Project End to End Design – *high-level version*

Project: - ***Insert name of project***

Author: - ***End to End Designer for project***

Date:-

Version:-

VERSION INFORMATION

LAST UPDATED	MASTER VERSION LOCATION

CHANGE HISTORY

VERSION NO.	DATE	CHANGE DESCRIPTION	APPROVED BY

REVIEWERS

VERSION NO.	DATE	NAME	TITLE / ROLE
			Delivery Manager

BRIEFED

VERSION NO.	DATE	NAME	TITLE / ROLE
			TDA lead
			Tower Lead Back End
			Tower Lead Semantic Layer
			Tower Lead Front End
			Project Director
			SME / Business Contact
			ES IT
			AM

This document contains many tables and diagrams. This reflects the remark that the E2E design is mostly used as a reference. In that case, tables are easier to use.

Since tables and diagrams are often used in this document, it is important to use a common colour scheme in the tables and diagrams.

TABLE OF CONTENTS

1 INTRODUCTION 5

1.1 Document purpose 5

1.2 Scope 5

1.3 Out of scope 5

1.4 Related documents 6

1.5 Key design decisions 6

2 PROJECT REQUIREMENTS 7

2.1 Functional requirements 7

2.2 Non functional requirements 7

3 SOLUTION OVERVIEW 8

3.1 Executive Summary 8
3.1.1 Outline 8
3.1.2 Security 8

3.2 Design considerations 8
3.2.1 Availability 8
3.2.2 Response time 9
3.2.3 Data latency 9
3.2.4 Data reconciliation 9

3.3 Data Sources (DS) 10
3.3.1 Sources 10
3.3.2 Data files & volumetrics 10
3.3.3 Tooling 11

3.4 Source data layer (SA) 11
3.4.1 Tables 11
3.4.2 Transformation 12
3.4.3 Tooling 12

3.5 Enterprise data layer (EDL) 12

3.5.1 Tables 13
3.5.2 Transformation 14
3.5.3 Tooling 14
3.5.4 Data cleansing & data quality 14

3.6 Business Semantic Layer (BSL) 15
3.6.1 Views 15
3.6.2 Transformation 16
3.6.3 Tooling 16

3.7 Reporting & analytics 16
3.7.1 Usage of the data from Semantic Layer / EDL for reporting 16
3.7.2 Reporting tools 17

3.8 Allocation 17

3.9 User profiles and security 17
3.9.1 Users 17
3.9.2 Data Security for data at rest 17
3.9.3 Data Security for data in motion 17

3.10 Future BAU requirements 18

4 COMPLIANCE WITH PROGRAM STANDARDS 19

1 INTRODUCTION

1.1 Document purpose

This document is intended to provide a high level design of the solution for the *xxx* project. *This document is to be created shortly after the finalisation of the user requirements. It is then intended as a guiding document throughout the project. It is only supposed to be updated if a major change is applied within the project.*

This document only details project specific design.

The scope of this document is limited to the project only. The BAU phase is not part of this document. Nevertheless, some information is required that is relevant for this BAU phase. This information can easily be collected in this phase of the project and is very relevant for the BAU phase. Hence a section "Future BAU requirements" is included in this document to provide a placeholder for such BAU relevant information.

It is intended that this document will be presented to the Towers in a briefing session. This document will be written only once, shortly after the approval of the user requirements. After the briefing session, the Towers should have a clear picture on the scope and content of the project.

1.2 Scope

Provide a high-level overview of what is in scope for the project. One may address here:

- *Which reports are in scope*
- *For which geographical area is this project set up*
- *For which information subject areas is this project set up*
- *Which source systems are in scope?*
- *Etc.*

1.3 Out of scope

What is specifically out of scope for the project?

1.4 Related documents

Identify any related documents that should be referenced alongside this document in order to provide context or background to the contents within this document. Provide referencing documents and their version numbers to fully understand the document is based on.

As a minimum, provide reference information on (links only):

- *The Business Capability Architecture*
- *The Criticality Assessment*
- *The Functional Requirements*
- *The Non Functional Requirements*

1.5 Key design decisions

List any key design decisions that have been made as part of this project, specifically where they may not form part of the strategic roadmap for information delivery and the background / justification around these decision.

*Note: this should be a **short** overview. Hence, limit yourself to a high level overview only. More detail can be given in the detailed version of this document. If you think that it is necessary to provide an architectural discussion, one may use the chapter on "Compliance With Program Standards".*

Decision taken	*Root cause for the problem that necessitates the decision*	*Justification for the decision taken*	*Likely consequences from the decision*	*Is the decision compliant with the IG-TDA-OneEDW. (If not compliant,) has an exception being requested?*

2 PROJECT REQUIREMENTS

2.1 Functional requirements

The functional requirements are captured in *reference to Functional Requirements*. *Only a link needs to be provided here.*

This section may include additional comments on the Functional Requirements.

2.2 Non functional requirements

The functional requirements are captured in reference to *Non-Functional Requirements*. *Only a link needs to be provided here.*

This section may include additional comments on the Non-Functional Requirements.

3 SOLUTION OVERVIEW

The solution is provided at high level – "*How* can this be achieved?" It also addresses the platforms and software that will be required to deliver the solution.

3.1 Executive Summary

3.1.1 Outline

The purpose of this section is to clearly define what the solution will provide. Try to avoid technical language here. It is aimed at decision makers who need to have an idea on what the project is about. Therefore, try to limit this section to 3 pages.

One may use the Business Capability Architecture (BCA). One may use this BCA as starting point for a non-technical description of the project. Indicate any deviations from the strategic reference architecture.

3.1.2 Security

Provide information on security as well. This section provides a solution on the non-functional requirements and the outcome from the Citicus assessment. Provide the consequences for the project in laymen terms.

Security is seen as a subject that is of uttermost importance to the user community as senior officers may be held personal accountable for breaches in security. Hence, such potential threats must be explained in understandable terms to the user community.

3.2 Design considerations

3.2.1 Availability

This section is intended to provide an overview on the availability of the solution that is being deployed and highlight service level requirements (SLR). This section provides a solution on the non-functional requirements on availability on reports and data.

Example. In the non functional requirements, it might be stated that reports must be available in the period 08.00 – 18.00 CET. The project might decide to arrange the schedule process in such a way that this requirement can be met. Whenever decisions are taken, such can then be mentioned in the key decisions list. The high level design can be included here.

3.2.2 Response time

This section is intended to provide an overview on the response times of the solution that is being deployed and highlight service level agreements (SLR). This section provides a solution on the non-functional requirements on response times on reports and queries.

Example. In the non-functional requirements, we might have a requirement to see reports within a 20 second time frame. The project might decide to materialise views in the semantic layer. Such a decision is included in the key decisions list. In this list, it is also indicated that an exception will be obtained from the standard architecture. Here, one may provide the high level design on how requested response time is met.

3.2.3 Data latency

What data latency is acceptable for the solution and what measures will be put in place to ensure the required latency can be met?

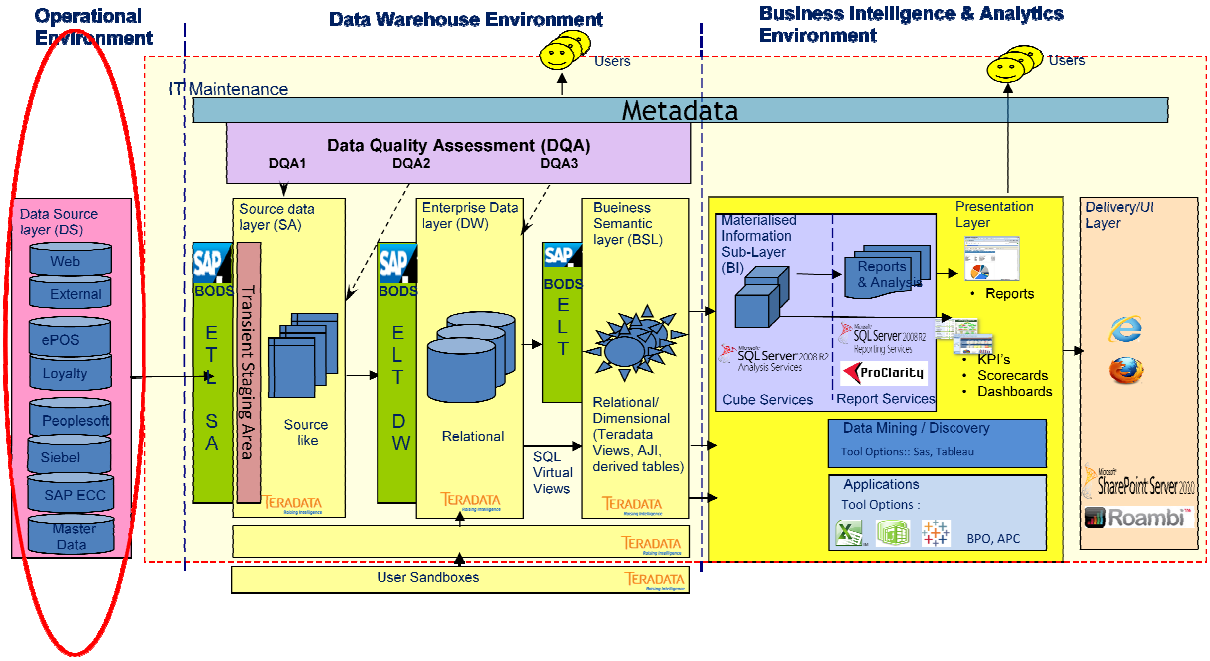
Example. In the non-functional requirement, one may read that events should be visible in the reports with a maximum delay of 24 hours after their occurrence. The project might decide to create a scheduling process that is able to meet this requirement. Here, the high level design on scheduling might be provided.

3.2.4 Data reconciliation

What measures will be taken to ensure data reconciliation during the series of transformations between source data and final reports?

Example. In the series of transformations, data rows might get lost (as result from a fault in a filter) or data might get corrupted (due to formatting issues). Mechanisms can be created to detect such inconsistencies, such as comparisons of the number of rows, or comparison on hash totals.

3.3 Data Sources (DS)



3.3.1 Sources

Provide an overview of the sources that are accessed. These sources bring in data that are not yet captured in Enterprise Data Layer. Existing data structures that will be re-used in the project can be left out.

Only sources can be included that are seen as a recognised authority of the data involved. Note: this implies that the BAA is in general not the adequate data source. The approach is that data should always be sourced as far as possible from the true source of the data – not another BAA, EDL or data mart – because one doesn't know what transformations, filtering, aggregations etc may have been done and one end up potentially missing details that you would otherwise get.

Indicate for what data requirement the sources are accessed. Provide an indication as to whether the creation of such source data requires manual intervention.

Below, an example is given on how to display the sources that will be accessed in the project.

Data Requirement	Expected Source System	Source	Manual intervention
For example: invoice information	*For example: Sirius Sirius ECC*	*For example: 2LIS_06_INV*	*For example: no*

3.3.2 Data files & volumetrics

Provide high level information on data files that are to be provided, along with preliminary estimated volumes and frequency of availability. Indicate which files are stored in an encrypted form.

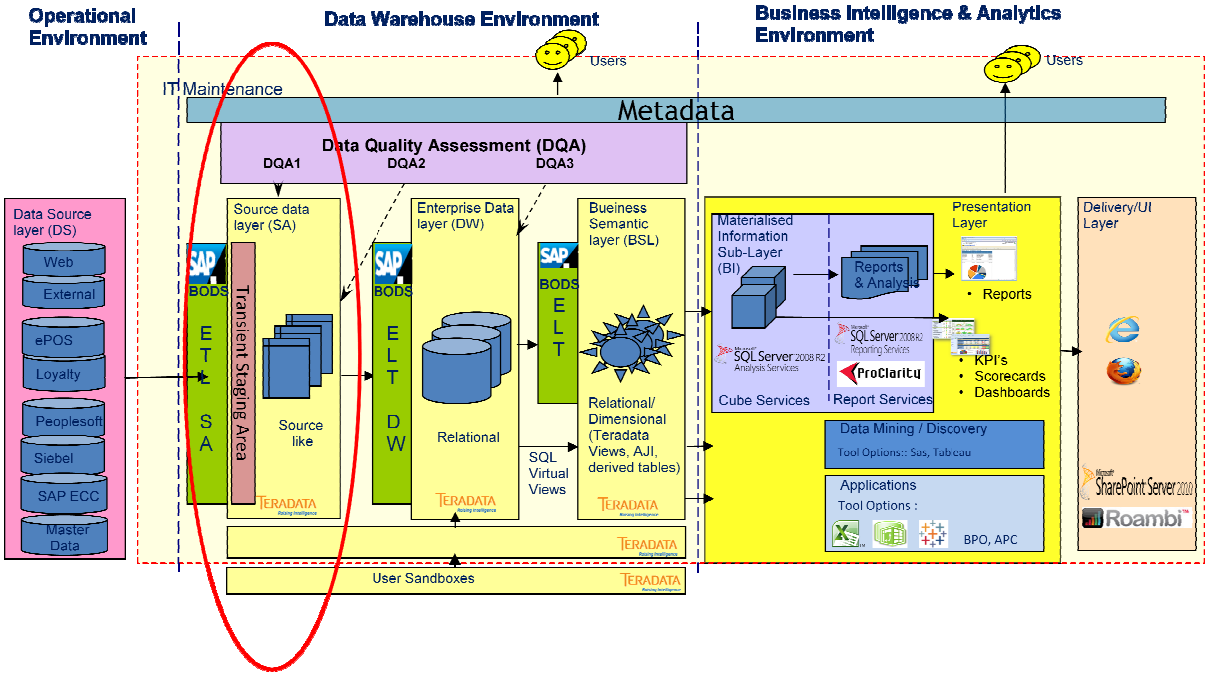
As an example:

Expected Source System	*Source*	*Estimated Volume*	*Frequency*	*Encrypted?*	*Number of objects?*
For example: Sirius ECC	*For example: 2LIS_06_INV*	*For example: 3 GB*	*For example: Monthly*	*No*	*For example: 3 Files*

3.3.3 Tooling

What tooling is used to extract the data from the source and store the data in the Data Source Layer?

3.4 Source data layer (SA)



3.4.1 Tables

Provide an overview on the environment where source data will be captured. Make a distinction for the different DEV/TST/PRD environments. At this stage, limit yourself to a logical indication only; the technical details will be provided later. Provide also information on the retention period.

As an example:

Data Requirement	*Stored on*	*Database*	*Retention Period*
For example: invoice	*Teradata Dev Server*	*Staging environment*	*Example: Not Applicable*

information			
	Teradata TST Server	*Staging environment*	*Example: Not Applicable*
	Teradata PRD Server	*Staging environment*	*Example: 5 Years*

3.4.2 Transformation

Provide the logic that is applied when lines from the Sources are loaded into the Source data layer (SA). Provide this logic on entity level.

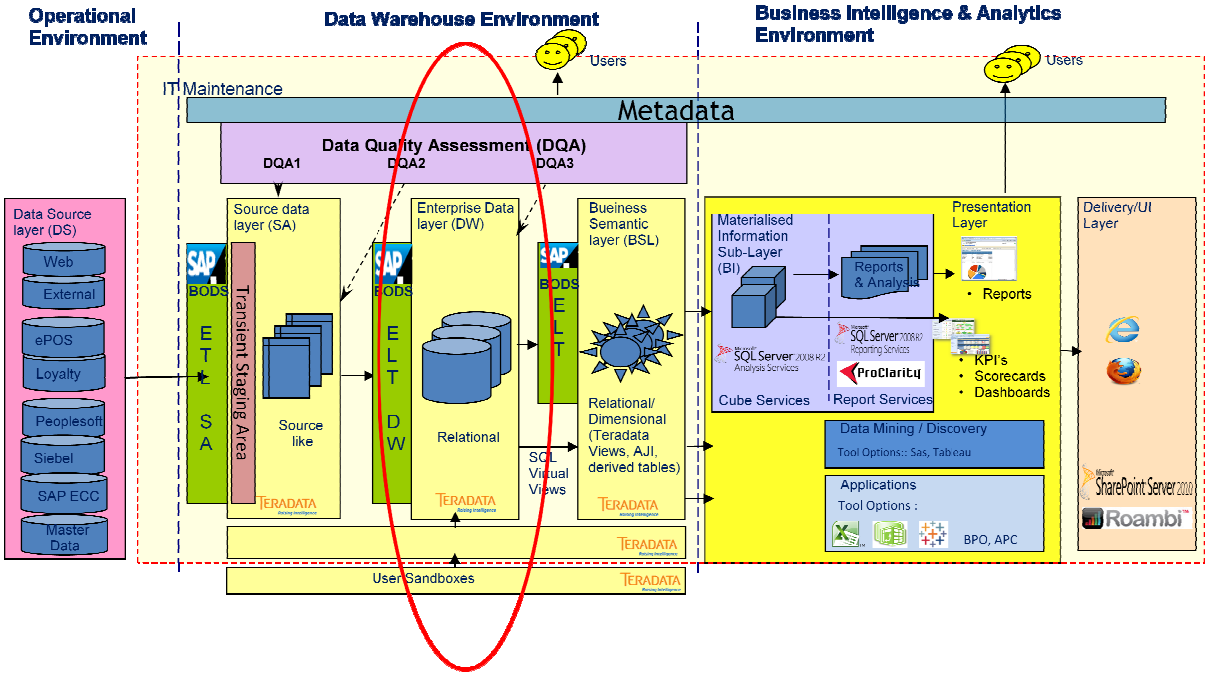
It is assumed that a one-to-one mapping is used here. If other mappings are used, please indicate so.

3.4.3 Tooling

What tooling is used to extract the data from the Data Source Layer (DS) and store the data in the Source data layer (SA)?

If other tools are used than Business Objects Data Services (BODS), what is the reason for that deviation?

3.5 Enterprise data layer (EDL)



3.5.1 Tables

Provide a high level overview on the environment on which the tables in EDL will be stored. Make a distinction for the different DEV/TST/PRD environments. Provide also an indication on the retention period. At this stage, limit yourself to a logical indication only; the technical details will be provided later.

As an example:

Data Requirement	*Stored on*	*Database*	*Retention Period*
For example: invoice information	*Teradata Dev Server*	*EDL environment*	*Example: Not Applicable*
	Teradata TST Server	*EDL environment*	*Example: Not Applicable*
	Teradata PRD Server	*EDL environment*	*Example: 5 Years*

Provide a reference (link only) to the EDL Logical Data Model (LDM) that is created by the project. Indicate below the entities that are used in the project. Indicate their size. Use a scheme like the one below:

Entities in the project	*New Entities?*		*Re-used Entities?*	
	Entity Name	*Size*	*Entity Name*	*Size*
Master Data	*√*		*√*	
Transaction Data	*√*		*√*	

Note 1: It is appreciated that the Logical Data Model (LDM) is not yet signed off when writing this document. If you think that the LDM is already at such a state that is can be usefully used, plse use the link to this LDM.

Note 2: It might be that the Conceptual Data Model (CDM) is better suited to provide this information on an entity level. It that is the case, plse use the CDM instead of the LDM.

3.5.2 Transformation

Provide the high-level logic that is used to load the entities in Enterprise Data layer (EDL) from the Source data layer (SA). Only new entities need to be addressed here. Provide this logic on entity level. Provide also an indication whether this logic is complex or simple.

On a high level, it is expected that events in the Source Data Layer (SA) will generate a one-to-one mapping while inserting records in the Enterprise Data layer (EDL). This is considered as a simple mapping. Indicate whether the proposed logic differs from this. Likewise, it is assumed that master data in the SA generates updates in the EDL. Also here, such mappings are considered as simple mappings. Indicate if the proposed logic will differ. If this information is available in the DIR, a reference (link only) to the DIR is sufficient.

In this chapter, one may include the business specific logic that is needed to create business info. This business info is the solution to the functional requirements. The generation of the business info can be described in this section.

3.5.3 Tooling

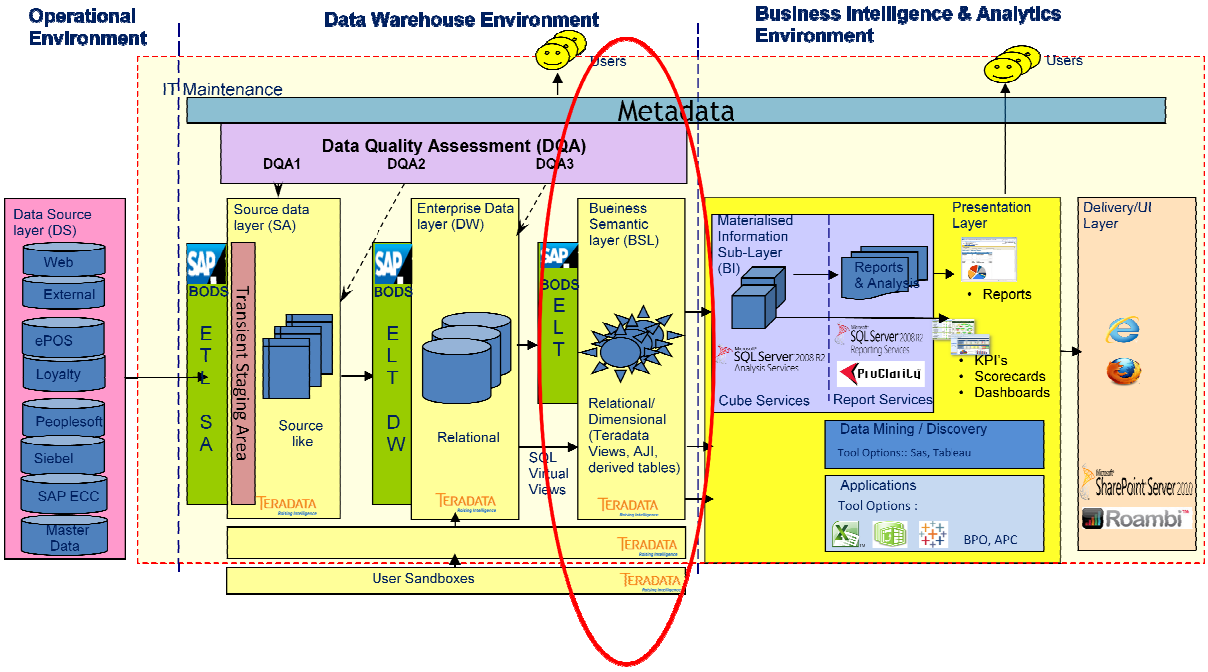
What tooling is used to extract the data from the Source data layer (SA), transform these data and store the data in the Enterprise Data layer (EDL)?

If other tools are used than Business Objects Data Services (BODS), what is the reason for that deviation?

3.5.4 Data cleansing & data quality

What high level data quality / data cleansing processes will be put in place to ensure data is of sufficient quality to be used by the business and support key business processes?

3.6 Business Semantic Layer (BSL)



3.6.1 Views

Provide a high level overview on the environment on which the views in BSL will be stored. Make a distinction for the different DEV/TST/PRD environments. At this stage, limit yourself to a logical indication only; the technical details will be provided later.

As an example:

Data Requirement	*Stored on*	*Database*
For example: invoice information	*Teradata Dev Server*	*Semantic environment*
	Teradata TST Server	*Semantic environment*
	Teradata PRD Server	*Semantic environment*

Provide a reference (link only) to the Logical Data Model for the Semantic Layer that is going to be approved by the Tower.

Indicate which Global Master Hierarchies are used. Indicate which local hierarchies are used.

3.6.2 Transformation

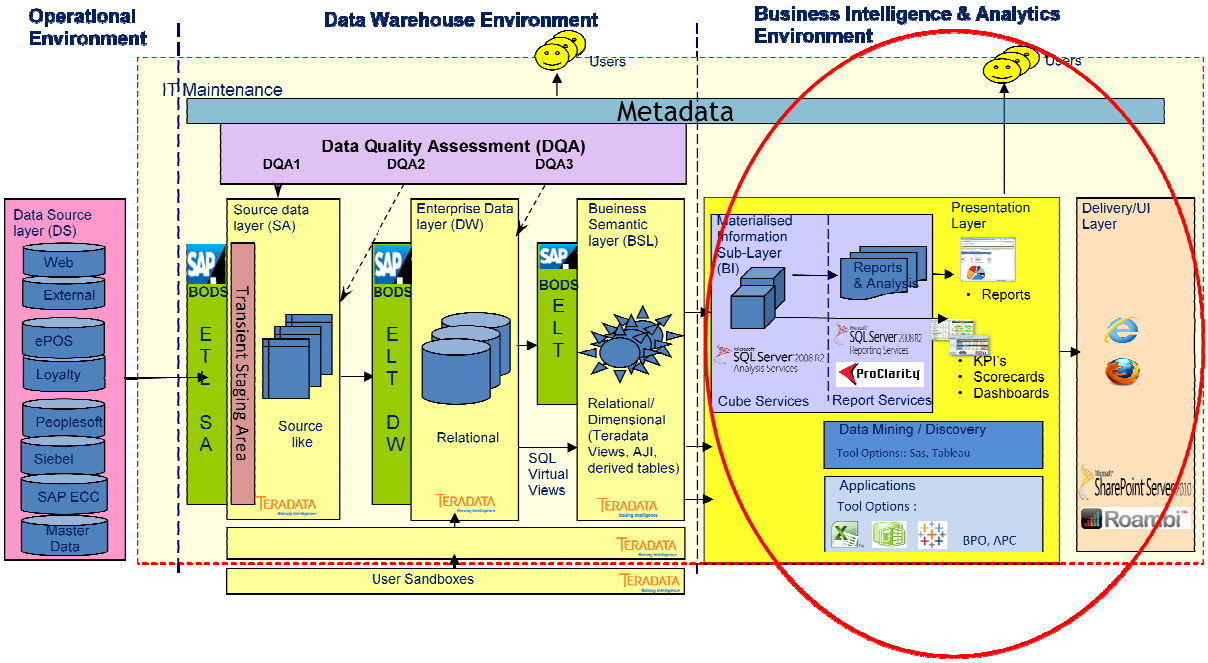
What views will be instantiated in the semantic layer as part of this project and what information will they contain. Provide this logic on entity level. If this information is available in the DIR, a reference to the DIR is sufficient.

It might be necessary to materialise some views for performance reasons. If information is already available at this stage, plse provide this information.

3.6.3 Tooling

What tooling is used to create the views in the Business Semantic Layer?

3.7 Reporting & analytics



3.7.1 Usage of the data from Semantic Layer / EDL for reporting

In this section, the usage of the Semantic Layer / EDL as source for reporting is discussed. Items to be addressed are:

- *What is purpose of the Reporting environment? Is the purpose limited to Reporting and/or Planning only?*
- *Is Semantic Layer the only data source for the Reporting environment?*
- *If the Reporting is limited to master data only: is it really necessary to use the Semantic Layer as a data source (e.g. as master data must be integrated for reporting purposes).*

- *Some systems use downstream data from the Data Warehouse Environment as a System of Records (or the Golden Source). In that case, legal considerations might be applicable.*
- *A project might have written a logical design for the Front End. If relevant information is already captured in this logical design, plse provide a link to this document here.*

3.7.2 Reporting tools

What reporting tools will be deployed / utilised as part of this solution – SQL Server Analysis Server, SQL Server Reporting Services, Excel PowerPivot, Excel etc ?

3.8 Allocation

It might be that EDL is used for allocation purposes. This is understood as data in EDL are used to split external data over different targets. In that case, this section can be used to provide a high-level design. Attention should be given as to what allocation rules are foreseen? What is the level of simplicity of allocation rules; in general EDL is not meant for complicated allocation rules.

3.9 User profiles and security

3.9.1 Users

What will the users be doing with the system when delivered? Will they be heavy analytical users or "lighter" operational users? How many of each type of users are expected and when? Where are the users located and how will they access the tools?

3.9.2 Data Security for data at rest

What high level security measures are implemented to protect data that reside in a database or in a file? Make a difference between the extracts from the Data Sources (DS) and data that are stored in databases in the DSL, EDL and BSL.

3.9.3 Data Security for data in motion

What high level security measures are implemented to protect data in motion? Think of data moving in and out the Data Warehouse Environment. Also think of data between the layers within the Data Warehouse Environment.

3.10 Future BAU requirements

This document is limited to the project only. However some information can easily be collected in this phase that is relevant to the BAU phase. Hence, this section provides a platform to convey this.

BAU requirements		
number of users	*Initial Number Users*	
	Expected growth users / month	
Total Disk space required in Source Data Layer (SA)	*Initial Disk Space*	
	Expected Growth Disk Space/ Month	
Total Disk space required in Enterprise Data Layer (EDL)	*Initial Disk Space*	
	Expected Growth Disk Space/ Month	
New operational support foreseen? If so, which support is foreseen?		
New monitoring requirements raised? If so, which new monitoring requirements?		
Is an upgrade in tooling required? If so, which tools are affected?		
Can incidents or maintenance work on other services impact this service and vice versa?		

4 COMPLIANCE WITH PROGRAM STANDARDS

This section should detail the non compliance with program standards and should refer to the design compliance statement which should also be completed by the project and reviewed by the relevant TDA members.

This chapter builds upon the list "Key Design Decisions" that is provided earlier in this document. This chapter then elaborates on the architecture that lead to the listed decision.