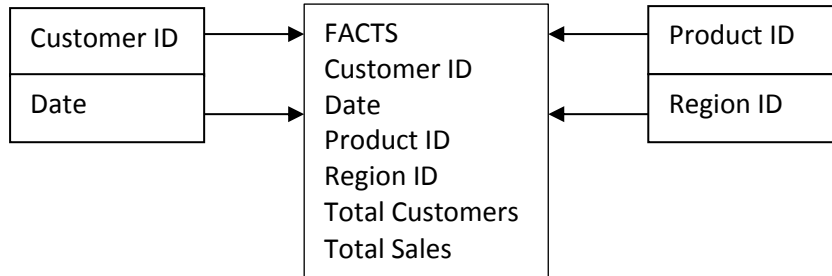| Semantic Layer	
Data Marts	
Data Warehouse	Data Lake
Event-Driven Architecture	
Operational Data Store	

5.1 Data Marts

This is a generic Design that we use as an example.
It features the most common Dimensions in a typical Dimensional Model, which are Customers, Dates, Products and Regions.

```
┌──────────────┐       ┌──────────────────┐       ┌──────────────┐
│ Customer ID  │─────▶ │ FACTS            │ ◀─────│ Product ID   │
├──────────────┤       │ Customer ID      │       ├──────────────┤
│ Date         │─────▶ │ Date             │ ◀─────│ Region ID    │
└──────────────┘       │ Product ID       │       └──────────────┘
                       │ Region ID        │
                       │ Total Customers  │
                       │ Total Sales      │
                       └──────────────────┘
```

6. Customer Data Platforms

6.1 Features

We can identify some features that are specific to Customer Data Platforms, including :-

- Provide an integrated Database of Customer-related data.
- Provide external access to this Database, perhaps by SQL.
- Tailored interaction to Social Media
- Customer-specific access to Data and Services
- Customer-specific Publish and Subscribe facilities
- Planning of Customer Journey Mapping

6.2 Data Architecture

The Architecture for a Customer Data Platform that support KPIs (Key Performance Indicators) is on this page :-

* http://www.databaseanswers.org/data_models/customer_data_platform_for_kpis/index.htm

and looks like this :-
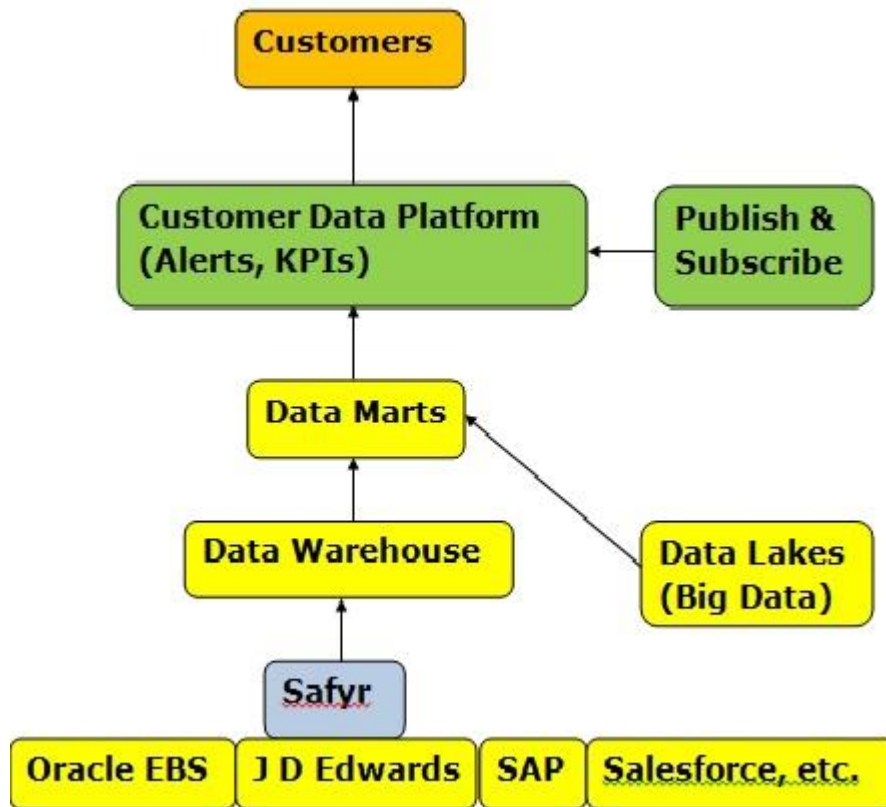
We show our view of Safyr Data Platform in 7.6

6.3 Single View of a Customer

Our Approach is to Compare, contrast and validate the data from separate Customer data sources.

Here we discuss three stream of Customer data where Streams 1 and 2 both have Customers called Jo Smiths.
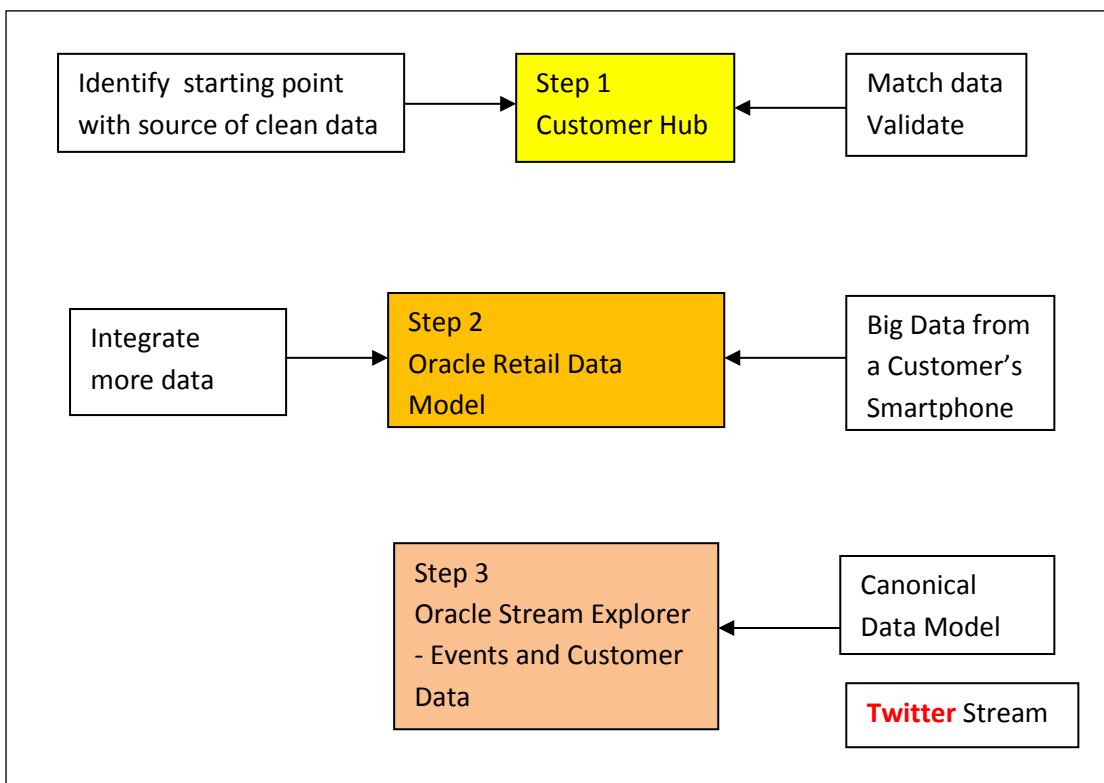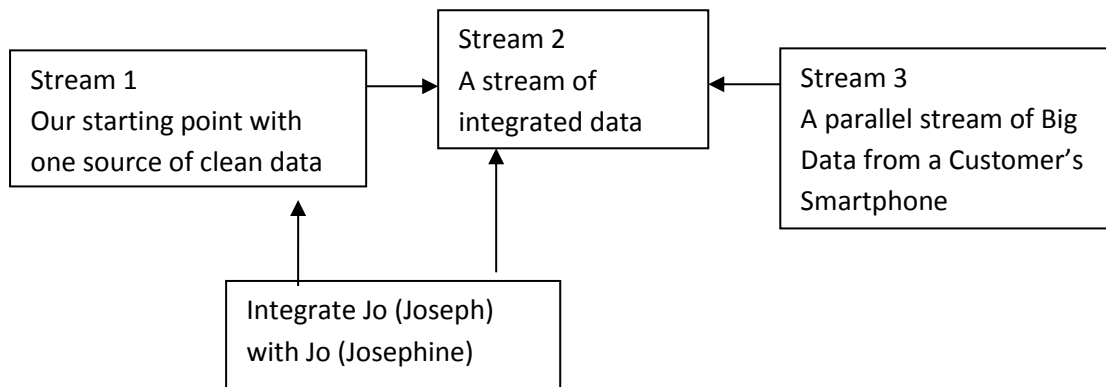
Stream 3 is a Big data Stream coming from a smartphone.

We find they share the same address and the same last name, but have a different Date of Birth and different Genders, like this :-

Jo Smith – Joseph

Jo Smith - Josephine

We conclude they are either Husband and Wife or Brother and Sister and we flag them as separate individuals.

```
┌──────────────────────┐      ┌──────────────────────┐      ┌──────────────────────┐
│ Stream 1             │ ───▶ │ Stream 2             │ ◀─── │ Stream 3             │
│ Our starting point   │      │ A stream of          │      │ A parallel stream of │
│ with one source of   │      │ integrated data      │      │ Big Data from a      │
│ clean data           │      │                      │      │ Customer's Smartphone│
└──────────────────────┘      └──────────────────────┘      └──────────────────────┘
          ▲                             ▲
          │                             │
       ┌──────────────────────────────────┐
       │ Integrate Jo (Joseph)            │
       │ with Jo (Josephine)              │
       └──────────────────────────────────┘
```

```
┌──────────────────────────────────────────────────────────────────────────────────┐
│  ┌────────────────────┐     ┌────────────────┐     ┌────────────────┐             │
│  │ Identify starting  │ ──▶ │ Step 1         │ ◀── │ Match data     │             │
│  │ point with source  │     │ Customer Hub   │     │ Validate       │             │
│  │ of clean data      │     │                │     │                │             │
│  └────────────────────┘     └────────────────┘     └────────────────┘             │
│                                                                                    │
│  ┌────────────────┐         ┌────────────────┐     ┌────────────────┐             │
│  │ Integrate      │ ──▶     │ Step 2         │ ◀── │ Big Data from  │             │
│  │ more data      │         │ Oracle Retail  │     │ a Customer's   │             │
│  │                │         │ Data Model     │     │ Smartphone     │             │
│  └────────────────┘         └────────────────┘     └────────────────┘             │
│                                                                                    │
│                             ┌────────────────┐     ┌────────────────┐             │
│                             │ Step 3         │ ◀── │ Canonical      │             │
│                             │ Oracle Stream  │     │ Data Model     │             │
│                             │ Explorer       │     └────────────────┘             │
│                             │ - Events and   │     ┌────────────────┐             │
│                             │ Customer Data  │     │ Twitter Stream │             │
│                             └────────────────┘     └────────────────┘             │
└──────────────────────────────────────────────────────────────────────────────────┘
```

6.4 Oracle Offerings

Oracle has a number of Customer-related Offerings, including these :-

- Customer Hub (Introduction)
  - https://docs.oracle.com/cd/E24010_01/doc.111/e23082/chap1_understanding_pip.htm
- Customer Hub (Terminology)
  - https://docs.oracle.com/cd/E24010_01/doc.111/e23082/chap1_understanding_pip.htm#BA BGFBGG
- Oracle JDEdwards - Setting up Customer Master Information
  - https://docs.oracle.com/cd/E16582_01/doc.91/e15085/setupcustomermasterinfo.htm#EOAAR00884
- Oracle Retail Data Model (Tutorial) :-
  - http://download.oracle.com/oll/tutorials/ORDMTutorial/player.htm
- Oracle Retail Data Model (Glossary of Terms) :-
  - http://download.oracle.com/oll/tutorials/ORDMTutorial/html/start_here/glossary.htm
- Oracle Streams
  - http://docs.oracle.com/cd/B28359_01/server.111/b28321/strms_over.htm#i1006084

6.5 Oracle Product Analysis

This analysis helps us to choose the specific Oracle product to meet our requirements.

'Cloud' shows which features are available in the Clouds as well as in-premise.

Feature	Attribute Matching	Customer Master Index	Big Source	Cloud
Oracle Customer Hub				
Oracle Retail Data Model				
Oracle Streams				

6.6 Data Streams

We analysis the Customer Data in three Streams

1. Basic
   a. Customer Hub (OCH)
   - https://docs.oracle.com/cd/E24010_01/doc.111/e23082/chap1_understanding_pip.htm

   "A core feature in OCH is the ability for a data steward to merge customer records."
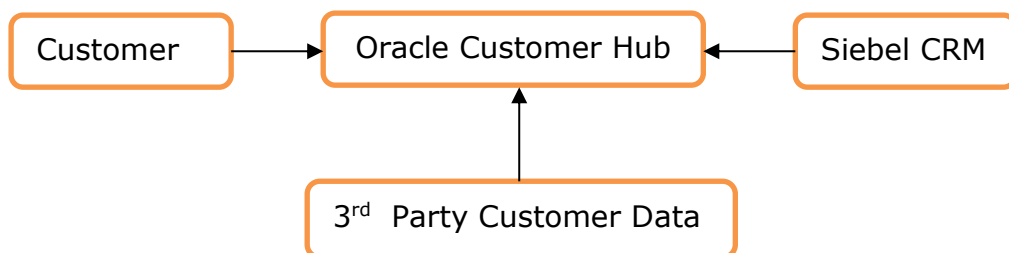
2. Intermediate
    a. Oracle Retail Model Tutorial
    o http://download.oracle.com/oll/tutorials/ORDMTutorial/player.htm

3. Advanced
    a. Oracle   Streams
    o http://docs.oracle.com/cd/B28359_01/server.111/b28321/strms_over.htm#i1006084

This analysis helps us to choose the specific Oracle product to meet our requirements.

7. Development Platforms

7.1 Low-Code Platforms

Wikipedia has an entry for Low-Code Development Platforms on this page :-
- https://en.wikipedia.org/wiki/Low-code_development_platforms

where it states :-
On June 9, 2014 Forrester Research, defined a Low-Code Development Platform as a type of technology that allows for creating apps through configuration of functions, rather than coding those functions. Low-code development platforms enable accelerated delivery of business applications by reducing the amount of hand-coding required as well as the initial investment in setup, training, and deployment.

Low-code development platforms are based on the principles of model-driven design, automatic code generation, and visual programming.

Here's a review of the Best Low-Code Development Platforms, (which includes Zoho Creator)
- https://www.g2crowd.com/categories/low-code-development-platforms

with some useful definitions :-
"Low-code development platforms provide development environments that allow businesses to develop software quickly with minimal coding, minimizing the need for extensive coding experience. The platforms provide base-level code, scripts, and integrations so companies can prototype, build, or scale applications without developing complex infrastructures.

To qualify for inclusion in the Low-Code Development Platforms category, a product must:
- Generate source code as a base for customization
- Allow developers to customize html markup and source code
- Integrate with databases, web services, or APIs to connect data"

7.2 No-Code Platforms

Here is a review of No-Code Platforms (which includes Salesforce):-

- https://www.g2crowd.com/categories/no-code-development-platforms

The Review states :-

"No-Code Development Platform provide drag-and-drop tools that that allow businesses to develop software quickly without coding. The tools provide templates for workflow, element libraries, and interface customization tools to create fully functioning applications without any coding.

To qualify as a No-Code Development Platform, a product must:

- Drag-and-drop elements to assemble applications
- Provide no tools for customizing code within the platform
- Integrate with databases, web services, or APIs to connect data

7.3 Microsoft PowerApps

Here is a review of Microsoft's new PowerApps :-

- https://adtmag.com/articles/2015/12/02/microsoft-powerapps.aspx

Microsoft has introduced PowerApps, the company's new low-code cross-platform mobile app development solution designed to let non-developers create enterprise apps.

It uses prebuilt templates, wizards, and a visual drag-and-drop designer to enable business users to leverage their Office skills to create simple apps.

PowerApps comes with the Azure App Service, letting users code in .NET, NodeJS, PHP, Python, or Java, leverage continuous integration services such as Visual Studio Team Services or GitHub. PowerApps can also integrate with Azure Active Directory and Azure Media Services.

8. My Platforms

This is a selection of 10 from

8.1 Airport Management

This Platform is on this Page :–

- http://www.databaseanswers.org/data_models/airport_mgt_platform/index.htm

and shows the data that lies between Analytics and BI Layer (at the top) and Reference Data (at the bottom) :-
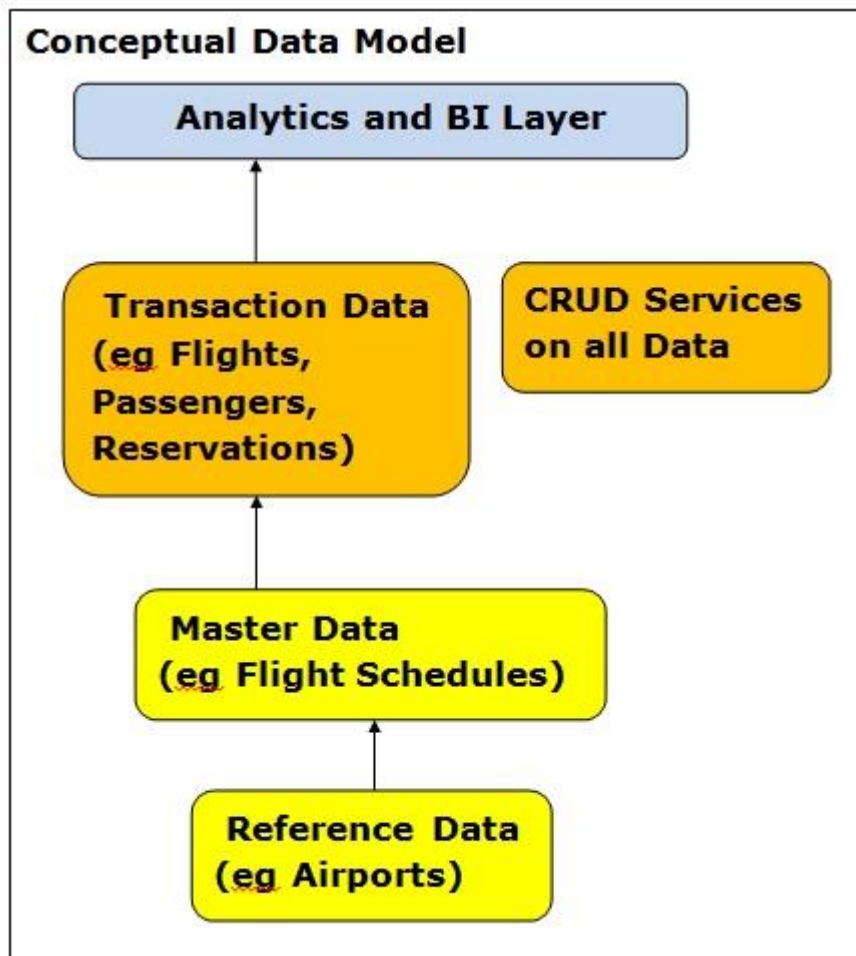
8.2 Banking

This Platform is on this Page :–
- http://www.databaseanswers.org/data_models/banking_platform_2020/index.htm

and shows three classical layers of Transaction, Master and Reference data :-

Banking Platform

**Transaction Data
eg Payments, Withdrawals**

**Master Data
eg Customers**

**Reference Data
eg Payment Methods**
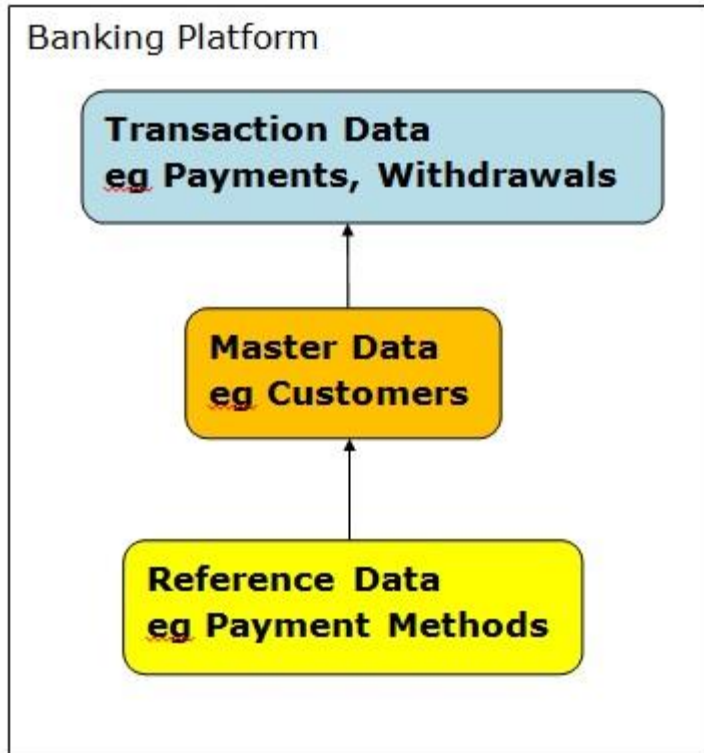
8.3 Customers

8.3.1 Customer's Activities

This Platform is on this Page :–

- http://www.databaseanswers.org/data_models/customer_data_platform_for_kpis/index.htm

Here we show a Customer Platform to build a profile of the Customer's Activities :-

- http://www.databaseanswers.org/data_models/customer_data_platform/index.htm
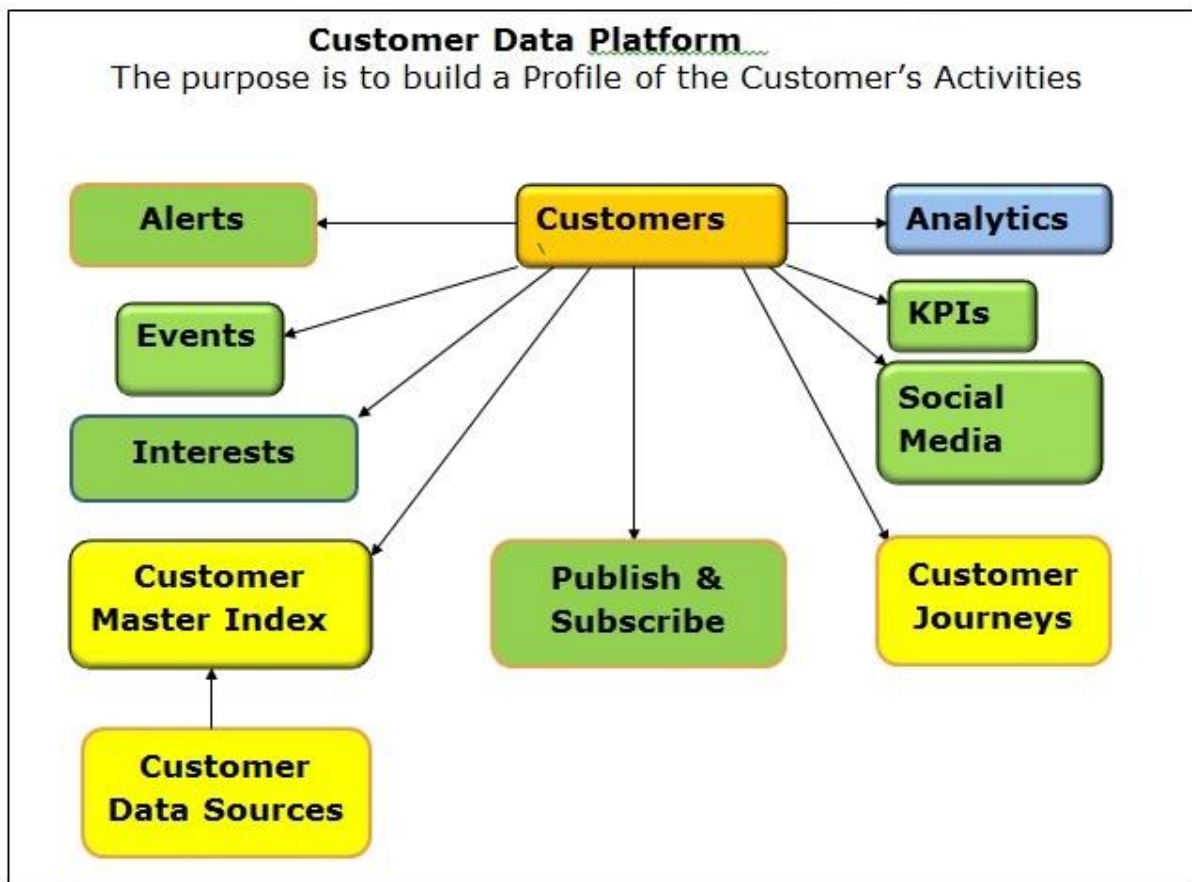
8.3.2 Customer KPIs

This Platform is on this Page :–

* http://www.databaseanswers.org/data_models/customer_data_platform_for_kpis/index.htm

and shows how the Customer Data Platform sits between the Customer and the data that the Customer is interested in :-
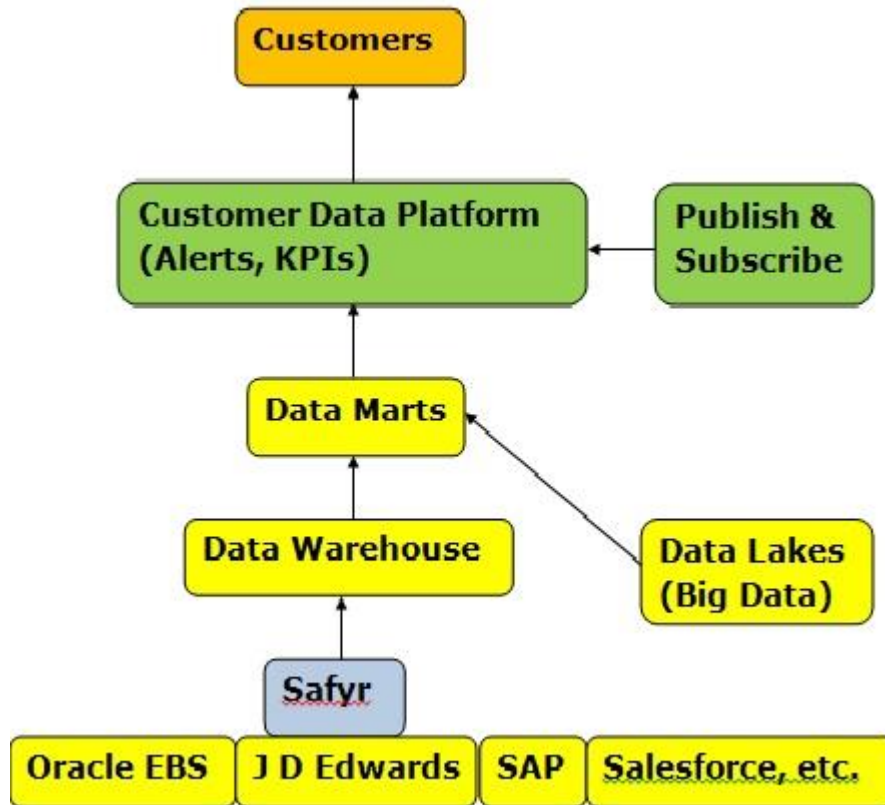
8.4 Event-Driven Entities

This Platform is on this Page :–

- http://www.databaseanswers.org/data_models/event_driven_platform/index.htm

and shows the layers above and below an Enterprise Service Bus :-
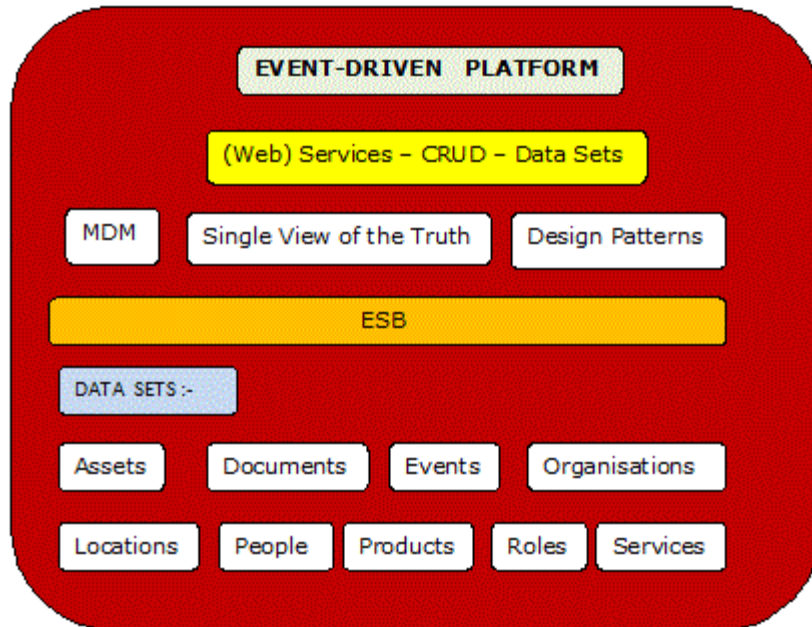
8.5 Event-Driven Layers

This Platform is on the same Page :–

- http://www.databaseanswers.org/data_models/event_driven_platform/index.htm

and it shows the four layers above and below the Event-Driven Platform :-
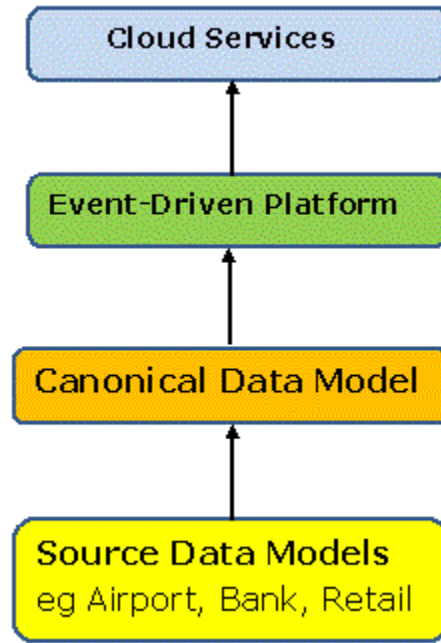
8.6 Generic

This Platform is on this Page :–

- http://www.databaseanswers.org/data_models/generic_platform/index.htm

and shows four basic Layers of Transaction, and Reference Data
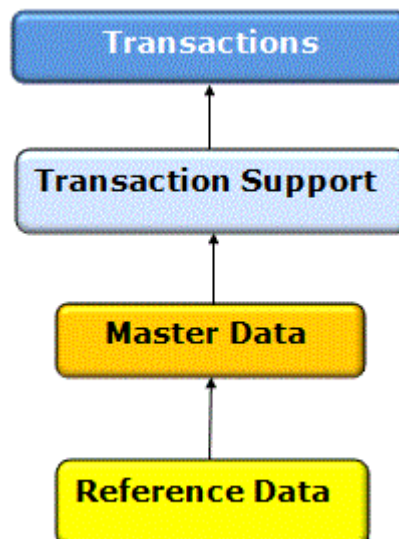
8.7 Hedge Funds

This Platform is on this Page :–

- http://www.databaseanswers.org/data_models/hedge_fund_platform/index.htm

It shows the layers between the Customer and the data :-
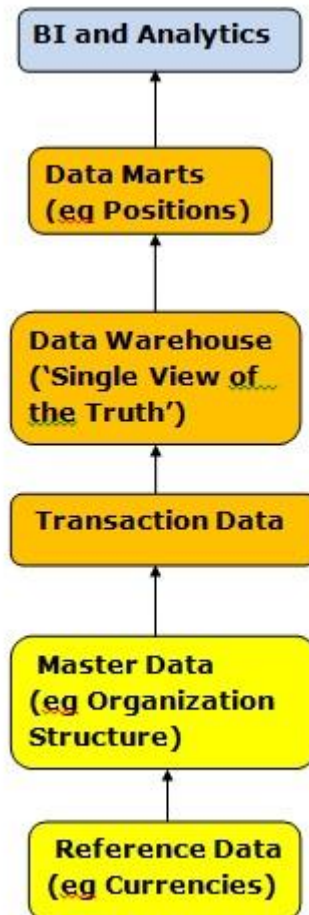
8.8 Master Data Management

This Platform is on this Page :-

- http://www.databaseanswers.org/data_models/master_data_mgt/index.htm

and looks like this and shows how an ERD can show a Platform :-
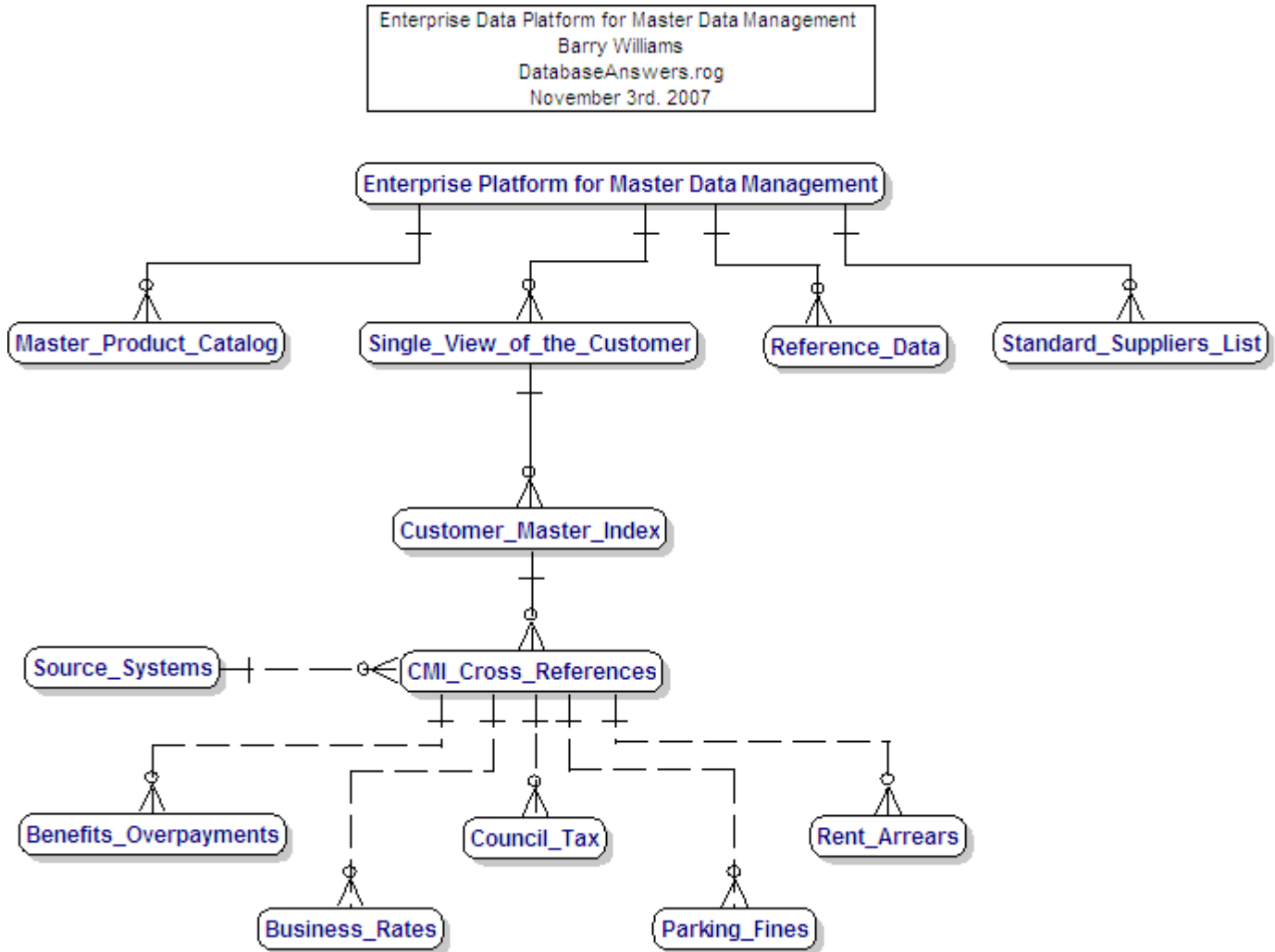
8.9 Payments Platform

This Platform is on this Page :-

- http://www.databaseanswers.org/data_models/payments_platform/index.htm

and shows how the Payments Platform can be on top of a number of Layers :–
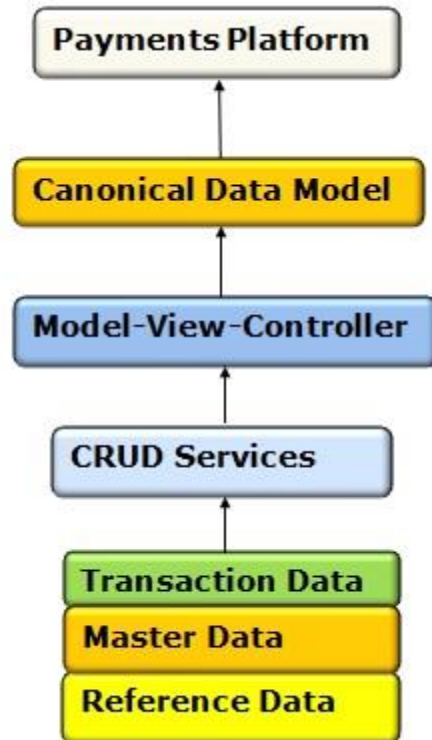
8.10 Service Delivery Platforms

This Platform is on this Page and shows how Platforms can be layered :–

- http://www.databaseanswers.org/data_models/layered_service_delivery_platforms/index.htm
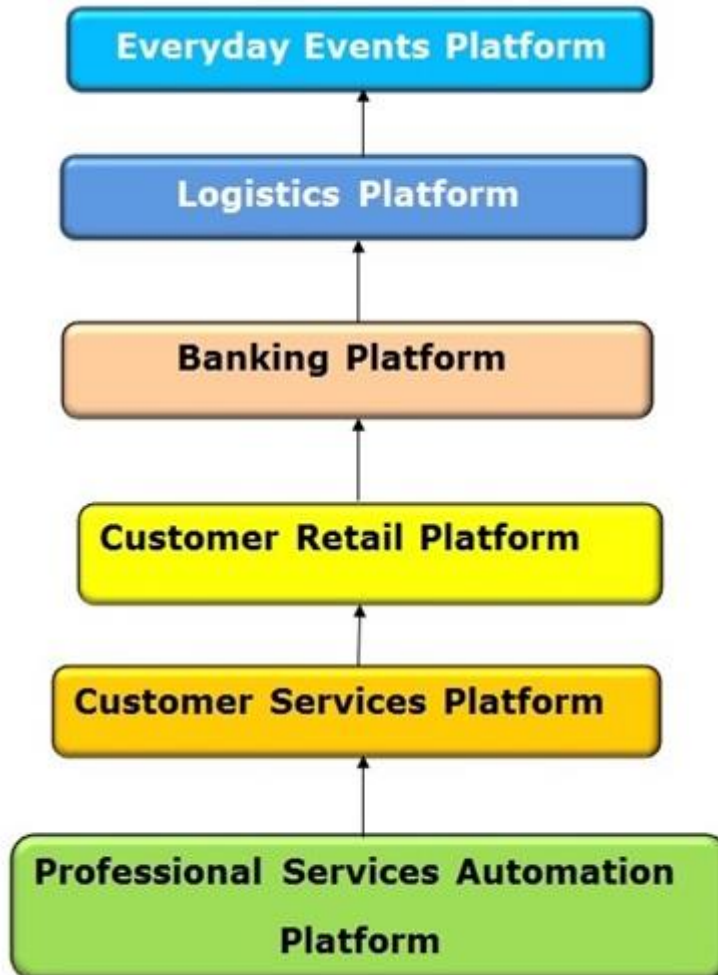
8.11 United Nations

This Platform for the UN Global Compact is on this Page :–

- http://www.databaseanswers.org/data_models/un_global_compact_platforms_for_2017/index.htm

This follows our 'Best Practice' Approach with the user-oriented Analytics Entities at the top layer and the basic Financial Instruments Entities at the bottom.
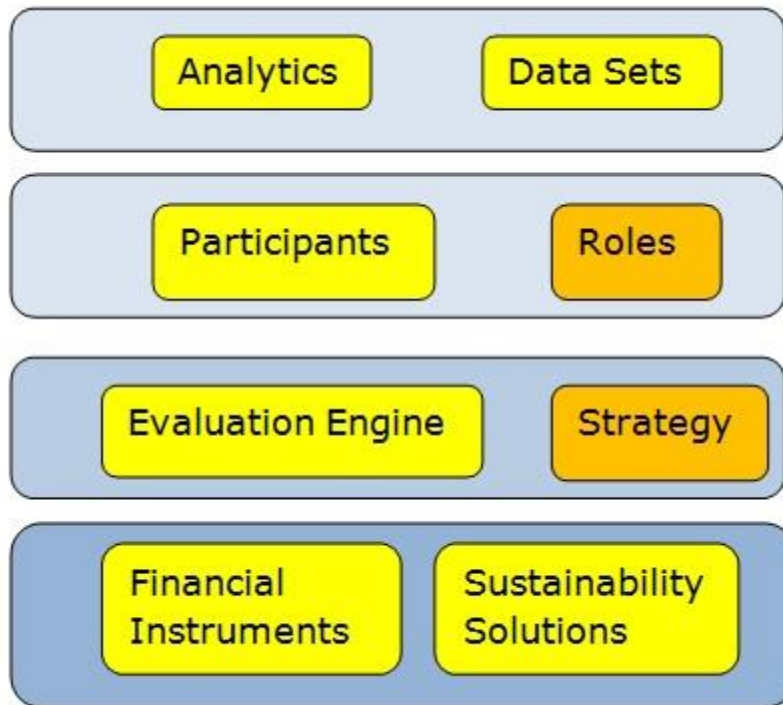
9. Commercial Platforms

9.1 Amazon and Google

Amazon and Google's Platform compared by zdnet.com –

- http://www.zdnet.com/article/comparing-amazons-and-googles-platform-as-a-service-paas-offerings/

Useful Links :-

- Facebook Platform - https://en.wikipedia.org/wiki/Facebook_Platform
- Google Cloud Platform – https://cloud.google.com/
- Google Developers Platform - https://developers.google.com/+/web/

9.2 IBM

IBM offers 'The IBM Big Data Platform' that we show below and appears in this PDF file :-
"Architecting a Big Data Platform for Analytics"

- http://citia.co.uk/content/files/architecting-a-big-data-platform-for-analytics_24606569.pdf
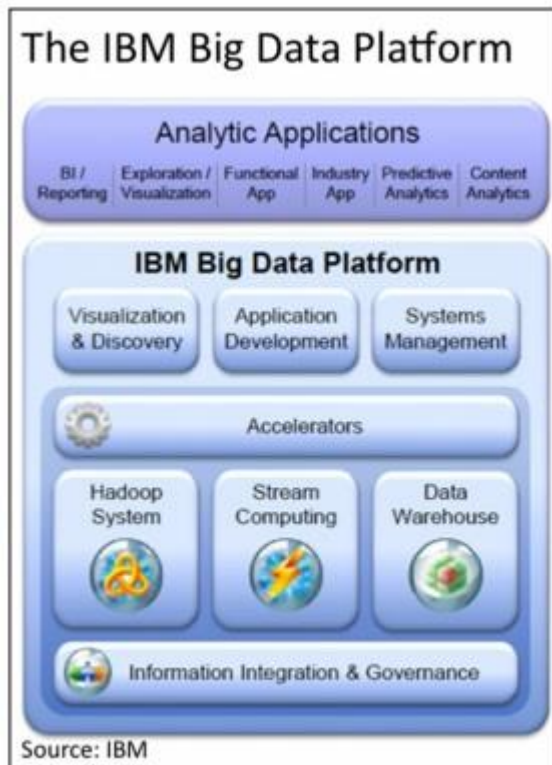
This Table shows how our Layers are supported by this IBM offering.

| Layer	IBM Support
Data Presentation	BI Reporting ?
Data Analytics	Analytic Applications
Data Integration	Information Integration and Governance

9.2.1 Useful Links for IBM

Alan Chatt – alan.chatt@uk.ibm.com

Big Data and Data Warehouse Augmentation :-

- http://www.ibm.com/developerworks/library/ba-augment-data-warehouse2/index.html

SOA and ESB http://www.ibm.com/developerworks/library/ws-soa-eda-esb/index.html

9.3 Microsoft

Here is a statement by Microsoft's view on Data Platforms :-

- *https://blogs.technet.microsoft.com/dataplatforminsider/2014/01/29/what-drives-microsofts-data-platform-vision/*

9.4 Oracle

Oracle offers this 'One bank Customer-Oriented Platform' which is taken from an Oracle Presentation on 'Modern Core Banking' available on this page :-

- http://www.slideshare.net/lanceslides/idg-banking-vietnam2010-customercentricplatform100527ltayv07

This table shows how our Layers are supported by this Oracle offering.

Layer	Oracle Support
Data Presentation	Multichannel
Data Analytics	BI and Analytics
Data Integration	Process and Data Integration

9.5 Progress Software

Progress offer their Digital Factory Platform, which includes an Event Processing Platform and a Rules Engine ('Corticon').

This Table shows how our Layers are supported by this Progress offering.

Layer	BOB Support
Data Presentation	BI Reporting ?
Data Analytics	Analytic Applications
Data Integration	Information Integration and Governance

FIND URL and Image
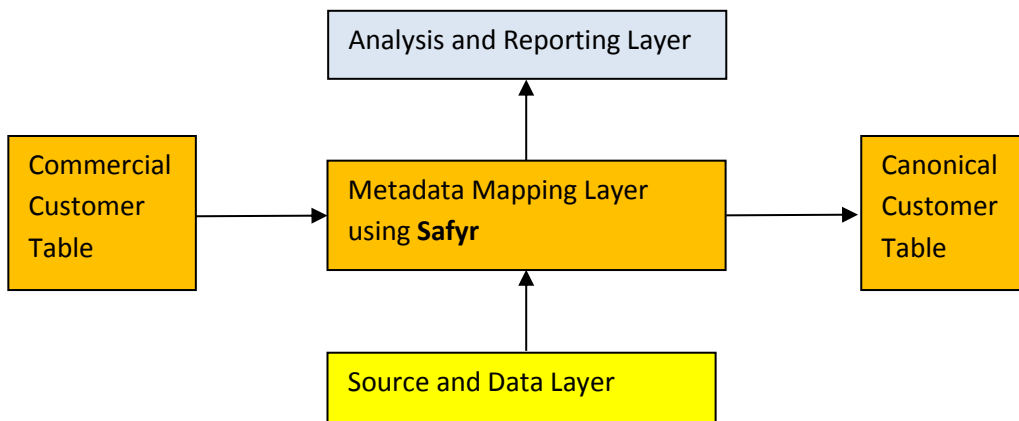
9.6 Safyr

This is our view of the Safyr Data Architecture, taken from this page :-

- http://www.databaseanswers.org/data_models/customer_data_platform_for_kpis/index.htm
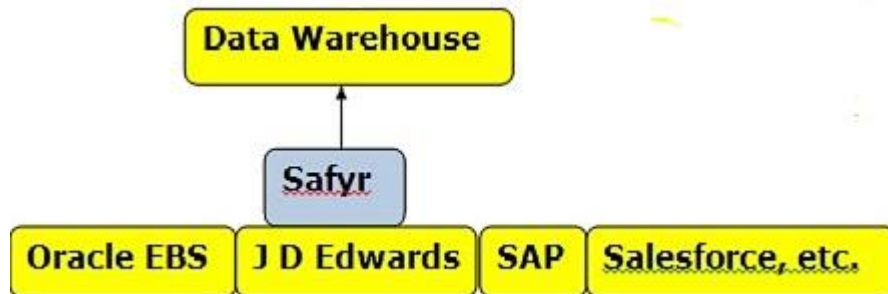
Here we show our view of how this maps on to our Safyr Data Platform :-

9.7 Salesforce

Salesforce offers this 'Cloudforce Platform' for the Social Enterprise Architecture.
Our analysis shows that it offers no support for our Generic Data Platform.

This shows that the word 'Platform' can be used in many different ways, even in the restricted scope of 'Enterprise Data Management'.

Layer	Cloudforce Platform Support
Data Presentation	None
Data Analytics	None
Data Integration	None
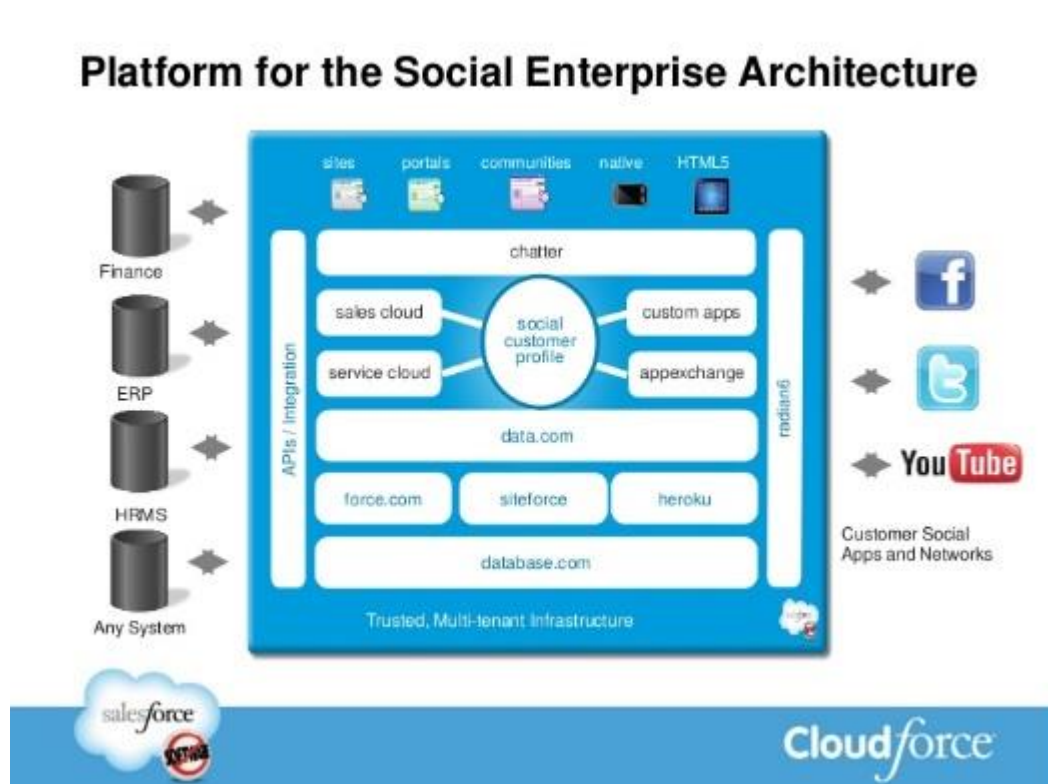
9.8 Task Rabbit

TaskRabbit brings together Freelancers for hire and people looking to hire freelancers.
- https://www.taskrabbit.co.uk/

9.9 Teradata

Teradata offers a 'Unified Data Architecture' that I have published on this page of my Database Answers Web Site :-

- http://www.databaseanswers.org/data_models/big_data/Teradata_Unified_Data_Architecture.htm

and we show it below.

It includes the Aster Discover (Analytics) Platform.

Layer	Teradata Support
Data Presentation	'Business Intelligence'
Data Analytics	'Aster Discovery Platform'
Data Integration	'Integrated Data Warehouse'
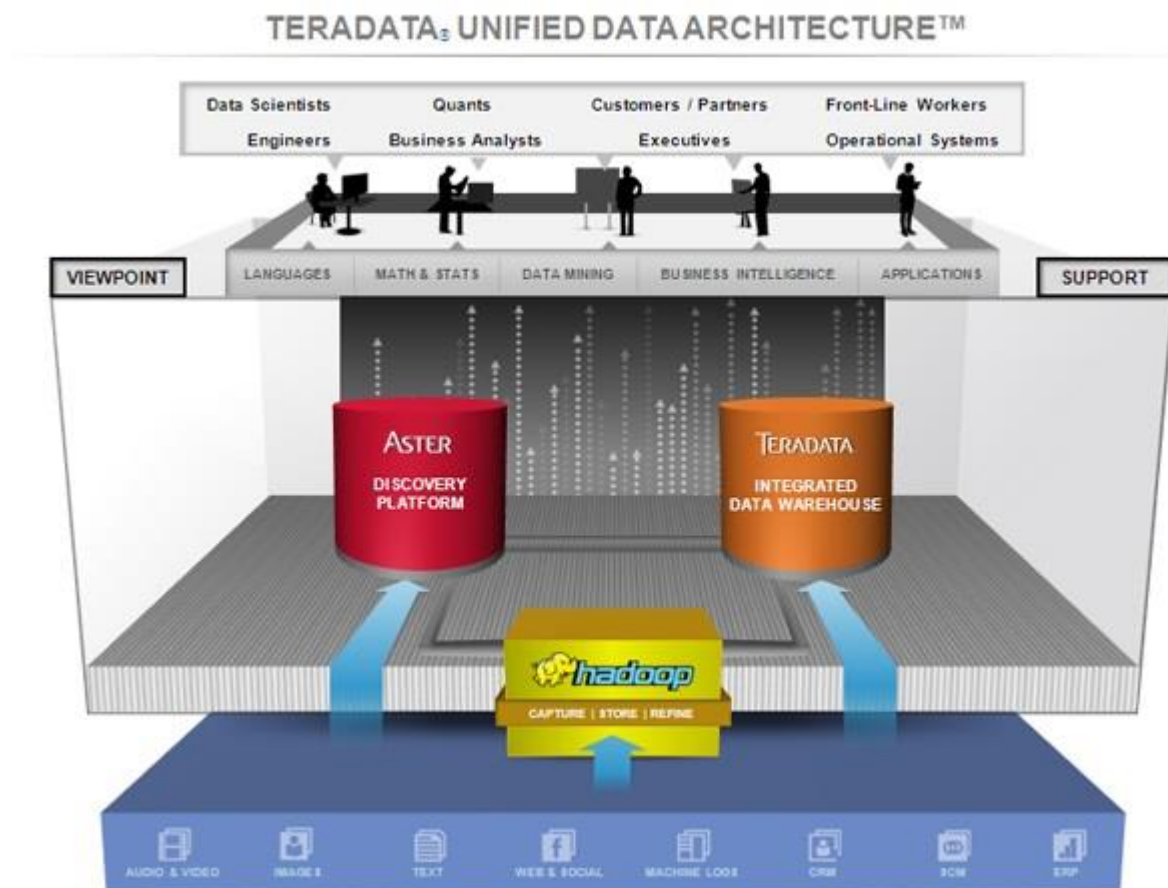
9.10 United Nations (UNGC) Global Explorer Platform

This is a Publication Platform from the United Nations Global Compact Global Explorer Platform which is published on this page :-

- http://www.databaseanswers.org/data_models/ungc_global_opportunity_explorer/index.htm

1

9.11 Upwork

Upwork is a Platform for Freelancers and People looking to hire freelancers to have work done:-

- https://www.upwork.com/

9.12 Verbling

Verbling is a Platform for Students and freelance Teachers.
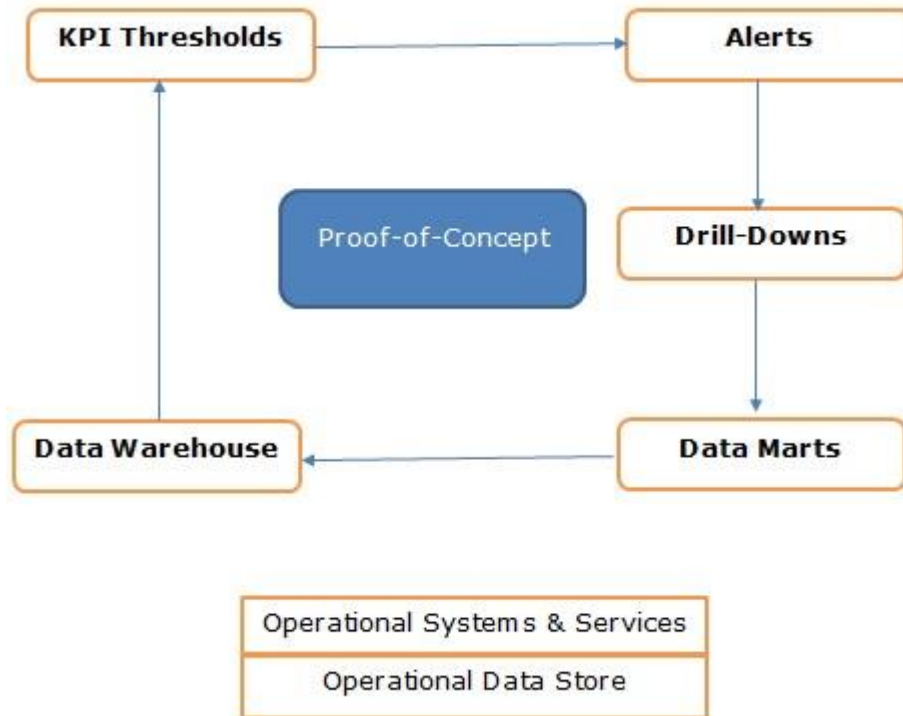
- https://www.verbling.com/

It offers high-level facilities for Teachers to advertise their Services and for Students to register with Teachers of their choice.

Appendix. KPIs and Alerts

This shows our approach to KPI Thresholds and Alerts.

- http://www.databaseanswers.org/data_models/platform_project/index.htm

Step 1 : The CEO can play golf and be alerted when an emergency occurs at the Airport.

- http://www.databaseanswers.org/data_models/airport_in_a_box/airport_in_a_box_POC_end2end.htm