# 

***Concept dossier***

|  |
| --- |
| SMART Reporting Dashboard |



|  |  |  |  |
| --- | --- | --- | --- |
| **Responsibilities** | **Name - Function** | **Department - company** | **Date** |
| Written by | Yiming GUI | Université Toulouse 1 Capitole | 06/03/2023 - 15/03/2023 |
| Validated by | Business Work Package Leader | SMART – Airbus | 21/03/2023 |

# TABLE OF CONTENT

[**1. Contexte 3**](#_heading=h.etnc7x6i4hch)

[1.1. Project presentation: objective, project area 3](#_heading=h.ensmx7w1ya98)

[1.2. Expression of needs 3](#_heading=h.fwm9xvqiv5rx)

[**2. SIAD Functional Architecture 4**](#_heading=h.c5mj5iz2tajc)

[2.1. Datamart : 4](#_heading=h.c4mzur3xdm9u)

[2.2. Source data 4](#_heading=h.kb1hnp6qqp0y)

[2.3. Overall architecture of the decision support system 5](#_heading=h.e31tfwz4ua48)

[**3. Datamart 6**](#_heading=h.2nj6tmg6wk96)

[3.1. Conceptual modelling 6](#_heading=h.h5645psm7cuj)

[**4. Data Warehouse 10**](#_heading=h.fc0knxno3s66)

[4.1. Data modeling 10](#_heading=h.ahlnxaxf86bj)

[4.2. Detailed design 10](#_heading=h.6pqm5cz97182)

[4.3. Treatment modelling 12](#_heading=h.3zlejwm154ac)

[**5. Treatment 13**](#_heading=h.ydy35ek4y6l)

[**6. Conclusion 13**](#_heading=h.ekiskxeub6om)

1. Contexte

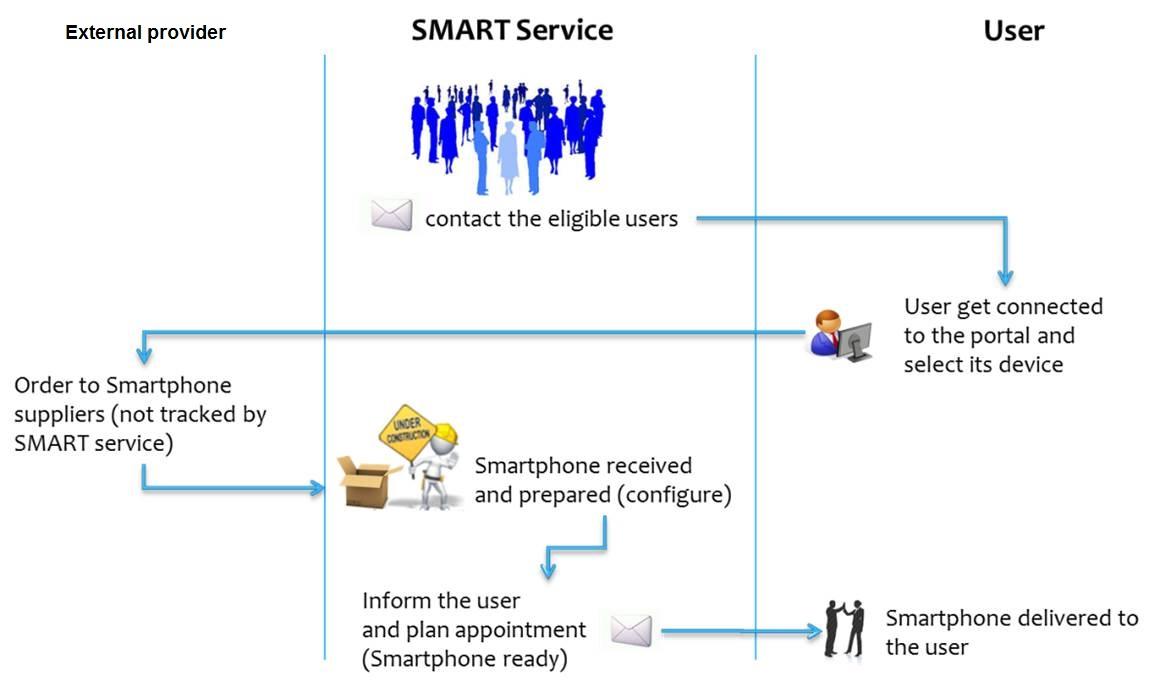
1.1. To-be process overall description

Airbus plans to implement an automated dashboard to track **SMART Service** performance through a set of key performance indicators (KPIs). The dashboard will support end-to-end monitoring of the process shown in the figure: eligible users are identified and contacted, users select devices via the portal, smartphones are received and configured by the SMART team, users are informed and appointments are scheduled, and devices are ultimately delivered to users.

In accordance with existing **Data Governance** rules, access control and information security must be ensured throughout the solution.

**Main objectives:** Monitor the SMART team’s operational activity and overall performance (e.g., follow-up, preparation, delivery, inventory/backlog, and rollout/forecast tracking).  
**Frequency of use:** On demand, with data refreshed **weekly**.  
**Target users:** SMART team members in the four main Airbus countries.  
**Target administrator:** SMART team manager.

**Diagram of the overall process**



Management defines who is eligible, in coordination with the work councils.

1.2. Requirement & Business Specifications

**Nota on priority levels:**

* High = mandatory
* Medium = important
* Low = nice to have

### 1.2.1. General Requirements & Priorities

Airbus wants to implement an automatic dashboard presenting the SMART Service performance via a series of KPIs. However, security must be ensured according to the existing Data Governance rules.

|  |  |  |  |
| --- | --- | --- | --- |
| **Req. ID/ Business Spec. ID** | **Short Name** | **Description** | **Priority** |
| GC1 | Security | Access and information security must be insured | High |
| GC2 | Roles | A minimum of 2 roles must be defined to segregate administrator and users | Medium |
| GC3 | SSO | Single-sign on should be available for the solution | Medium |
| GC4 | One-pager | All 4 KPIs should be presented on a single page | High |

### 1.2.2. Data Sources and extraction rules & data retention period:

*A detailed file is available and delivered by Business Technical Representatives.*

| **Req. ID** | **Short Name** | **Description** | **Priority** |
| --- | --- | --- | --- |
| DS1 | Loading | No user action during loading (automation) | Medium |
| DS2 | Loading frequency | Data will be loaded every week | Medium |
| DS3 | Extraction Data quality log | Data quality log is available during file integration process with all errors | High |
| DS4 | Empty cells | Replace empty cells by N/A, in order to improve data quality except for data rules described below | High |
| DS5 | Specific character | Description column: Specific characters have been found and should be taken into account by the future analysis tool, such as “ ” and # and [ ] | Low |
| DS7 | Week number | Week number need to be captured during the data integration (if not available in the file itself) | High |
| DS8 | Retention period | Data should be kept for 2 years | Medium |
| DS9 | Data files | Each data file is unique (no overlap period) | High |

### 1.2.3. GUI (graphic user interface):

*The GUI is at first for one user, but could be extending for many users.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Req. ID** | **Short Name** | **Description** | **Priority** |
| RT1 | GUI | The tool is Web based | Medium |
| RT2 | Local | Tool is accessible in local mode (no network) offline mode | Low |
| RT3 | Dashboard | All KPI must be presented in a common page | High |
| RT4 | Colour interface | Possibility to change colours of graphs | Low |
| RT5 | Logos | Possibility to add logos | Medium |
| RT6 | Comments | Possibility to add comments by report | Low |
| RT7 | PowerPoint | Export analysis in PowerPoint presentation | Low |
| RT8 | Drill down | Possibility to use drill down in analysis | High |
| RT9 | Country filter | Possible to choose one or many countries | High |
| RT10 | Company filter | Possible to choose one or many Business Unit | High |
| RT11 | Flexible filter definition | User is able to setup any filter on any data | Medium |
| RT12 | Unified CLAF | Common Look & Feel is needed across KPIs | Medium |
| RT13 | Colour interface | No red in the graph | High |

### 1.2.4. Cross KPIs business rules:

*All KPIs are calculated according to single business rules.*

|  |  |  |  |
| --- | --- | --- | --- |
| **Req. ID** | **Short Name** | **Description** | **Priority** |
| CR1 | Reporting period | Reporting period is a variable that can be changed | High |
| CR2 | Refresh | Historisation and data refresh are on a weekly basis | High |

1.3. KPIs

### 1.3.1. KPI1 : Smartphone logistic weekly follow up

This KPi aims at measuring the overall process status at a particular moment of time:

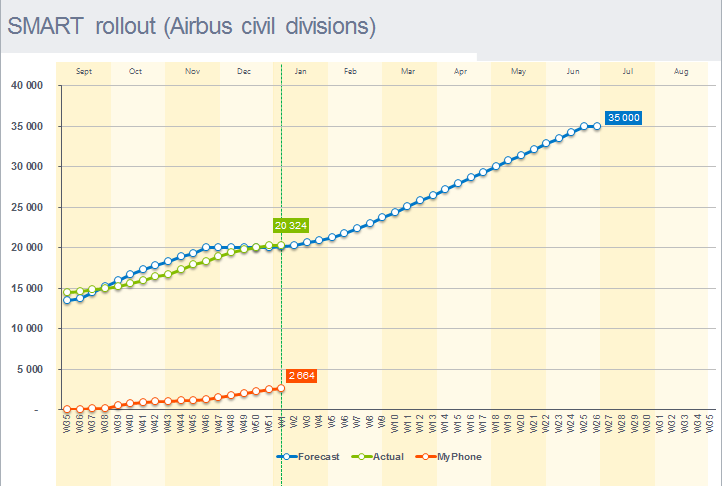


***Format examples***

|  |  |  |  |
| --- | --- | --- | --- |
| **Req. ID/ Business Spec. ID** | **Short Name** | **Description** | **Priority** |
| KPI1-1-1 | Eligible users contacted | Number of users contacted to announce they are eligible | High |
| KPI1-1-2 | Smartphone ordered | Number of smartphones selected by eligible users | High |
| KPI1-1-3 | Smartphone received | Number of smartphones received from the suppliers | High |
| KPI1-1-4 | Smartphone prepared | Number of smartphones prepared and ready to be distributed to the users | High |
| KPI1-1-5 | User informed | Number of appointments booked with users | High |
| KPI1-1-6 | Smartphone delivered | Number of Smartphone delivered (appointment done) | High |
| KPI1-2 | Period | Report will be launch a specific week # or month or quarter or year | High |
| KPI1-3 | Value (+x) | Current period value compared to previous period value (warning if a range is selected) | Medium |

### 1.3.2. KPI 2: SMART roll out follow-up

This KPi aims at presenting the deployment performance compared to the objectives:

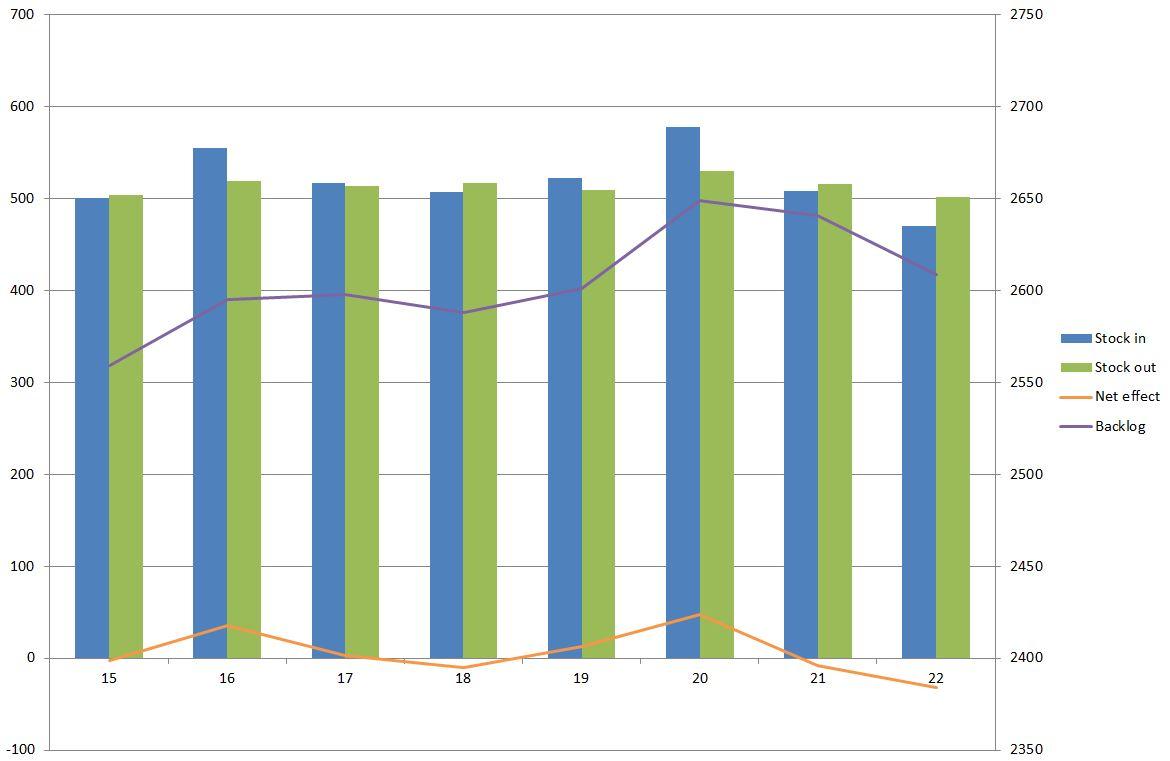


***Format examples***

|  |  |  |  |
| --- | --- | --- | --- |
| **Req. ID/ Business Spec. ID** | **Short Name** | **Description** | **Priority** |
| KPI2-1-1 | Target | Forecasted number of smartphones to be delivered | High |
| KPI2-1-2 | Inputs | A forecast input interface shall be provided (figures entered manually and saved within the database) |  |
| KPI2-2 | Actual | Number of smartphones really delivered to users | High |
| KPI2-3 | MyPhone | Number of smartphones BYOD | High |
| KPI2-4 | Time range | The report should allow user to select a time range | Medium |
| KPI2-5 | Selection criteria | The report should allow user to select a time a country / a business unit etc. | High |
| KPI2-6 | Forecast | The system should calculate the future deliveries according to the past ones | High |
| KPI2-7 | Alert | The system must raise an alert in case of deviation between forecast and planned targets | High |

### 1.3.3. KPI 3: SMART logistic Inventory

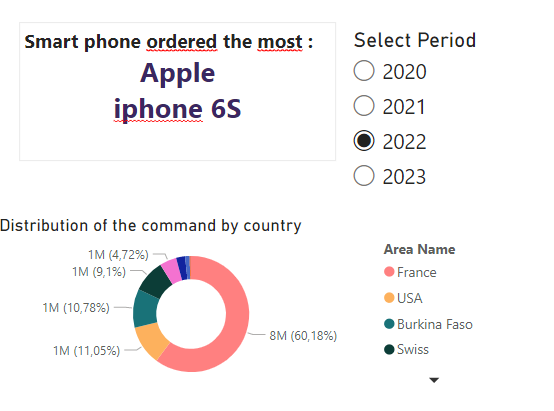
This KPI aims to show number of people contacted and still in progress by week



***Format examples***

|  |  |  |  |
| --- | --- | --- | --- |
| **Req. ID/ Business Spec. ID** | **Short Name** | **Description** | **Priority** |
| KPI 3-100 | Stock in | = Smartphones received from the suppliers | High |
| KPI 3-101 | Stock out | = Smartphones delivered to the end users | High |
| KPI 3-103 | Net effect | = Stock in – Stock out | High |
| KPI 3-104 | Backlog | = Cumulative value of the net effect | Medium |
| KPI3-2 | Jump | By clicking on a the graph we should be able to go to the others KPI filtered on this week | Low |
| KPI3-3 | Selection criteria | The report should allow user to select a time range / a country / a business unit etc. | High |

### 1.3.4. KPI 4: The most popular smartphone statistics



Q1

Q2

Q3

|  |  |  |  |
| --- | --- | --- | --- |
| Req. ID/ Business Spec. ID | Short Name | Description | Priority |
| KPI 4-1-1 | Popular Brand | The most selected smartphone’s brand during a period | High |
| KPI 4-1-2 | Popular Model | The most selected smartphone’s model during a period | High |
| KPI 4-1-2 | Country/subsidiary | The percentage of the orders of the popular phone by country | High |
| KPI 4-2 | Period selection | The rapport should allow to select a specific period (week, month, quarter, or year) | High |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **REFERENCE DOCUMENTS** | | | | | |
|  |  |  |  |  |  |
| **A/R** | **Index** | **Title** | **Reference** | **Issue** | **Date** |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

Note on column A/R:

*‘A’ indicates that the document is applicable.*

‘*R’ indicates that the document is referred to in the text.*

|  |  |  |
| --- | --- | --- |
| **APPROVAL** | | |
|  |  |  |
| **Name – Function** | **Department -**  **Company** | **Date** |
| Line MARTINEZ – Business Representative | M2 ISIAD | 06-Dec-2019 |
| Franck DEBREILLY – Business Representative | M2 ISIAD | 06-Dec-2019 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **TABLE OF REVISIONS** | | | | |
|  |  |  |  |  |
| **Issue** | **Date** | **Modified by** | **Modified sections** | **Observations** |
| 1.0 | 06-Dec-2019 | Business Representatives | All | Document Creation |
|  |  |  |  |  |

1.4. Project presentation: objective, project area

SMART service aims at providing to each Airbus employee a smartphone with Airbus master and Airbus apps .

They had called on our service to help them improve their tracing system including the principal missions bellow :

• Defining the smartphones catalog which is offered to employees

• Deliver smartphones to employees in each location of Airbus in the world taking into consideration local specificities (Germany do not have same legal rules as France)

• Manage the backbone that manages the professional tools on smartphones

• Manage the service and after sales support

• Define strategy for future evolutions

There are 3 systems not intergrated with major referential allow activity follow-ups : FLEET, BES, SMART portal. Therefore, we will create a dashboard to follow the SMART service team activities based on data generated by these 3 referentials.

1.5. Expression of needs

|  |  |  |  |
| --- | --- | --- | --- |
| Requirement reference | Decision-making class | Decision-making class | Description |
| B1 | Service Smart | Smartphone logistic weekly follow up | B1.1.1 : Number of users contacted to announce they are eligible |
| B1.1..2 : Number of smartphones selected by eligible users |
| B1.1.3 : Number of smartphones received from the suppliers |
| B1.1.4 : Number of smartphones prepared and ready to be distributed to the users |
| B1.1.5 : Number of appointments booked with users |
| B1.1.6 : Number of Smartphone delivered (appointment done) |
| B1.2 : Period |
| B1.2 : Value (+x) |
| B2 | Service Smart | SMART roll out follow-up | B2.1.1 : Forecasted number of smartphones to be delivered |
| B2.1.2 : A forecast input interface shall be provided (figures entered manually and saved within the database) |
| B2.2 : Number of smartphones really delivered to users |
| B2.3 :  Number of smartphones BYOD |
| B2.4: The report should allow user to select a time range |
| B2.5: The report should allow user to select a time a country / a business unit etc. |
| B2.6: The system should calculate the future deliveries according to the past ones |
| B2.7: The system must raise an alert in case of deviation between forecast and planned targets |
| B3 | Service Smart | SMART logistic Inventory | B3.1.1 : Smartphones received from the suppliers |
| B3.1.2 : Smartphones delivered to the end users |
| B3.1.3: Net effect (= Stock in – Stock out ) |
| B3.1.4 : Cumulative value of the net effect |
| B3.2: Jump on the graph (to go to the others KPI filtered on this week) |
| B3.3 : select criteria (a time range / a country / a business unit etc.) |
| B4 | Service Smart | The most popular smartphone statistics | B4.1.1: The most selected smartphone brand compared to other brands |
| B4.1.2: The most selected smartphone model compared to other brands |
| B4.2:The rapport should allow to select a specific period (week, month, quarter or year) |

1. SIAD Functional Architecture

2.1. Datamart :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Identifiant | Description | Utilisateurs | Outils de stockage | Outils de restitution |
| Mag\_SMART | To create a tool to follow the SMART service team activities | Service Smart | DMS [Database Management System]  (Oracle) | indicators, graphs, … |

2.2. Source data :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Source code** | | **FLEET** | **BES** | **SMART portal** |
| Description | | Follow fleet activities •Based on employee situation | •Follow smartphones activities •Based on smartphone status | •Provide selfservice offers to employees on their smartphones •Based on requests status |
| Owner | | SMART service | SMART service | Smart service |
| Localisation | | internal | internal | internal |
| Type | | not integrated referencial | not integrated referencial | not integrated referencial |
| Support | Stockage | Excel | Excel | Excel |
| Logiciel | Oracle | Oracle | Oracle |
| Accessibility | | Easily | Easily | Easily |
| Refreshment | | Weekly | Weekly | Weekly |
| Legal limitation | | /NA | /NA | /NA |
| Volumetry | | small | small | small |
| Data quality | | Incomplete | Incomplete | Incomplete |
| Data dictionary | | Yes | Yes | Yes |

## 2.3. Overall architecture of the decision support system :

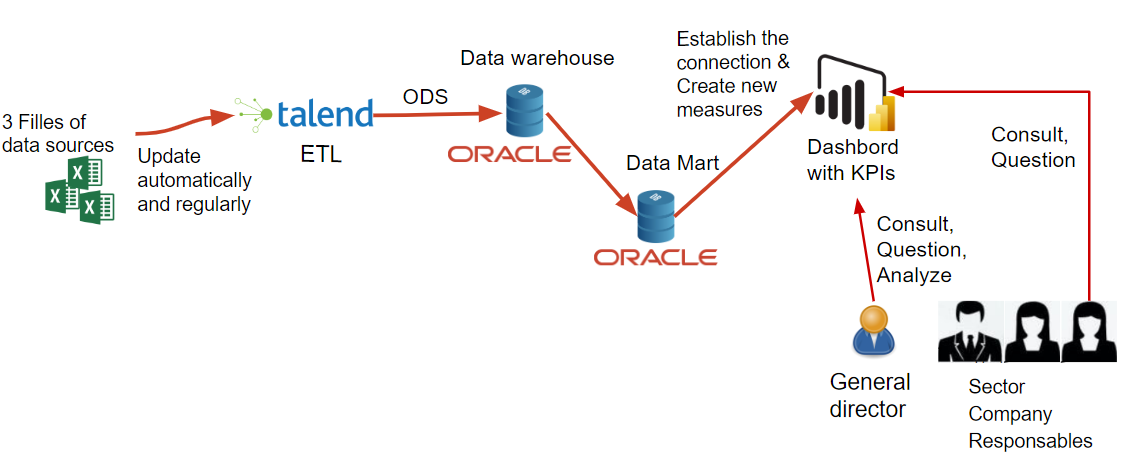
* **Source/store matching table**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **BES** | **DB Portal** | **Stock files** |
| Mag\_SMART | X | X | X |

* **Detection of the need for a Data Warehouse or not**

Firstly, for the automatisation of loading the data into User Interface regularly with the data source generated by 3 references of company’s system, a stockage of data and a connection with UI in the upsteam is needed to realize this requirement. Secondly, the selection of Business Unit, Company and time Period criteria require a separate data structure from the other tables.

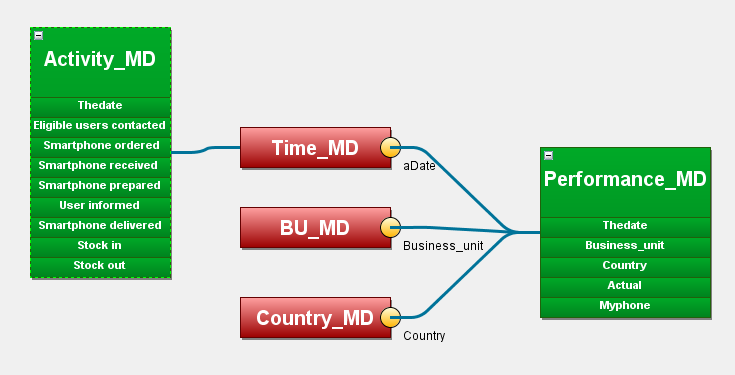
* **Global scheme**



First the process of ETL is needed to clean the data with errors of format and null cases etc., and then to automatically receive the sources files, 3 ODS jobs through Talend is need to transform the first version of cleaned data into our data warehouse, then to build our warehouse in order to reorganize its’ structure into Data mart for further creation of measures for KPIs. At the last process, once the KPI are formed, they will be shared within the responsables within the companies.

1. Datamart

3.1. Conceptual modelling

* **Multidimensional scheme**

Data model (star schema) for the SMART Service dashboard

* **Detailed definition of facts: measurement dictionary**

|  |  |  |  |
| --- | --- | --- | --- |
| **Measure name** | **Description** | **Type** | **Extraction formula** |
| Eligible users contacted | nb of eligible users contacted in a given week/month | Integer | = COUNT(EMPLOYEE.LoginID) |
| Smartphone ordered | nb of smartphone selected by eligible users in a given week/month | Integer | =COUNT(SMARTPHONE.IDPHONE) |
| Smartphone received | nb of smartphone received from the suppliers in a given week/month | Integer | = sum(STOCK\_IN.QteReceived) |
| Smartphone prepared | nb of smartphones prepared and ready to be distributed to the users in a given week/month | Integer | =sum(STOCK\_OUT.QtePrepared) |
| User informed | nb of appointments booked with users in a given week/month | Integer | = count(APPOINTMENT.Appointment\_scheduled\_on) |
| Smartphone delivered | nb of Smartphone delivered (appointment done) in a given week/month | Integer | = count(DELIVERY.Delivery\_done\_on) |
| Target | Forecasted number of smartphones to be delivered | Integer | N/A |
| Actual | Number of smartphones really delivered to users who work in a company belong to 1 BU in 1 country in 1 given month | Integer | =count(DELIVERY.MODEL) |
| MyPhone | Number of smartphones BYOD | Integer | =count(Activate.Owership) |
| Stock in | = Smartphones received from the suppliers | Integer | =SUM(Stock\_in.QtyReceived) |
| Stock out | = Smartphones delivered to the end users | Integer | =count(Delivery.DeliveryDate) |
| Net effect | = Stock in – Stock out | Integer | =SUM(Stock\_in.QtyReceived) - count(Delivery.DeliveryDate) |
| Backlog | = Cumulative value of the net effect | Integer | =SUM(SUM(Stock\_in.QtyReceived) - count(Delivery.DeliveryDate)) |
| Popular Brand | The most selected smartphone’s brand during a period | String | =(count(selectphone.model)) |
| Popular Model | The most selected smartphone’s model during a period | String | =(count(selectphone.model)) |
| Country/subsidiary | The percentage of the orders of the popular phone by country | Float | =(count(selectphone.model)) |

* **Detailed definition of dimensions: attributes dictionary**

|  |  |  |  |
| --- | --- | --- | --- |
| code | designation | type | extraction formula |
| LoginId | The identification of the user who logged in smart system | Integer | employee.LoginId |
| Eligible\_Flag | indicates wether the user is eligible or not | CC(50) | employee.Eligible\_Flag |
| Department | the department / Business Unit of a company | CC(50) | employee.Department |
| Company | the code of company | CC(50) | employee.Company |
| CompanyName | the name of the company | CC(50) | employee.CompanyName |
| Country | the country of the company’s location | CC(50) | employee.Country |
| PhoneId | the identification code of the phone | Integer | Smartphone.PhoneId |
| Ownership | indicates wether the phone a BYOD or a corporate possession | CC(50) | Smartphone.Ownership |
| Model | the model of the smart phone | CC(50) | Smartphone.Model |
| Brand | the brand of the smart phone | CC(50) | Smartphone. |
| Date | the date that the stockage /sales took / will take place | Date (DD/MM/YYYY) | to\_char(Date DD/MM/YYYY) |
| Week | the week that the stockage or sale takes / will take place | Integer | to\_char(Date DD/MM/YYYY) |
| Month | the month that the stockage or sale takes / will take place | Inetger | to\_char(Date ‘MM’) |
| Quarter | the quarter that the stockage or sale takes / will take place | Inetegr | to\_char(Date, 'Q') |
| Year | the year that the stockage or sale takes / will take place | Integer | to\_char(Date, 'YYYY’') |

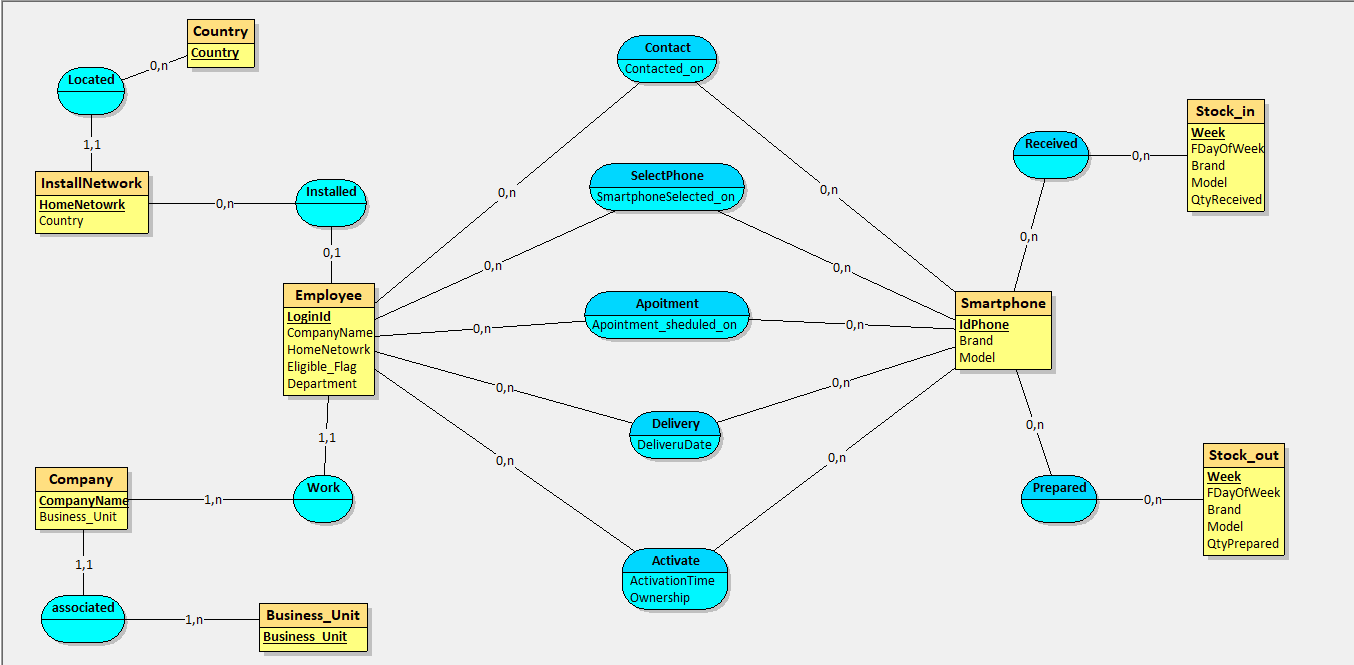
1. Data Warehouse

4.1. Data modeling

* **Necessary data**

|  |  |  |
| --- | --- | --- |
| Data store | Attributes | Source |
| Mag\_SMART | THEDATE  BUSINESS\_UNIT  COUNTRY  ACTUAL  MYPHONE  ADATE  AWEEK  AMONTH  AQUARTER  AYEAR  ELIGIBLE\_USERS\_CONTACTED  SMARTPHONE\_ORDERED  SMARTPHONE\_RECEIVED  SMARTPHONE\_PREPARED  USER\_INFORMED  SMARTPHONE\_DELIVERED  STOCK\_IN  STOCK\_OUT | =DIM\_TEMPS.thedate  =BU.BUSINESS\_UNIT  =COUNTRY.country  (details of calculation listed on the measurement dictionary) |

* **Detailed conception**
* The onceptual data diagram:



* The associated data dictionary:

|  |  |  |  |
| --- | --- | --- | --- |
| **Code** | **Description** | **Type** | **Constraints** |
| ActivationTime | The date when a smartphone was activated by an employee | Date (DD/MM/YYYY) | <= sysdate |
| Apointment\_sheduled\_on | The appointment date set with an employee for a smartphone | Date (DD/MM/YYYY) | <= sysdate |
| Brand | The brand of a smartphone. | CC(50) |  |
| Business\_Unit |  | CC(50) |  |
| CompanyName | The name of one company. | CC(50) |  |
| Contacted\_on | The date an employee was contacted to announce eligibility | Date (DD/MM/YYYY) | <= sysdate |
| Country |  | CC(50) |  |
| DeliveryDate | The date when the new smartphone is delivered to one employee. | Date (DD/MM/YYYY) | <= sysdate |
| Department | The department where one employee works. | CC(50) |  |
| Eligible\_Flag | The status to see if one employee is eligible for a new smartphone. | CC(50) | {Y,N} |
| FdayOfWeekP | The first day of the smartphone week. | Date (DD/MM/YYYY) | <= sysdate |
| FdayOfWeekR | The first day of the week of smartphone reception. | Date (DD/MM/YYYY) | <= sysdate |
| HomeNetowrk |  | CC(50) |  |
| LoginId | The login Id for one employee. | Integer |  |
| Model | The model of a smartphone. | CC(50) |  |
| Ownership | The status to show if one phone is owned by a person or corporate. | CC(50) | {Personal,Corporate} |
| PhoneId | The id for a smartphone. | Integer |  |
| QtyPrepared | The quantity of the smartphone prepared for delivery. | Integer | >=0 |
| QtyReceived | The quantity of the smartphone received. | Integer | >=0 |
| Smartphone\_selected\_on | The date an eligible employee selected a smartphone | Date (DD/MM/YYYY) | <= sysdate |
| WeekP | The week when the smartphone is being prepared. | Integer | >0 |
| WeekR | The week when the smartphone is received. | Integer | >0 |

* Relational logic diagram deduced

Country (Country)

InstallNetowrk(HomeNetowrk, #Country)

Company (CompanyName, #Business\_Unit)

Business\_Unit (Business\_Unit)

Employee ( LoginId, Eligible\_Flag, Department, #CompanyName, #HomeNetowrk )

Contact (Contacted\_on, #LoginId, #Model, #Brand)

SelectPhone (Smartphone\_selected\_on, #LoginId, #Model, #Brand)

Appointment (Apointment\_sheduled\_on, #LoginId, #Model, #Brand)

Delivery (DeliveryDate, #LoginId, #Model, #Brand)

Activate (ActivationTime, Ownership, #LoginId, #Model, #Brand)

Smartphone (IdPhone, Model, Brand)

Stock\_in (Week, QtyReceived, FDayofWeekR, #Model, #Brand)

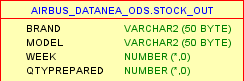
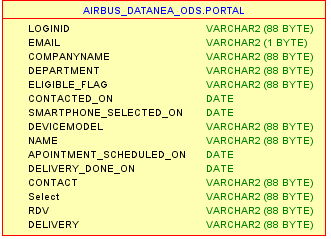
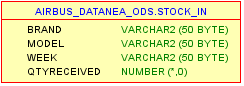
Stock\_out (Week, QtyPrepared, FDayofWeekR, #Model, #Brand)

4.2. Treatment modelling

* **Choice of architecture (ODS)**

For the first phase of ETL(extract, transform and load) processus, we have put a ODS (Operational Data Store Staging Area) for collecting the source data in amont. This step is to extract the flat data to form the same amount of source table for preparing the data warehouse and correct the inconsistent data format in an initial step.

Bellow are the schemes of ODS source tables structure.

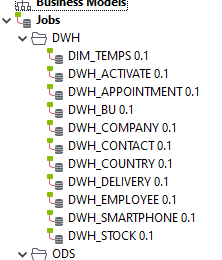


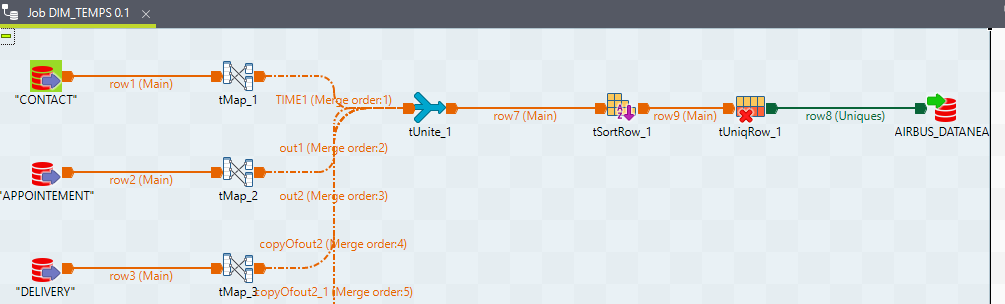
* **ETL**:

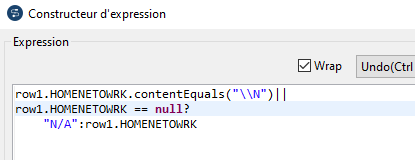
In this step, we have added the process of transforming the data into relational data warehouse from the ODS tables. In each jobs in Talend application, we had set a updating process to compare the old data with new incoming rows, the updated values comparing to existing ones etc. This is to make sure our data in the Warehouse is the latest version and corresponding the real data represented by the smart service systems.

The function of automatic updating will be lunched each monday at 7 am to catch up the new datas, to realize it we had also set a plan in the talend parameter to evoke a updating process once we will receive the data from our shared directory in Sharepoint.

Bellow is two exemples of firsty the jobs we created through talend, each job contain at least one output table connecting with data warehouse, and a specific job concerning the time table to prepare our time dimension in Datamart.







1. Treatment Description

* **1) Source Systems and Data Flow**

In the upstream layer, SMART Service operational data is collected from three primary repositories: **BES**, **SMART Portal**, and **Fleet** exports from the shared directory. These sources provide the raw events required to monitor the end-to-end logistics process (contacted → ordered → received → prepared → informed → delivered). Data is extracted on a scheduled basis and delivered as **delimited flat files**, which are then ingested into the ETL pipeline.

* **2) ETL Pipeline (Staging → Standardization → Cleansing → Load)**

The ETL process is implemented in **Talend** and follows a staged approach:

* **Ingestion/Staging:** Raw files are loaded into a staging layer with minimal transformation to preserve the original payload for auditability.
* **Standardization:** Fields are normalized to a consistent format (e.g., converting activation dates to **dd-MM-yyyy**). For malformed attributes (unexpected length, extra characters, mixed tokens), parsing rules are applied to **split strings**, isolate valid segments, and rebuild compliant values.
* **Data Cleansing & Validation:** Common data quality issues—**manual spelling errors**, **inconsistent values**, and **null/missing data**—are addressed through validation rules and correction logic.

To ensure reusability and consistency, data quality checks are encapsulated as **Java routines** stored in Talend’s **Routines** repository and reused within **tMap** components. This design supports centralized rule maintenance and reduces duplication across jobs.

* **3) Traceability and Data Lineage**

To guarantee traceability, the pipeline maintains end-to-end visibility from source to target by:

* Preserving raw extracts in the staging layer (source snapshot)
* Capturing technical metadata such as **source system**, **file name**, **load timestamp**, and **record identifiers**
* Enabling reconciliation between source counts and loaded counts at each step (staging → curated → warehouse)

This approach supports data lineage analysis and simplifies audit/replay scenarios when upstream data changes or corrections are required.

* **4) Error Logging and Exception Handling**

Robust operational controls are included to manage failures and data anomalies:

* **Row-level error handling:** Records failing validation rules are redirected to a reject flow (quarantine) with an error reason code.
* **Job-level exception handling:** Talend jobs implement try/catch-style handling and controlled exits for connection failures, schema mismatches, or corrupted files.
* **Error logs and monitoring:** Execution logs capture processed/rejected counts, transformation steps, and exceptions. Reject datasets and log files are retained to support debugging and root-cause analysis.
* **5) Loading into Oracle Data Warehouse and Analytics Layer**

After validation and cleansing, curated data is loaded into the **Oracle** data warehouse. The model is structured to support analytics with key dimensions such as **Business Unit**, **Country**, and **Time**. The curated warehouse layer produces a consistent **data mart** (e.g., *Mag\_Smart*) that serves as the governed dataset for reporting.

* **6) Power BI Consumption**

Once the data mart is available, it is connected to **Power BI** for KPI reporting. Measures and visuals are built on top of the curated model, ensuring consistent definitions across the dashboard and enabling interactive slicing (time, business unit, country) with reliable, traceable data.

1. Conclusion

|  |  |  |
| --- | --- | --- |
| Needs | Restitution | Data Store |
| B1 | B1.1.1  B1.1.2  B1.1.3  B1.1.4  B1.1.5  B1.1.6  B1.2  B1.2 | Mag\_SMART |
| B2 | B2.1.1  B2.1.2  B2.2  B2.3  B2.4  B2.5  B2.6  B2.7 | Mag\_SMART |
| B3 | B3.1.1  B3.1.2  B3.1.3  B3.1.4  B3.2  B3.3 | Mag\_SMART |
| B4 | B4.1.1  B4.1.2  B4.2 | Mag\_SMART |

To conclude, SMART Service expressed three main reporting needs: **(1)** weekly follow-up of smartphone logistics KPIs, **(2)** monitoring of the **SMART rollout** with forecast comparison and deviation alerts, and **(3)** tracking **SMART logistics inventory and backlog**. In addition, we proposed an extra analytical view—**order distribution by country and the most popular smartphone model**—to help the service manager quickly identify geographic concentration and product preferences.

All requirements are consolidated into a **single-page executive dashboard**, where each high-level KPI is broken down into more detailed indicators and can be filtered by **year, month, week, business unit, and country**. The dashboard also includes **“Run with Week”** and **“Run with Month”** controls to dynamically animate all visuals over time for presentation and trend analysis.

Finally, we will create a centralized data store, **Mag\_Smart**, that contains the curated data required to compute these indicators. A **single data store** is sufficient to support all reporting needs while ensuring compliance with existing **Data Governance and security** rules and enabling a **weekly refresh**.