



**HAL**  
open science

# A proposition of service architecture and best practice for business intelligence projects within an IT department of an international bank

Xu Gao

► **To cite this version:**

Xu Gao. A proposition of service architecture and best practice for business intelligence projects within an IT department of an international bank. Computational Engineering, Finance, and Science [cs.CE]. 2013. dumas-01270804

**HAL Id: dumas-01270804**

**<https://dumas.ccsd.cnrs.fr/dumas-01270804>**

Submitted on 8 Feb 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**CONSERVATOIRE NATIONAL DES ARTS ET METIERS**  
**PARIS**

---

Mémoire présenté en vue d'obtenir

Le DIPLOME D'INGÉNIEUR C.N.A.M

SPECIALITE : Informatique

OPTION : Architecture et Ingénierie des Systèmes et des Logiciels

par

**Xu GAO**

**Supervisor: Mr. Yves LALOUM**

---

**A Proposition of Service Architecture and Best Practice for  
Business Intelligence Projects within an IT Department of an  
International Bank**

Soutenu le 26 mars 2013

**JURY :**

Professeur *Yves LALOUM* du CNAM (Président)

Professeur *Nicolas TREVES* du CNAM

Mr. *Patrice LIGNELET* du CNAM

Mr. *Laurent AVET* de BNP Paribas

Mr. *Philippe MULLOT* de BNP Paribas

**A Proposition of Service Architecture and Best Practice for  
Business Intelligence Projects within an IT Department of an  
International Bank.**

Memoire d' ingénieur C.N.A.M., Paris 2013

## **Abstract**

Business Intelligence system plays an increasingly important role in the industry and commerce world in that it can help high level users making decision and figuring out business strategy. In this internship report, we will talk about the basic knowledge of BI system. More over, we will give a proposition of the rules and best practices of BI system in the specific context of the IT Group of BNP Paribas. A general software service reference model defined for BI systems in the context of the group will also be discussed.

## **Key words**

Business Intelligence, Software Architecture, Software Service, Best Practice

## **Résumé**

Le système d'information décisionnelle joue un rôle plus en plus important dans le domaine industriel et le domaine commercial parce qu'il peut aider des utilisateurs de haut niveau quand ils prennent décision et proposent stratégie métier. Dans ce mémoire , nous allons parler de connaissance basique du Le système d'information décisionnelle. Au même temps, nous proposons des règles et des bonnes pratiques dans le cadre de IT groupe de BNP Paribas. Aussi une architecture de reference serai diffusée.

## **Mots clés**

Système d'information décisionnelle, architecture logicielle, service logiciel, bonne pratique

## **Appreciation**

For my internship and my internship report, firstly I will thank my tutor from Conservatoire National des Arts et Métiers, Mr. Yves LALOUM. He arranges a meeting every month to talk about my internship and my report. He checks my report carefully and cares about my internship seriously.

And I will thank all the members present in my presentation.

And also I will thank Mr. Laurent AVET and Mr. Philippe Mullot. Thank them for giving me the opportunity of doing an internship in the team of Reference Technical Architecture of Infrastructure & Production Services, which is a branch of the department of Group IT Function in BNP Paribas group. I really appreciate them for their help during my learning the context of the team and the background of the project, as well as their patience when they taught me the knowledge of Business Intelligence and software architecture.

I will also thank all the colleagues of the team of Reference Technical Architecture. Thank them for all their support when they helped me integrate in the group.

Meanwhile I will thank the Business Intelligence Competency Center and the Master Data Management Community of BNP Paribas Group. Thank them for providing me the necessary documents for my project.

## Table of Content

Abstract.....	2
Key words.....	2
Résumé .....	3
Mots clés.....	3
Appreciation.....	4
Table of Content .....	5
Table of Figures.....	8
0 Introduction .....	9
1 Organization and Context .....	10
1.1 Organization .....	10
1.1.1 BNP Paribas.....	10
1.1.2 Information Technology and Processes (Technologies et Processus) .....	12
1.1.3 Group IT Function (Informatique & Technologies Groupe).....	13
1.1.4 Infrastructure & Production Services (Infrastructure et Production Services) .....	13
1.1.5 Reference Technical Architecture (Missions & Organization).....	14
1.1.6 Business Intelligence Competency Center.....	14
1.2 Context .....	16
1.2.1 Enterprise Architecture Group Language .....	16
1.2.2 Software Integration Patterns .....	21
1.2.3 Generic Software Service Reference Model .....	23
1.3 Reference Technical Architecture .....	27
1.3.1 IT Vision.....	29
1.3.2 Enterprise Technical Rules & Guidelines .....	29
1.3.3 Standard Technical Components .....	29
1.3.4 Standard Technical Services.....	30
1.3.5 Standard Technical Architecture.....	30
1.3.6 IT Roadmaps.....	31
2 Business Intelligence .....	32
2.1 Introduction .....	32
2.1.1 Historic .....	32

2.1.2	Definition .....	32
2.1.3	Usage.....	34
2.1.4	Advantage .....	35
2.2	Foundation of Business Intelligence System Architecture .....	35
2.3	The Component of Business Intelligence System.....	37
2.3.1	Collecting System .....	38
2.3.2	Integration System.....	38
2.3.3	Data Staging Area.....	40
2.3.4	Data Layer .....	40
2.3.5	On – Line Analytical Processing (OLAP) .....	44
2.3.6	Data Mining.....	45
2.3.7	Data Presentation .....	45
3	Presentation of the Task .....	47
3.1	Introduction.....	47
3.2	The Client.....	48
3.3	The Requirement (The Stake).....	49
3.4	The Product .....	50
3.5	The Plan .....	51
3.5.1	The Beginning of the Project.....	53
3.5.2	The Process of the Project .....	53
3.6	The Situation before BI ITRefCard .....	54
3.7	Working Environment.....	55
4	Project: ITRefCard .....	56
4.1	The Introduction of Business Intelligence in BNP Paribas Group.....	56
4.1.1	Benefits .....	56
4.1.2	Attention Points.....	57
4.1.3	BNPParibas Context and Strategy.....	58
4.2	BI - SSRM.....	59
4.2.1	Business Intelligence Software Service Reference Model .....	59
4.2.2	Reference Architecture .....	64
4.2.3	Platform Instantiations .....	67
4.3	Governance Elements.....	72
4.4	Rules and Guidelines .....	72
4.4.1	Modularity and Autonomy.....	73



4.4.2	Data Responsibility & Access to Information.....	73
4.4.3	Data Integration.....	74
4.4.4	Data Storage.....	75
4.4.5	Development and Group services integration.....	77
4.4.6	Data Restitution Architecture .....	78
4.4.7	Standardization, Industrialization & Reuse.....	78
4.4.8	MDM system Integration.....	79
5	Conclusion and Future Task .....	80
5.1	Conclusion .....	80
5.2	Future Task .....	80
	Glossary.....	81
	Bibliography .....	83
	Appendix – Product: Business Intelligence ITRefCard .....	85

## Table of Figures

Figure 1-1 Implantations of BNP Paribas All Over the World.....	10
Figure 1-2 BNP Paribas Integrated Business Model .....	11
Figure 1-3 The Six Layers of EAGLE .....	17
Figure 1-4 EAGLE Framework.....	18
Figure 1-5 Illustration of Software Integration .....	22
Figure 1-6 Generic Software Service Reference Model.....	26
Figure 1-7 Example of GSSRM.....	27
Figure 1-8 Reference Technical Architecture .....	28
Figure 2-1 BI in the context of a Métier’s Information System .....	34
Figure 2-2 The Business Intelligence Functional Architecture Defined by BICC.....	36
Figure 2-3 Business Intelligence Architecture .....	38
Figure 2-4 Topologies of Operational Data Store.....	41
Figure 3-1 Reference Technical Architecture .....	47
Figure 3-2 Application Software Domains.....	48
Figure 3-3 ITReferenceBook.....	50
Figure 3-4 Process of the Project.....	52
Figure 3-5 Roadmap of the Team Reference Technical Architecture.....	53
Figure 4-1 Mapping from BI Services to GSSRM .....	59
Figure 4-2 Business Intelligence Software Service Reference Model.....	60
Figure 4-3 Business Intelligence Software Service Reference Model (Modified).....	62
Figure 4-4 Business Intelligence Reference Architecture .....	65
Figure 4-5 Business Intelligence Reference Architecture (Modified) .....	66
Figure 4-6 Reference Architecture for BI system.....	66
Figure 4-7 Business Intelligence Patterns Implemented by ETL.....	69
Figure 4-8 Business Intelligence Patterns Implemented by ETL (Modified).....	70
Figure 4-9 Business Intelligence Patterns Implemented by Message .....	71
Figure 4-10 Business Intelligence Patterns Implemented by Message (Modified) ..	72

## 0 Introduction

This internship report states the task of figuring out the definition of Business Intelligence project in the context of Enterprise Architecture Group Language (EAGLE) within the BNP Paribas IT Group. At the same time, to propose guidelines and best practices, prepared for the Business Intelligence systems within the group, on the group level is also a main task of this article.

The first part of the document talks about the organization where the writing of this Business Intelligence ITRefCard takes place, as well as some basic and general information of BNP Paribas Group.

In the second part of this internship report, we introduce the context of this task. We give a short introduction of Enterprise Architecture Group Language - EAGLE of BNP Paribas. The main topics here are Software Integration Patterns and Generic Software Service Reference Model of the EAGLE. And also we talk about the team of Reference Technical Architecture who is responsible of the task of Business Intelligence ITRefCard.

The third part of the paper gives the basic knowledge of the Business Intelligence systems. We will talk about their functions and features. Meanwhile the main components of a Business Intelligence system are described in this paragraph.

In the forth part of the report, we will give a description of the project. We will present the plan and the client of this task. We state the fore work about Business Intelligence system before Reference Technical Architecture and BI ITRefCard within BNP Paribas. We will get a simple conclusion of their guidelines and definitions concerned.

The fifth part of the article presents the concrete work we did for the Business Intelligence ITRefCard, including group definition, Business Intelligence Software Service Reference Model, guidelines and best practices on the group level involving Business Intelligence systems.

The last part indicates the conclusion and the future task to be done to fulfill the task completely.

# 1 Organization and Context

## 1.1 Organization

### 1.1.1 BNP Paribas

As one of the European and global leaders in the area of banking and finance services, BNP Paribas has a huge international net covering 78 countries and more than 200,000 collaborators all over the world.

A group of **198 400** employees

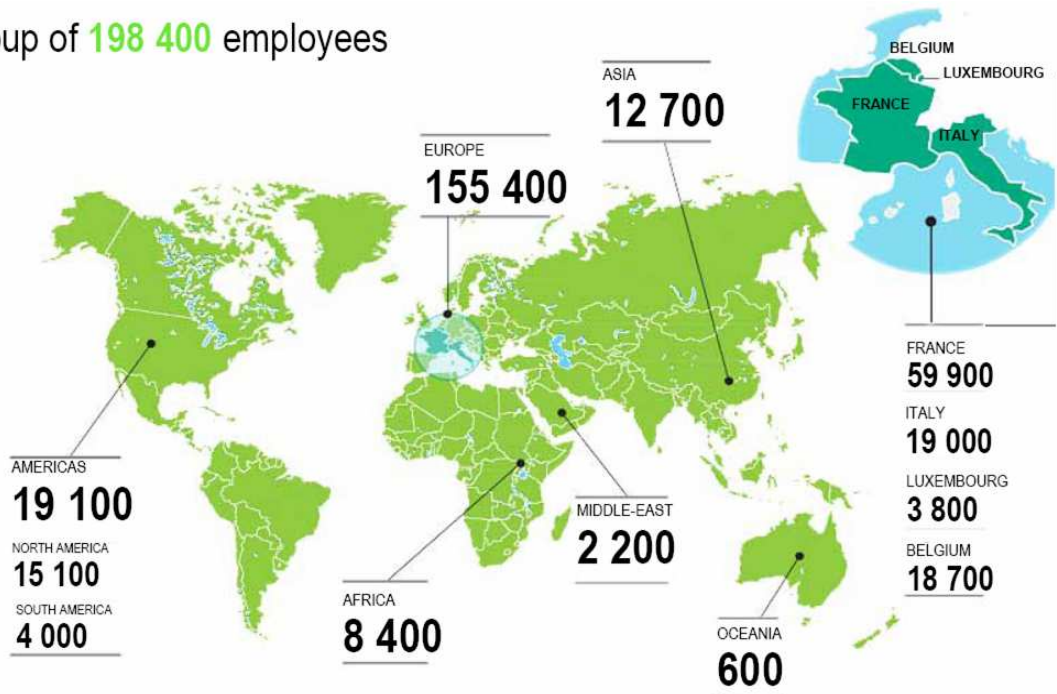


Figure 1-1 Implantations of BNP Paribas All Over the World

The new BNP Paribas Group was founded in 2000 with the merge of BNP with Paribas. The new group inherited two major traditional banks:

- BNP: the number one French bank dating back to 1848
- Paribas: a commercial and investment bank set up in 1872

After the combination in 2000, the group also acquired some other components:

- Banc West, from the USA, joined in the group in 2001
- TEB, from Turkey, joined in the group in 2005

- BNL (Banca Nazionale del Lavoro), from Italy, joined in the group in 2005
- BNP Paribas Fortis (originally Fortis Banque), from Belgium, joined in the group in 2009
- BGL BNP Paribas (Banque générale du Luxembourg), from Luxembourg, joined in the group in 2009

As a giant enterprise, BNP Paribas has an Integrated Business Model as following:

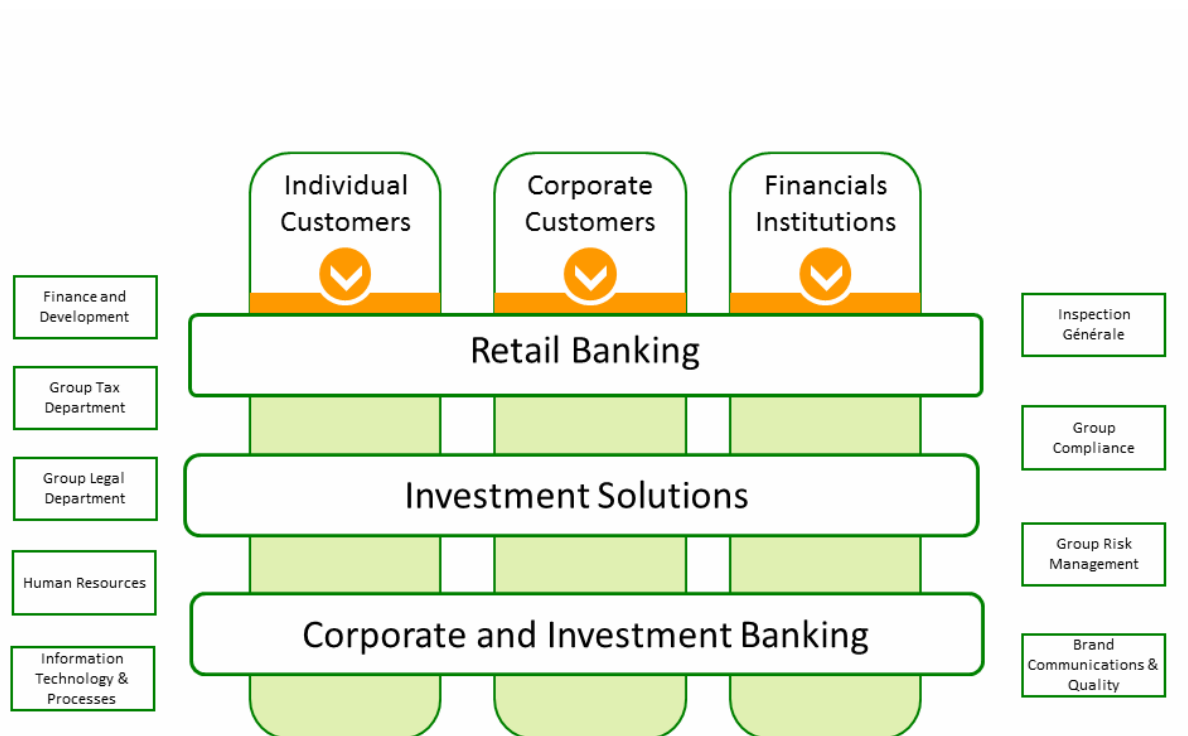


Figure 1-2 BNP Paribas Integrated Business Model

- The nine domains of activities:
  - Retail Banking

In 2011, BNP Paribas regrouped the web of agencies in Europe and in the world. In 2012, the organization of activities of Retail Banking evolves according to the particularly defined models of a set of Domestic Markets, an International Retail Banking entity and a Personal Banking entity.
  - Investments Solutions

BNP Paribas regrouped the activities of BNP Paribas related to the collecting, the management, the valorization, the protection and the administration of the deposit and the feature of clients.

Investment solution proposes a great number of products and services with strong added value, satisfying the requirements of particular enterprise and institutional investors.

- Corporate & Investment Banking

The principal objects of the teams of BNP Paribas Corporation & Investment Banking are developing and maintaining long-term relationships with their clients. They accompany their clients with their development and investment strategies, and also with their daily operations management. At the same time, they response to their requirement in the areas of finance, investment, financial advice and risk management thank to global solutions.

- The nine functions:

- Group Tax Department
- Group Legal Department
- Group Compliance
- Group Development & Finance
- Group Risk Management
- Inspection Générale
- Information Technology & Processes
- Brand, Communications & Quality
- Human Resources

### 1.1.2 Information Technology and Processes (Technologies et Processus)

The organization of Information Technology and Processes (ITP for short) is described as following:

- Group IT function
- Corporate Real Estate Services

Responsible for providing building and basic related services to Group entities; provide technical support for running the building.

- Procurement and supplier Relationship

Help to control external group expenditure and optimize relations between suppliers.

- Group Global Security

Align security of information, people and equipment with the operational risks of Core business Division, Business Lines and Functions.

- ITP Finance

Support function. Consolidate the ITP financial global vision.

- Human Resources and Communication

Support functions. In charge of recruitment, career management, job mobility, training, diversity and ITP communication.

- Compliance, PCOR & Business Continuity

Support functions. Oversee the monitoring and risk system and draw up applicable rules on ethics and information security.

- Corporate Insurance

Transversal program. Draw up and deploy Group strategy with regard to insurance of Group risk.

### 1.1.3 Group IT Function (Informatique & Technologies Groupe)

Group IT Function (ITG for short) provides a more global vision of IT, which firstly is a transversal one upon ITG and then upon specialist committee created according to a future federal governance model. It will, for each client, develop levers which offer it the assistance of improvement of its service quality and of its IT economic performance. ITP will also have a greater confidence and it is very important because we do not forget that IT can be dangerous. Actually, the benefit is to have a better IT with more security and a better controlled cost.

### 1.1.4 Infrastructure & Production Services (Infrastructure et Production Services)

The department of Infrastructure & Production Services (IPS for short) has the principal missions as following:

- Guarantee the quality of information productions operated for the businesses, operational units and BNP Paribas functions.
- Realize the technology evolutions according to the best practices of process.

- Carry out infrastructure structuring projects for the right functions of the Bank.
- Assure the coordination of centers of shared services operating in the domain of the informatics infrastructures.
- Continue deploying the actions necessary for controlling risks of informatics production presentations operated by ITG.
- Define and implement the governance upon the production activities within the Group.

#### 1.1.5 Reference Technical Architecture (Missions & Organization)

The team of Reference Technical Architecture (RTA for short) defines and manages the reference technical architecture and participant the alignment between the group production service package and the requirement of direction of information system within the Métier.

At the same time, the RTA supports the convergence and the mutualisation of the technical architectures, technical components and group production services.

Reference Technical Architecture animates the net of group technical architectures and brings out the evolution of the technical architecture within the group.

The team of Reference Technical Architecture is responsible for providing the technical architecture rules and architectures, as well as the solutions of referenced technical architectures.

The team of RTA defines the technical architectures at the IT Group level. (The functional entities and the operational entities can develop their own referenced technical architecture.)

#### 1.1.6 Business Intelligence Competency Center

The Business Intelligence Competency Center is an inner entity focus on the domain of Business Intelligence system and projects. The missions are:

- To coordinate Business Intelligence users and players within the group, with the aim of sharing experience and disseminating best practices.
- To coordinate the community of BI experts with a view of defining norms and standards for the product family.



- To provide advice and expertise on BI solutions for projects and Business Lines.
- To propose shared services specifically for BI solutions.

ITP is one of the contributors of the BICC.

## 1.2 Context

### 1.2.1 Enterprise Architecture Group Language

Enterprise Architecture Group Language, EAGLE for short, provides the common language for Enterprise Architecture within the BNP Paribas Group.

#### 1.2.1.1 EAGLE Scope

In that BNP Paribas Group is such a huge entity, it is not necessary to have an Enterprise Architecture scope which is able to cover the whole group. In fact, in the most of times, scopes are focused on given parts of the enterprise. In any case, the precise limits on which Enterprise Architecture is applied must be defined when Enterprise Architecture is launched.

In the context of BNP Paribas, Enterprise Architecture works can focus on:

- The whole BNP Paribas Group
- A pole
- A business
- A group function
- A line of businesses

Enterprise Architecture needs to manage the scale and complexity of an enterprise. EAGLE breaks down the enterprise into complementary projections called “layers”. Each layer covers specific concerns for different EA stakeholders.

Business & IT Managers



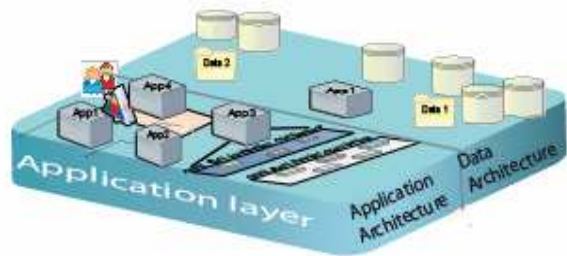
Business Representatives



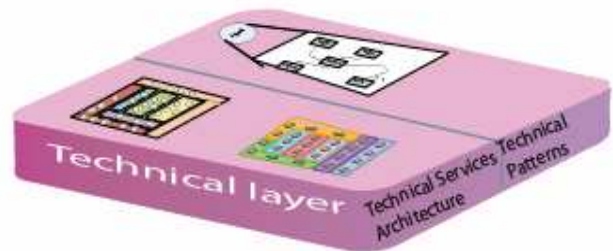
Enterprise Functional Architects



Project & Application Portfolio Managers



Enterprise Technical Architects



IT Development & Production Managers



Figure 1-3 The Six Layers of EAGLE

Here, an “Enterprise Architecture Stakeholder” means a major actor of information system evolution, whose point of view is considered as a priority for Enterprise Architecture. The list of Enterprise Architecture stakeholders provided here is not intended to be comprehensive and completed. For each layer, we just propose one stakeholder for the illustration purpose.

EAGLE is a way of providing an overall and integrated set of concepts and models organized into different layers that reflect the point of view of all the Enterprise Architecture stakeholders in such a manner that these different points of views are consistent and aligned.

These six layers complete each other to provide a completed and integrated vision of an enterprise architecture scope for all interested stakeholders.

### 1.2.1.2 EAGLE Framework

This is the EAGLE Framework of six layers:

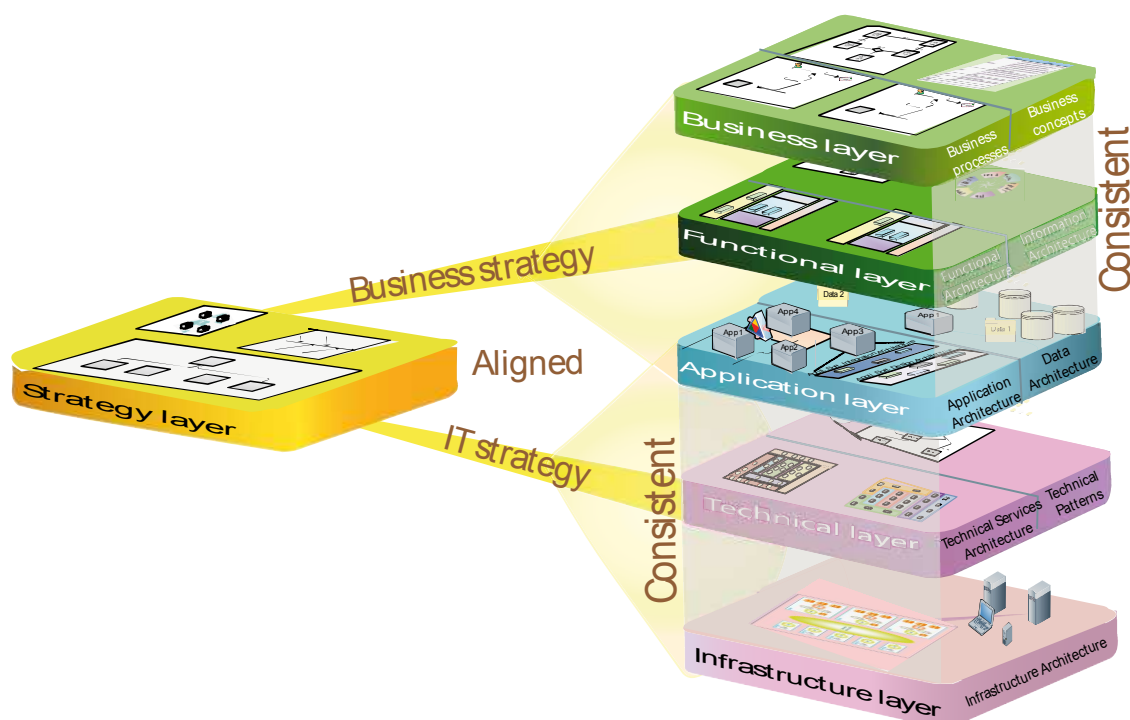


Figure 1-4 EAGLE Framework

Among the six layers, the strategy layer plays a specific role that is enlightened by the EAGLE Framework.

Business, functional, application, technical and infrastructure layers must follow the standard and requirements of business and IT strategy goals formalized by the strategy layer. Within each layer EA concepts must be identified and models realized with a particular concern for consistency with other layers.

Consistency between layers is a key element for the success of any Enterprise Architecture work.

The functional and technical layers play a central role in overall consistency, due to their higher level of stability: they provide reference architectures which change when strategy changes. These reference architectures allow making consistent decisions about information systems evolutions (on processes, applications, infrastructures), on a day to day basis, thanks to common stable references.

Each layer is connected by one or more of its components to the components of other layers.

Decisions made while building the architecture of one layer must take into account the decisions made within the other architecture layers.

- Strategy Layer

The strategy layer provides IT Group with the strategic vision of the business. It is focused on the description of strategic orientations formalized by the Group or organization targets. It provides orientations for all the other layers, which are necessary to take and back up the right structuring decisions.

The strategy layer analyzes the business and IT strategic orientations into goals and requirements. The strategy layer also provides models of the enterprise organization.

- Business Layer

The business layer provides an overall vision of the business processes of a Métier with a final customer orientation supported by the end to end scope promoted by value streams.

The business layer also provides a definition of business terminology, more or less formalized depending on whether or not business representatives are familiar with modeling.

Business layer documents the current and final states of business processes. These processes evolve step by step with organizational and IT projects, on the roadmap towards target.

- Functional Layer

The functional layer plays a pivot role between business and application layers, based on the central concept of Information System function for

managing indirect links. It also provides a reference Information System architecture made out of functional sets and blocks that group or separate IS functions based on criteria such as reusability, stability, differentiation or required standardization levels. These criteria must be chosen according to the business strategy and the business goals.

The functional layer allows translating the business strategic orientations and the goals/requirements that refine them into reference architectural elements. Some IT strategy elements might also need to be taken into consideration.

The functional layer documents IS functions that provide a common reference for business process activities and application software components. It allows sustaining and improving alignment between business and application layers, throughout all the evolutions they undergo as they follow their respective roadmaps.

The functional layer provides a list of the functions and information entities of an Information System. It provides also functional patterns for implementation into the business and application layers.

- Application Layer

The application layer consists of:

- Data Architecture

Data Architecture defines the list and scope of the data blocks existing in the current IT system, and required within the target IT system.

- Application Architecture

Application Architecture describes the breakdown of an Information Technology system into application software components, from high level components (modules) to low level components (application software components).

The application layer describes IT solutions both at IT system and Application software levels. Its intention is to promote convergence of solutions and alignment on business and IT goals.

Application target architecture derives from both functional and technical reference architectures and provides a meeting ground for functional and technical architects.

- Technical Layer

The technical layer plays a pivot role between application and infrastructure layers, based on the concept of services. It provides a translation of the IT strategy into reference architectures encompassing both application and infrastructure layers' elements.

The technical layer defines Reference Architectures via Rules and Patterns meant for consistent decisions between application software and infrastructures. It doesn't describe the technical components necessary to implement services and patterns.

- Infrastructure Layer

The infrastructure layer implements infrastructure reference architectures defined in the technical layer. The infrastructure layer defines IT infrastructure services offered to application software components.

The infrastructure layer documents the current and target states of hardware and software infrastructure components. These components evolve step by step together with business and infrastructure projects, on the roadmap towards target. It defines IT infrastructure services to be used by application software components of the application layer (or human end users) and by other infrastructure services.

## 1.2.2 Software Integration Patterns

As technology continues to evolve, it is important to choose an appropriate design for a distributed system. Decomposing these architectures in patterns and functionalities, architects will often have to choose between multiple architecture solutions. The Software Integration Patterns enforce use of common Enterprise Architecture vocabulary and share best practices for designing communication between applications. It deals with three domains of issues.

### 1.2.2.1 Software Integration

Software integration deals with interoperability and information exchange between software components.

When designing a distributed computing system, defining the way that software components communicate and interact is an important and structuring technical architecture activity ("software integration architecture"). These technical choices related to software integration are essential because they have a direct impact on the performance, the liability and the ability to evolve of the whole system.

Software integration architecture is focused on the description of the data flows and specific software components required for implementing the exchanges between software components.

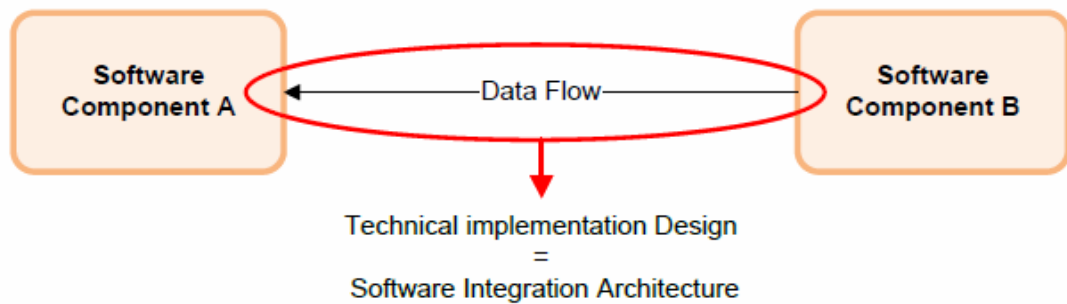


Figure 1-5 Illustration of Software Integration

### 1.2.2.2 Software Integration Patterns

The role of patterns is to capture and formally codify good designs and best experience-based practices in such a way that it is possible for designers to reuse them. By using these patterns, architects may decrease the perceived complexity of the solution. A pattern represents indeed as a simplified view of a given functionality of a system. As such, patterns are a good way to communicate between IT and business users as well as between two IT teams.

Provided as guidance to architects at this stage, this dictionary provides architectural and design patterns for software integration:

- Software Integration Architecture Patterns
 

An Architecture Pattern expresses a fundamental structural organization or schema for software systems that solves a specific, recurring, coarse-grained design problem.
- Software Integration Design Patterns
 

A Design Pattern describes a commonly recurring structure of collaborating Software Components that solves a design problem within a particular context.

### 1.2.2.3 Software Integration Building blocks

Software Integration Building Blocks are atomic software elements defining an aspect of architecture integration. Each Software Integration Building Block features intrinsic characteristics and can be seen as a fundamental element of integration architecture. Software Integration Architectural Patterns and Software Integration Design Patterns assemble Software Integration Building Blocks in order to realize particular use cases.



Software Integration Building Blocks form a toolkit for describing architectural and design patterns, or packaged integration solutions like, for instance, an Enterprise Service Bus (ESB).

Software Integration Building Blocks are divided into four independent categories which address distinct topics:

1. Topologies, describing the overall arrangement of Software Components interactions (e.g. point-to-point, via a broker...).
2. Communication vectors, describing the medium used to enable the communication (e.g. a file, a message...).
3. Communication sequences, describing which data flows in which order are needed to establish communication (e.g. synchronous or asynchronous request-response, one-way...).
4. Communication connectors, describing data sending and consumption strategies, e.g. how a consumer Software Component retrieves incoming data.

Each exchange between two Software Components can be described by at least one Software Integration Building Block per category. Sometimes two Software Integration Building Blocks from the same category are needed to describe an exchange. For instance, an exchange can contain both synchronous and asynchronous sequences.

During the design process, some Software Integration Building Blocks may appear more structuring than others. For instance, topology issue should be considered early as it has consequences in the decoupling level of the system.

### 1.2.3 Generic Software Service Reference Model

To describe the IT Systems they design, build, deploy and operate, IT Application Development activities and IT Service Management activities share technical concepts that form the basis of an IT language. This language must be used by all IT stakeholders in the Group working together at different grain of technical activities in order to ensure easy and fluent mutual comprehension of how IT Systems are organized, how they perform and evolve.

In this context, the GSSRM enables common identification methods by providing a reference classification to describe Software Services within software architectures at the conceptual level.

#### 1.2.3.1 Definition

Generic Software Service Reference Model (GSSRM) is a classification scheme of Software Services materialized by a normalized representation and color conventions. IT enables better software architectures through shared software Service Catalogs.

GSSRM is used to describe Software Architectures for decisions in patrimonies and projects, and to define reference architecture. It provides a map of Software Services, Blocks and Sets upon which Software Services, used by application or infrastructure software, may be dispatched according to their functional coverage.

Its main value is to provide a reference list to describe software architectures, and enable using rules and good practices shared by development and infrastructure teams.

This will ultimately lead to enhance or, in some cases, enable:

- Communication between stakeholders

GSSRM facilitates separation of concerns for software development, integration and deployment. At high level of description, it is used to show the Software Services organization and interactions to a non-technical audience. At a detailed level, it is used by developers to structure software development and packaging, and by production team members to analyze and optimize infrastructure services.

- Software modularity and reuse

Describing and categorizing Software Services helps to promote modularity and identify appropriate levels of Service granularity. Aggregating Software Services in standardized Sets facilitates the Software Services orientation in terms of architecture and/or development and enables to construct re-useable and loosely coupled software.

- Software Architecture rules definition, usage and measurability

The Set structure defined by the GSSRM helps to setup generic rules between Sets and with Infrastructure Services. Some of these rules can then be automatically checked and/or enforced by development tools and infrastructure solutions (e.g. code analyzers, deployment tools, firewalls, etc.).

It is also possible to define interaction rules between Blocks or rules to be used for packaging.

- Software selection among candidates

In case of multiple solutions opportunities in a make/buy/re-use choice of Software Service, the GSSRM summarizes the Software Service differences (in term of modularity and compliance with defined rules) in a graphical way and helps to choose the “right one”.

### 1.2.3.2 Classification

Software Services are grouped based, on common characteristics, into seven Software Service Sets.

The aggregation logic into Sets is based on the separation of:

- Access management, defined in a broad way, and distribution of external communications flows
- Human machine interactions
- Mechanisms related to the exposition and consumption of Software Services
- Functionalities that are specific to a Métier
- Data management: storage, optimization persistency and access
- Mechanisms related to software operations
- Other technical functionalities shareable between part of or all of the Service Sets.

### 1.2.3.3 Representation

The GSSRM graphical representation guides the possible choices for grouping the various Software Services, enabling to identify the services that can be factorized within the same application or infrastructure software at the conceptual level.

The GSSRM representation is a normalized map of the seven Software Service Sets, associated with color conventions and what must be used to position Software Service Blocks without overlaps.

***Important remark:*** The GSSRM normalized representation touch upon the well-known “N-Layer” deployment model. However, Software Service Sets must not be confused with deployment “layers”. A Set is a conceptual group of software services, which may or may not be deployed on one single physical “layer” (e.g. depending on security policies, performance or geographical constraints...).

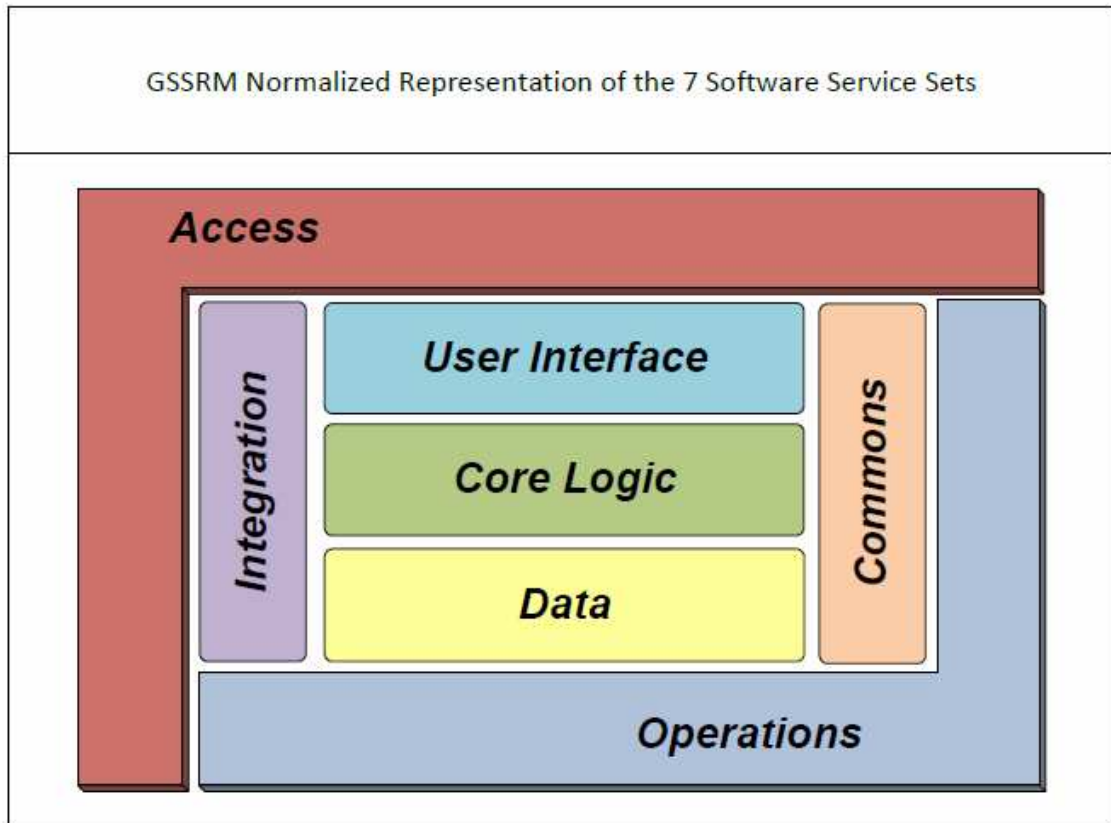


Figure 1-6 Generic Software Service Reference Model

The GSSRM provides a sample list of major Software Service Blocks and their definitions. This list is not exhaustive as the relevance of a particular Block may vary depending on the context. This is why GSSRM detailed representations (i.e. including Software Service Blocks and Software Services) are not normalized and must be defined at the Métier level.

The GSSRM detailed representations are necessary to identify redundant, single or missing software services, and to assess compliance with rules defined between Sets or between Sets and Infrastructure Services for packaging and mapping.

#### 1.2.3.4 GSSRM Software Service Sets

- Access Software Services
  - Handle all external interactions between the software, the users and other external software, and manage the access context associated with each user or external software service.
- User Interface Software Services

Provide interfaces for the end user of the software. They transform system information into output usable and readable by the end user.

- Integration Software Services
  - Enable interactions and interoperability with external software services and enhance internal functionalities.
- Core Logic Software Services
  - Provide core functionalities of the software (processing of data, transactions and business rules).
- Data Software Services
  - Provide data persistency, access and storage functionalities.
- Commons Software Services
  - Provide basic technical functionalities and/or capabilities used by others services.
- Operations Software Services
  - Provide functionalities enabling or facilitating the technical and functional administration of the software.

The following figure shows an example of a full SSRM recommended representation, showing Sets and Block:

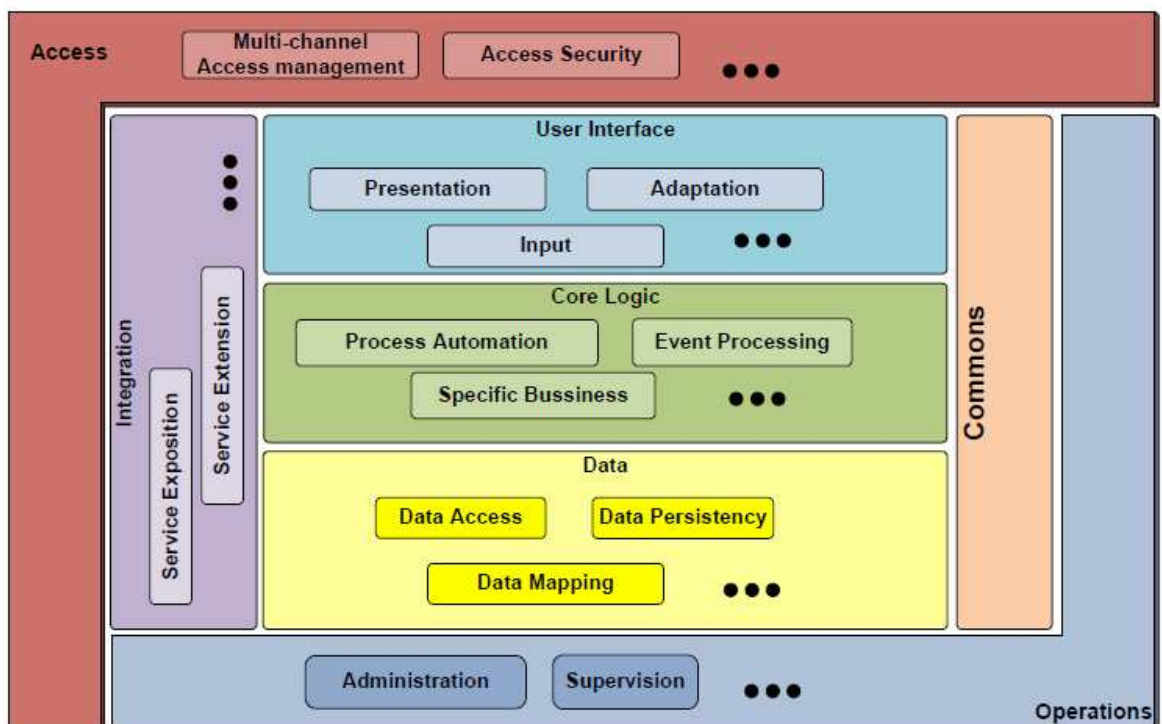


Figure 1-7 Example of GSSRM

### 1.3 Reference Technical Architecture

As we see in 1.1.5 Reference Technical Architecture, my internship took place in the team “Reference Technical Architecture”. We are responsible for a huge project to define group level “Reference Technical Architecture” which will last about 5 years.

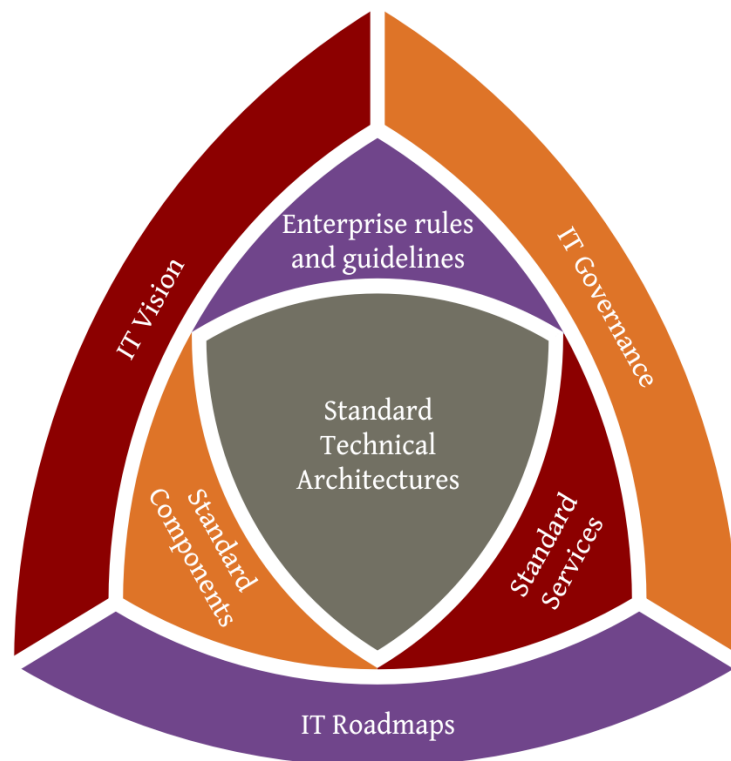


Figure 1-8 Reference Technical Architecture

Reference Technical Architecture encompasses four main areas:

- Enterprise rules and guidelines, intending to ensure interoperability, favor coherent and perennial choices, and maintain IT under control.
- Standard services, regrouping infrastructure services catalogue and shared applications software services catalogue.
- Standard components, today Group norms and standards catalogue.
- Standard technical architecture, detailed, operational, ready-to-use architectures, compliant with enterprise rules and using standard services and components.

The content of the task is supported by:

- An IT Vision approach aiming at defining an outlook for Group IT Services at 3 years, in order to guide the choices of rules and standards and define the direction of IT evolution roadmaps
- A Group governance for technical architecture, through dedicated governance committees at project, entities and Group level, with their indicators, control and waiver management mechanisms

- IT roadmaps managing tactical evolutions of standard IT components and services

### 1.3.1 IT Vision

In this area, we build a vision of the Group IT services.

The IT vision of Reference Technical Architecture provides a global and transversal view through all the IT projects by involving leaders from business and operation IT. IT vision focuses on a service oriented view at the center of technologies, organization and process. IT vision responds to reusability, interoperability and scalability goals. IT vision should be based on the business and group needs, best internal practices and market standards. The target of IT vision is providing a four–five years horizon shared and validated by group IT players. The horizon goes past existing constraints and allows a better control of costs and risks.

### 1.3.2 Enterprise Technical Rules & Guidelines

The Enterprise Technical Rules & Guidelines proposes interoperability, consistency and durability of BNPP’s Technical Architectures. They are references for both infrastructure architecture and application software architecture.

The contents of Rules & Guidelines include elements as following:

- Enterprise architecture definitions, principles, rules and guidelines
- Common high-level architecture models and patterns
- Strong focus on conceptual/logical architectures
- High-level decision criteria for appropriate selection of architecture patterns, infrastructure services, technology/component types... (recommended use cases)
- Elements of technical governance

Enterprise Technical Rules & Guidelines proposes benefits of enhancing visibility and adaptation of BNPP’s technical standards and accelerating architecture decisions & reviews.

### 1.3.3 Standard Technical Components

Today, the catalogue of norms & standards is mainly a product usage inventory inside the group. It is not focused on determining what the group standards are for a given need. It scares the top IT management by its volume. Standard Technical Components have a real value, especially in its ecosystem.

There is a strong will the management to quickly refocus the catalogue on a limited number of group standards:

- Driven by usage among the group, IT sourcing strategy, perennially and editor support, answer to needs, technical aspects, etc.
- In order to reduce costs, risks, and identify a limited number of main suppliers
- Other technologies will not be forbidden, but labeled “specific”

#### 1.3.4 Standard Technical Services

IT productions want to deliver more and more IT services to their customers, and less and less products and components. It is a strong drive to decouple Métier IT from IT production lifecycles. This will enable more added values from IT productions and more control of technological spawn that has to be supported. This is coherent with the “cloud” and “X as a service” approach.

The role of Reference Technical Architecture on these IT production service catalogue is still to be defined:

- Group standard components -> label services as “standard” or “specific”
- Group IT roadmaps -> evolution of services and their “standard” status
- Standard technical architectures -> can be delivered “as a service”
- Experience sharing (and maybe synergies) between Group IT productions
- Accompany the maturation of catalogues from components to services

#### 1.3.5 Standard Technical Architecture

The purpose of Standard Technical Architecture is to provide technical architecture accelerators for projects and ensure project compliance with enterprise rules. At the same times, it promotes use of standard services and components. It aims at being widely deployed and cheaper than specific architectures, for standard needs, and eases industrialization of IT operations.

The characteristics are as following:

- Answer to Métiers and Group needs
- Are based on Group and market best practices
- Are not mandatory
- Are aligned with/inherit from the IT vision (HITS)
- Are based on standard services and components
- Aim at being detailed/operational, covering all representation levels defined in Eagle (from “Conceptual” to “Deployed”)



### 1.3.6 IT Roadmaps

IT roadmap aims at tactically manage evolutions of standard IT components and services:

- identification of future standard technologies / components for the Group for a given need
- Identification of obsolete technologies/ components that need to be decommissioned

## 2 Business Intelligence

### 2.1 Introduction

With increasing amount of information systems and gradually accumulating enterprise data, enterprises need a more efficient and precise method to utilize the data and to discover useful information from the data. However, due to the great amount, the methodologies with which we deal with the regular relational database and structured data pattern are not same with what we will adopt here. At the same time, due to the fact that the data come from various types of source systems, how to unify the patterns and structures of data is also a big issue for IT department of every company.

The technology of Business Intelligence or the solution of Business Intelligence, with many existing products, is now a popular trend to solve the problem of huge amount of data and make advantage of them to produce strategy and take decisions.

#### 2.1.1 Historic

In 1958, the term Business Intelligence was used by Hans Peter Luhn, one of the IBM researchers. That was the first time we see this name in the world. He defined Intelligence as “the ability to apprehend the interrelationships of presented facts in such a way as to guide action towards a desired goal”.

Actually the Business Intelligence systems today come from the decision support systems which began in the 1960s. They were created to assist with decision making and planning. From that time on, the technologies like data warehouse and on-line analysis processing came into focus in the late 1980s.

In 1989, the Gartner Group proposed that “Business Intelligence” could be used to describe “concepts and methods to improve business decision making by using fact-based support systems”.

In 1990s, in the context of Gartner Group in 1996, Business Intelligence was defined as the application of a set of methodologies and technologies, among which data warehouse and data mining are the most famous.

#### 2.1.2 Definition

Nowadays, Business Intelligence covers a broad category of applications and technologies for gathering, providing access to, and analyzing data for the purpose of helping enterprise users to make better business decisions and to acquire a global view of the business activities.

Business Intelligence symbolizes a process including collecting data from multiple functions of an enterprise, formalizing and cleansing information,

consolidating and diffusing data according to the usage of each specific domain or subject. It consists of a series of interactive activities of exploring and analyzing structured domain-specific information to discern business trends or patterns, thereby deriving insights and drawing conclusions.

At the same time, Business Intelligence which usually includes a portfolio of decisional applications relying on a specific dedicated Information System segregated from operational information system, permits piloting the enterprise by providing a detailed or synthetic view, by facilitating the process of making decision and pilotage of activities, and by contributing the performance of commercial activities.

It is more than an application or a certain kind of software. Actually it is a solution of different software components combining, traversing, and excluding the possibility of data silos in order to making decisions from the point of view of the whole enterprise. It is an integrated solution which focuses on the commercial requirements and business needs. In the developing process, the understanding of business process and how to define business issues are always the key challenge for the team.

From the IT perspective, Business Intelligence (BI) is a broad category of software applications and technologies for gathering, storing, analyzing, and providing access to operational and historical data for the purpose of helping enterprise users to acquire a global view of the Métiers<sup>1</sup> activities and to support their business decision-making.

From the process perspective, BI symbolizes a process including collecting data from multiple functions of an enterprise, formalizing and cleansing information, consolidating and diffusing data according to the usage of each specific domain or subject.

From the end-user perspective, BI is a portfolio of decisional applications (query & reporting, dash boarding, analytical processing, discovering, data mining, score carding...) that enable access to a detailed or synthetic view on short or long term historical information about their business and its environment. It supports exploration and analysis of structured Métier-specific information, generated across different activities (products, clients, contracts, sales, etc.) and gathered from multiple operational applications, to discern business trends or patterns, thereby deriving insights and drawing conclusions.

From the IS perspective, the decisional functions of a Métier which relate to Business Intelligence are included in the information system of this Métier, and constitute a single BI system, subpart of the Métier's information system; there is no BI system which is not a part of an information system.

---

<sup>1</sup> "Métier" is the generic name (ref. EAGLE Glossary v4) indicating a pole, a Group function, a métier or a subsidiary (i.e. CIB, Group Human Resources, Banque de Bretagne, Fixed Income).

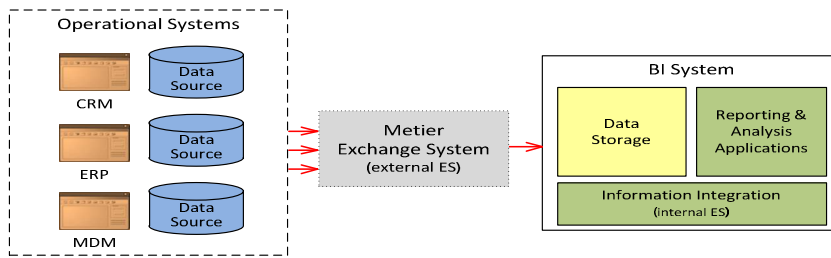


Figure 2-1 BI in the context of a Métier's Information System

This BI system, also called a “Decision Support System” (DSS), is a specific and dedicated information subsystem (highly analytic), segregated from operational systems (highly transactional). Data are collected from multiple operational applications (customer relationship, sales, marketing, referentials, etc.), and loaded in a central data storage<sup>2</sup> for additional operations (consolidation, normalization, historization, aggregation...) in order to satisfy BI purpose (reporting & analysis).

### 2.1.3 Usage

With Business Intelligence superior tools, now employers can also easily convert their business intelligence via the analytical intelligence to solve many business issues.

At the same time, an enterprise can identify its most profitable customers and underlying reasons for those customers' loyalty, as well as identify future customers with comparable if not greater potential.

And also, Business Intelligence analyses click-stream data to improve ecommerce strategies.

It can discover money-laundering criminal activities, detect and deter fraudulent behavior, such as from usage spikes when credit or telephone cards are stolen.

Meanwhile, Business Intelligence helps to reduce risk exposure through more accurate financial credit scoring of their customers.

A Business Intelligence system determines what combinations of products and service lines customers are likely to purchase and when and determines with attribution and churn analysis why customers leave for competitors and /or become the customers.

It also sets more profitable rates for insurance premiums.

The final object for a Business Intelligence system is to improve quality and availability of information and business data. More precisely, Business Intelligence can reveal the position of the company in the comparison with its competitors, changes in the customers' behavior and spending patterns, market future trend and the behaviors of competitors.

<sup>2</sup> In french, “Gisement de données” encompasses DWH and DM.

#### 2.1.4 Advantage

Business Intelligence, being an interactive method of construction and structured information analysis, provides many advantages and benefits to the enterprises which launch it.

It can pilot an enterprise by providing a synthesis view of the whole enterprise and the business as well as by assisting the process of decision making through consulting all the commercial activities.

With the implementation of Business Intelligence, a company or a department can reduce the possibility of “guessing” work within an organization in that they have more reliable data integration and logically organized data restitutions. It enhances communication among departments through coordinating activities and enables companies to response quickly to the outer environment. Business Intelligence can totally elevate the performance of an enterprise because it can improve the quality and efficiency of data.

Well-built Business Intelligence systems facilitate the capabilities to research on business information so as to coordinate management and to improve decision making from a point of view de enterprise level.

Meanwhile a vestment in Business Intelligence systems helps to realize and maximize the value of the historical data of the company. By improving the structuring the large volume data, a BI system intends to discover the regulars hidden in the information and reflex them to the decision makers. The regulars usually reveal the relationship among clients, products, sales, providers, employees and so on.

A Business Intelligence system helps to accumulate daily operational data in the routine operations and transactions. In this way, we can prevent enterprise’s ignorance and supposition upon the data.

## 2.2 Foundation of Business Intelligence System Architecture

In the definition and the work of Business Intelligence Competence Center before us, they defined this Information Management Framework which describes the necessary components for a Business Intelligence system.

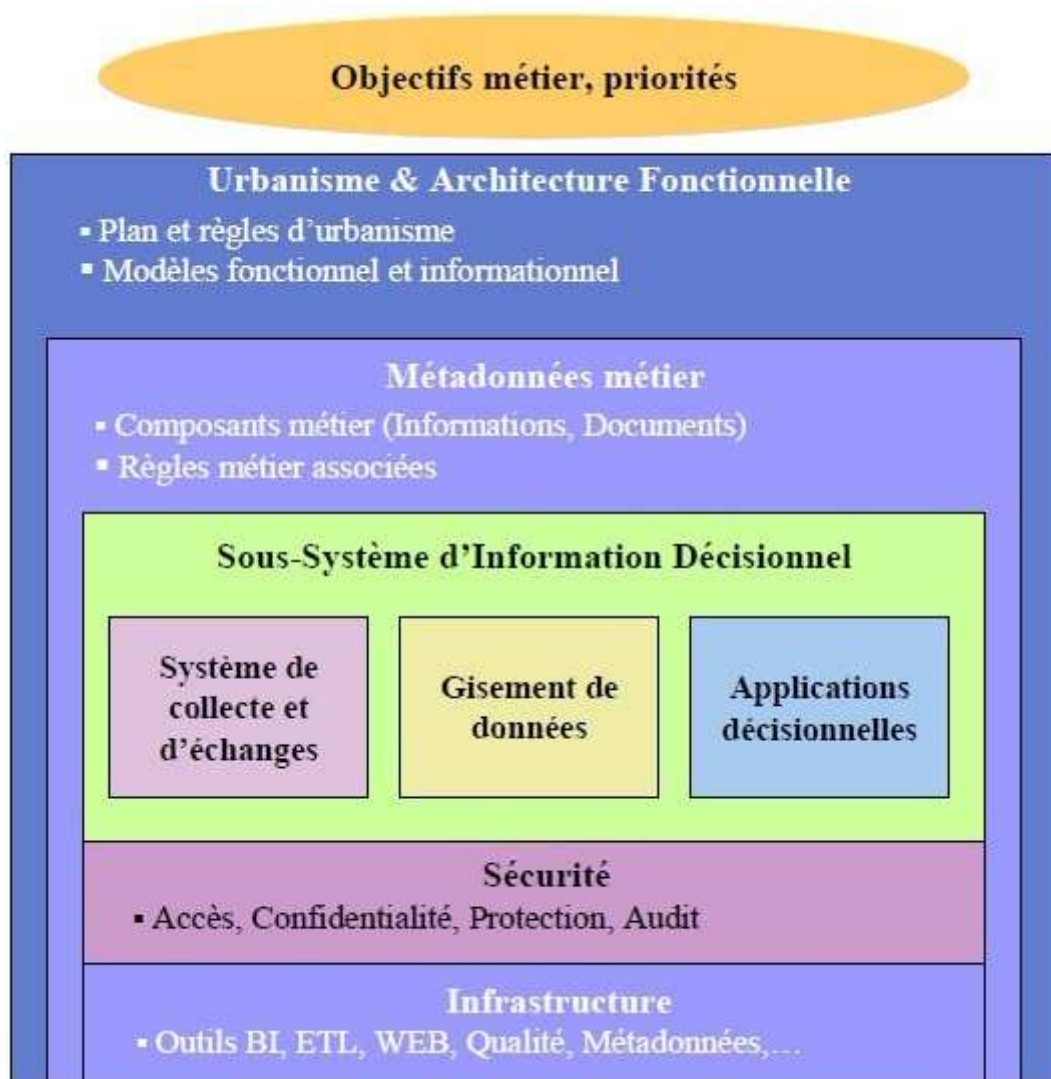


Figure 2-2 The Business Intelligence Functional Architecture Defined by BICC

The Business Intelligence system is defined by the urbanization map and the functional architecture.

- The urbanization map identifies and delimits the scope of s Business Intelligence system: identification of Métiers, objectives of a Métier Business Intelligence system, etc.
- The functional model specifies the functions assured by the Business Intelligence system.
- The informational model fixes the structure and the definition of the measures and indicators on the Group level. It is derived and detailed on the level of each Métier.

The Métier meta data defines the information and the restitution that the Business Intelligence system should deliver.

- They are defined at the same time with the users' requirements.

- The information, such as the description, the rules of calculation and transformation, the periodicity of carrying out, the depth of the history, are specified in this part.
- The meta data of restitution provide and describe the supports, the tables of borders and the analyze environment, such as list of the information, media for diffusion, security regulars and periodicity of updating.

According to the definition of BICC, a Business Intelligence system is a collection of decisional applications based on the data layer. It is designed for supporting different application topologies which could be an application available just in a certain area or department.

From the architecture of BICC, a Business Intelligence system consists of three parts:

- A data collection and exchange system. This permits transforming and integrating the real time data from source systems. This system can also provide the services for accessing real time data.
- A data layer. This layer permits stocking the collected data, and distributing information to the decisional applications in an organized and historical way.
- A group of applications. They reconstitute structured information and put forwards the information to a great number of users, who come from general direction, analysts, and leaders of departments..... in order to finish this target, we make use of different methods: information delivery, information analyze.....

At the same time, a Business Intelligence system rests on:

- Security mechanisms which permit authority upon users and safeguard the access to the data.
- A technical infrastructure deposed into two parts: the first part is contributed to the integration of data, the second part to the restitution of data.
- A meta data management system which helps the disposition of users in the process of development with the organization, the description and the relation of meta data of a Business Intelligence system. It can be on the user level, technical level or the operational level.

## 2.3 The Component of Business Intelligence System

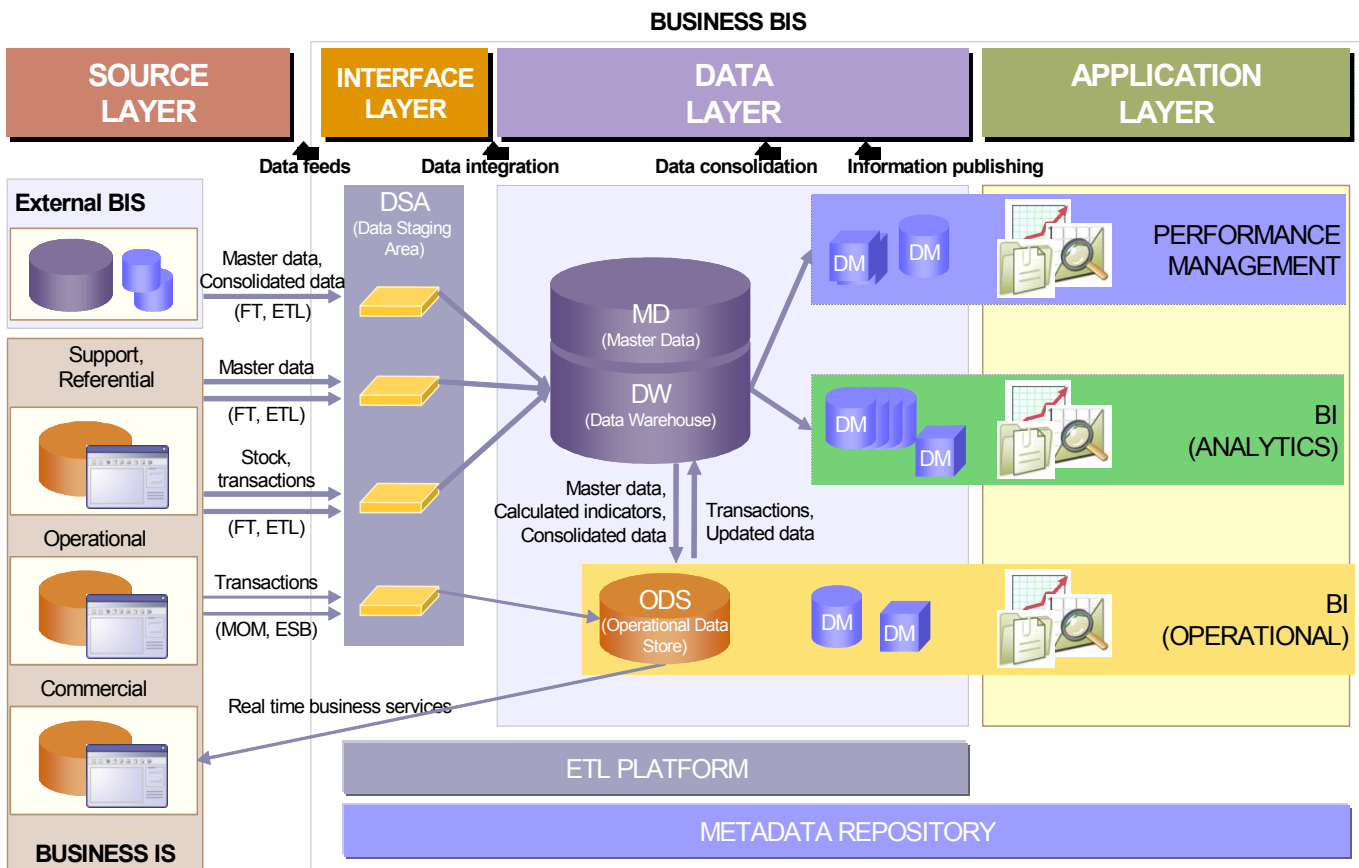


Figure 2-3 Business Intelligence Architecture

### 2.3.1 Collecting System

The collecting system satisfies the requirements of accumulating referential data, of business applications, of supporting applications or external systems.

It is responsible of the acquisition of data flow, usually which is important for; multiple heterogenic data sources and manages the reconciliation and the synchronization of these flow to produce target data.

It is responsible for the qualification of the data: controlling the data, management of the rejects and reprise of the treatments.

It is also at this level that the task of transformation and integration occur. (e.x. alignment of the attribute domains, creations of new indicators, keys management, passage of source data model in the form of target model.)

### 2.3.2 Integration System

The information integration platform is a technical component, usually initialized by a platform of ETL (Extraction, Transformation and Loading).

Within this platform, three types of ETL treatment are developed and executed.



- The integration treatment: it provides the components of data layer through the collecting system. The integration treatment is responsible for researching, controlling, accumulating and transforming source data.
- The management treatment: it is responsible of managing data archiving, the management of data techniques to measure the data quality.
- The distribution treatment: it is in charge of producing adapted data for the use of particular users' requirement and distributes them towards a component, for example, from data warehouse to data mart.

Obviously, the integration of Business Intelligence is not simply a process of accumulating data. Instead it is process of Extraction, Transformation and Loading, which are usually named process "ETL". ETL is an important unit of the construction of a Business Intelligence, which has a great impact upon the following parts.

#### 2.3.2.1 Extraction

The data of a data warehouse are subject – oriented and not all the data from source systems are valuable for it. Therefore, we should decide which data in the database is relative with strategy making before we import the data of source database into the BI system and data warehouse. The main process of extraction is as following:

- Identify the data of data source and their meaning.
- Extracting. Decide which tables and files to get access to and identify the attributes to extract.
- Make sure the frequency of extraction.
- Output. The address of output and the format.
- Solution for exceptions.

#### 2.3.2.2 Transaction

The data prepared for a Business Intelligence system usually come from various data sources. Different data may be developed and accumulated by different platforms and different applications, and also in diverse of data structures. The source data need to be transformed and even modified before being transferred into the data warehouse. The main task is to unify granularities and correct the inconsistency of the data.

#### 2.3.2.3 Cleansing

The quality of data is an important issue. Often, "dirty" or "polluted" data may lead to strategies of low reliability even wrong decisions. Meanwhile,

these bad data also bring out the problem of non-coherence among reports. So it is necessary to check data quality and eliminate errors.

#### 2.3.2.4 Loading

The data which have been transformed and cleaned are then loaded into data warehouse.

### 2.3.3 Data Staging Area

Data Staging Area is the interface between the source applications and the data layer. It entirely contributes to the acquisition, the control and the collected data. The flows of integration (between DSA and the data layer) are made used of by the data integration platform.

Data Staging Area permits decoupling the Business Intelligence system from source applications and referential sources. It also decharges the data layer from collect treatment. The data collection progresses nearly from their original point, which promotes the quality and facility of the treatment of collecting. It simplifies the source data treatment and proposes the autonomy to it. The collect programs are independent, redundancy and more easily joinable.

### 2.3.4 Data Layer

#### 2.3.4.1 Operational Data Store

Operational Data Store is in charge of integrating and carrying out the current data rapidly. It plays a role particularly important for the operational piloting applications and those supporting Métier applications. An Operational Data Store contributing to operational support can cohabite with another ODS within a Business Intelligence system. The Operational Data Store is not always necessary component and sometimes its absence in a Business Intelligence system can be frequent.

The data can be integrated and deployed more quickly than in a data warehouse. The Operational Data Store is dimensioned for responding requirements of mixed, operational and decisional consultation.

The topology of Operational Data Store:

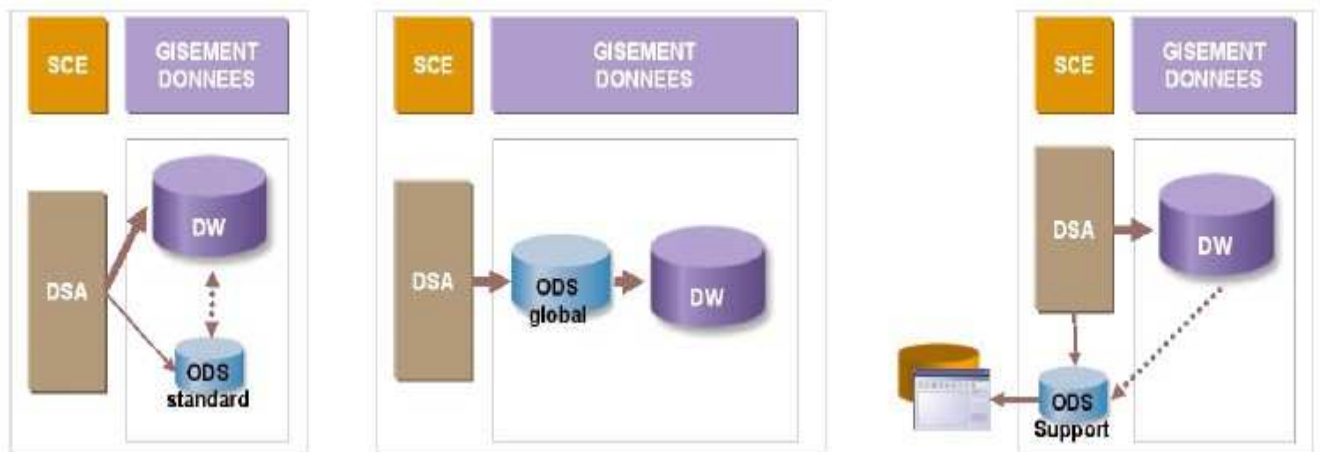


Figure 2-4 Topologies of Operational Data Store

We have three types of Operational Data Stores liked illustrated in the image above. Each one has a role according to the specific preoccupation.

The standard Operational Data Store and the global Operational Data Store do not cohabite simultaneously for the reason of optimization and avoiding data duplication. A choice must be done between these two sorts of Operational Data Stores.

The Operational Supporting ODS is usually implemented separately from a global ODS and a standard ODS for the reason of performance of access. It will be joined into an existing ODS.

- Standard Operational Data Store

In a standard configuration, the Operational Data Store is complementary for Data Warehouse. The data of Operational Data Store is current, namely valid and available in Métier no-historical and partial applications. (The domain of information are not covered exhaustively)

Therefore, it satisfies the need of consulting detailed data of a period of time going on.

The data are not susceptible to be modified before their integration into the Data Warehouse.

The issue indicators of the Métier applications work for the object of an iterative process of validation.

The data model is complementary with the model of the Data Warehouse.

- Global Operational Data Store

In the case of figuration, the Operational Data Store feeds the totality of the data of the Data Warehouse. So this plays a role of anti Data Warehouse.

This Operational Data Store has the same characteristics as the standard Operational Data Store, in the aspect of current data and elementary granularity. The difference rests in the completed coverage of the data.

The choice of a global Operational Data Store can be done to define a layer of statistics isolation (The layer materialized by a model of persisting data. It allows the editor of the Business Intelligence system not preoccupied the form of the collected data.), a "Software as Service" between the source data and the Data Warehouse when:

The data comes from plural heterogeneous information systems. It permits standardizing the source data in a great amount and therefore reduces dramatically the number of necessary integration treatments for feeding the Data Warehouse.

An entity of Métier level is responsible for the edition and for the distribution of a Business Intelligence system applicable on the level of branches or juridical entities. The Métier applications of these filialnesses are not identified.

The decision of creating an Operational Data Store should be naturally considered in that it can lead to double the model of data layer. Those will be translated with a proportional growth of the cost of conception and of the realization of the data and of the treatments.

- The Operational Data Store of operational support

The decisional information is more and more implicated in the execution of the operational process. The Operational Data Store is the component in charge of dialoguer directly with a Métier application.

Its usage is absolutely adapted in supporting the applications of market, applications of multi-canal or commercial management.

It permits, for example:

- Deliver, in a real time way, the evolution, the volume of effected transaction from the start of the day.
- Complete the knowledge of client.
- Calculate the score in real time.
- Calculate a commission.
- Deliver indicators of syntheses of a client's activity.

The most frequent method of an Operational Data Store of operational support is linked to the applications of client relation management.

- It permits Métier applications, through Métier services or services accessing to the data, finishing the following tasks:
- To access to decisional data.
- To effect treatments on a great amount of data very important for being treated by the technology of Métier applications.
- To treat data spread in the Métier applications but integrated in the Business Intelligences system.

#### 2.3.4.2 Data Warehouse

The Data Warehouse is the heart of the data layer: it centralizes the group of information of the Business Intelligence system. It is in this way a unique component in a BI system.

It assures the coherence, the right integration, the certification and the historical record of the data.

The major interest of the Data Warehouse is to provide coherent and reusable information to satisfy the need of the group of components and applications of the Business Intelligence system.

Offering structured and ready-to-use data, it prevents projects from systematically researching the source data from developing the treatments of control, from integration and from filling process.

#### 2.3.4.3 Data Mart

The Data Mart is a sub group of the data of the Data Warehouse adapted to a group of particular users of a Métier or to a particular application.

The data are selected and put into form to deliver concerned and understandable information to the final user. We can naturally find synthesis information there.

The decision of creating a data mart follows a certain number of criteria and demands:

- Carry out implement of a group of private data, for the confidential reason.
- Provide users with a functional and easily understandable model to measure that takes into consideration information and rules of specific management of the concerned group of users.

- Guarantee the performance of the application when a number of users use the system at the same time or the queries are completed through the creation of aggregates, or cubes.

It is noted that a Data Mart can be virtual if this does not surely go against the constraints of security and confidentiality. That is to say it does not generate the creation of a new instance of DBMS but makes use of, in the environment, the instance of DBMS of Data Warehouse. This is treated by the DBA.

The extraction of a sub set of the Data Warehouse data brings about more reactivity and simplicity for providing information to the final usage. The independence between the data model of Data Mart and that of Data Warehouse brings about more flexibility and change of development. It permits modifying model according to the wish of users without modifying the ETL treatment of collection neither modifying there referential data model.

It also permits guaranteeing more easily the performance of the consulted application. It always pays attention to the time of feeding data.

### 2.3.5 On – Line Analytical Processing (OLAP)

On – Line analytical Processing (OLAP) is a software technology of accessing and analyzing prepared for multiple dimension data and specified business issues. It permits the strategy people of an enterprise have a deep and profound view and research on the business data by providing a rapid, static and interactive view of the data. OLAP supports particularly complex analyzing operations, focusing on decision making of high level managers of enterprises.

#### 2.3.5.1 The characteristics of OLAP

- The OLAP technology is designed for analysis people and management people. It helps them have a direct view and impression upon the enterprise data and obtain what they need to support their decision making job.
- The operation of OLAP technology is read – only and intended to get access to huge volume of data. It does not have the average operations, such as adding, deleting and editing. What’s more, it is a searching process with the target of discovering useful information instead of simple query.
- The OLAP technology is subject – oriented and contributes to multiple – dimension data.

#### 2.3.5.2 Operations of OLAP

- Roll up and drill down

Roll up and drill down do not change the dimension of observing. Instead, the operations change the granularity of the data observed.

Roll up is the transferring from a micro view to a more macro view; drill down is the verse.

- Drill through

It is the process of transferring view from one model to another.

- Slice and dice

These two operations refer to choosing a certain dimension and analyzing data of other dimensions in the same type.

- Pivot

It refers to that analytical people change the direction of dimension in order to observe the relation of different sets and different dimensions.

### 2.3.5.3 Classification

- Relational OLAP

ROLAP refers to that fact that the data are stored in the traditional relational databases. Every model of ROLAP is based on some tables of this relational database. The typical models are star model and

- Multidimensional OLAP

MOLAP refers to the fact that the data on which it is based are stored in a multiple database.

- Hybrid OLAP

Due to the short comings of ROLAP and MOLAP, a mixed model is proposed as hybrid OLAP, which takes the advantages of the both.

### 2.3.6 Data Mining

Different from OLAP, Data Mining analyzes data by following the fixed regulars and existing data in the data warehouse. The goal is to recognize and discover hidden pattern and knowledge.

The main tasks of Data Mining:

- Association Analysis
- Clustering
- Classification
- Predication
- Time – Series Patter
- Deviation

### 2.3.7 Data Presentation

In order to demonstrate the data of the Business Intelligence system directly and coherently, we need to utilize tools to show them. The most usual tools are reports, dash - board, texts and ad – hoc query.



### 3 Presentation of the Task

#### 3.1 Introduction

As I said in 1.3, our team of Reference Technical Architecture is responsible for the project “Reference Technical Architecture”.

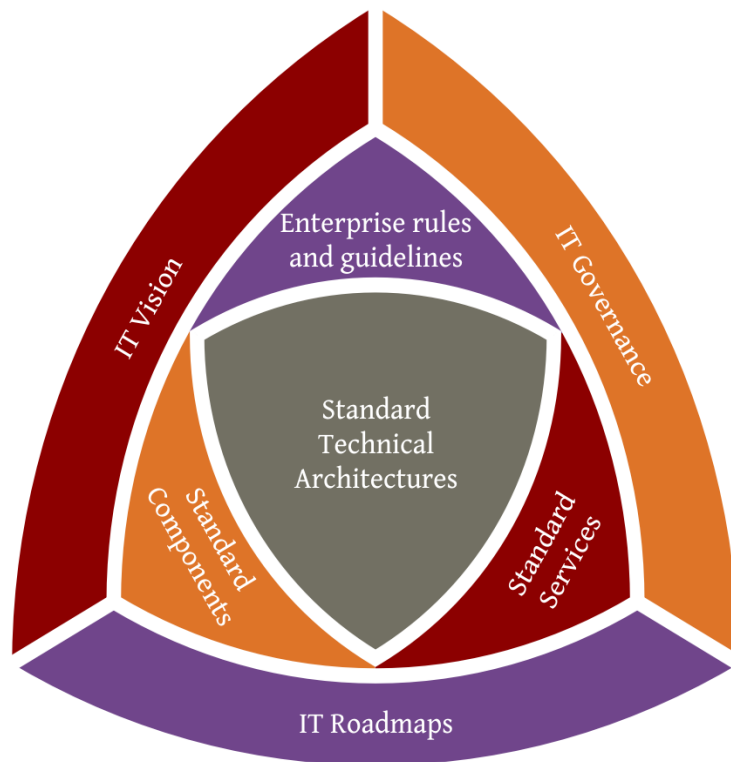


Figure 3-1 Reference Technical Architecture

We can see there is a small part called “Enterprise rules and guidelines”. Accordingly, in this domain we define rules and guidelines for IT projects on a group level.

In the context of Enterprise rules and guidelines, the team has three phases to define: Applications, Infrastructures and Transversal.

In the phase of Applications, we talk about rules and guidelines about software design and deployment or those specific to particular application software domains. Actually we have seven application domains to describe:

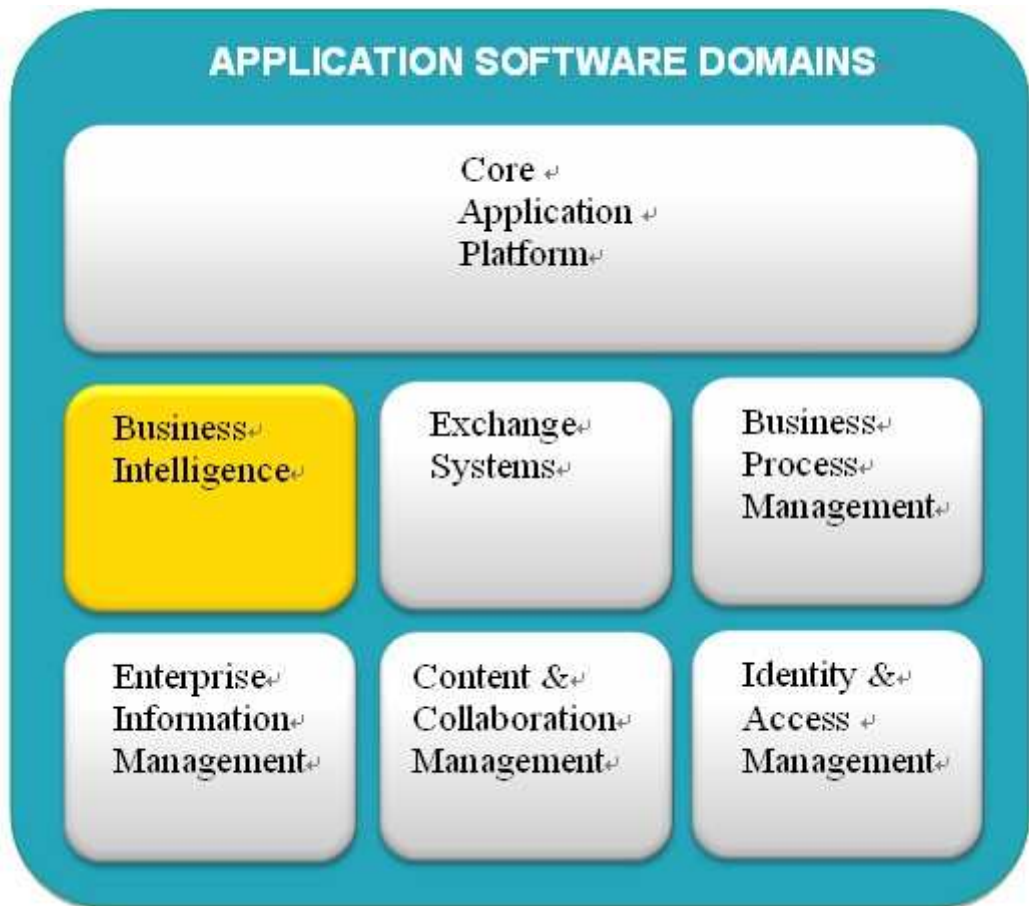


Figure 3-2 Application Software Domains

For me, my job according to the internship was to define the enterprise rules and guidelines in the domain of Business Intelligence.

I was expected to conclude the definition of Business Intelligence. And then I should talk about the benefits brought to the enterprise by Business Intelligence systems. I can not forget the paragraph of attention points in the design and development of a Business Intelligence system. The ecosystem and governance elements are also included in my work. Owing to EAGLE, the language of General Software Service Reference Model and platform Instantiations in the language of EAGLE Patterns will be designed and defined with great care and ordinance with the IT Group. Last but not least I must provide rules and best practices, which is the main target of the job, for Business Intelligence projects in BNP Paribas IT Group.

### 3.2 The Client

Actually, the team of Reference Technical Architecture does not have any external client. Our client exists in the internal of the group. Usually, they come from governance and project managers and architects.

For this task of ITRefCard document, the client is Technical Architecture Board, TAB for short. It is a Group level committee whose purpose is to pre-validate Technical Architecture standards and best practices. It gives directions on Technical Architectures at Group level, to meet the requirements of the Métier and of the Group, including non functional requirements (level of service, continuity, security, control, efficient use of IT budget, etc).

The TAB instructs and makes proposal regarding processes, methods and leadership of Technical Architecture within the Group (including technical innovation).

The TAB meeting takes place every two months. Sometimes the frequency can be adjusted. The agenda and working documents of the TAB are sent to the participants before the meeting so that all the members can prepare the decision and discussion. The agenda on the basis of current issues and requests are made by the Métiers, Filières and Poles. After each meeting, the organizer of the TAB collects the feedbacks from all the participants and make meeting minutes of the meeting.

The head of the Group Technical Architecture is the owner of the TAB. The members of the TAB include the main heads of the Technical Architecture of the Métiers, Poles, Retail Banking, Filières and IT operators. To facilitate the decision-making process by restricting the number of participants, one participant can represent several entities.

The TAB defines and pre-validates the evolution related to the Technical Architecture of EAGLE and to Technical Architecture standards and best practices before submitting them to the CNMA Committee for formal validation and deployment in the Métiers, Filières and Poles. It can be commissioned by the CNMA to define the Technical Architecture standards and best practices for the Group. The TAB proposes Technical Architecture standards and best practices to the CNMA.

At the same time, it can launch and drive actions related to the Community of Technical Architects: training, communications, capitalization, etc. Meanwhile, the TAB monitors an action portfolio, guides the activities, sets priorities, identifies and addresses interdependencies among issues.

### 3.3 The Requirement (The Stake)

For this specific project, some stakes must be cared about during the whole process of the project.

First of all, the data quality and governance are the premier topics for Business Intelligence. This is the first element that we will take into consideration during the process of the project. In this area, the task is to satisfy the increased needs for data consolidation and cross-silo reporting, which brought by increasing regulations.

The second issue to be talked about is performance. A Business Intelligence system in the Group must have a good performance which can make the best use of the rapidly growing volumes of data in the enterprise.

The next stake for the project is the problem of sourcing. In this area, the difficulties we must overcome include strong vendor consolidation, suites versus best-of-breed, niche solutions, open-source alternatives.

Last but not least, we will also take into account the latest innovations, which could add new interest and difficulties to Business Intelligence systems. For example, different kinds of appliances may bring out new challenges. The real-time BI, in-memory-BI, BI-aaS may appear in the future Business Intelligence system in BNP Paribas group.

### 3.4 The Product

The product for this job, however, is not an actual software product. Our responsibility is to make out internal documents to be put forward all around the IT Group. An ITReferenceBook is an assembly of ITRefCards published in a year.



Figure 3-3 ITReferenceBook

In this project, our product is ITRefCard. It is an internal document distributed all over the whole IT Group.

In an ITRefCard, we have enterprise architecture definitions, principles, rules and guidelines of a certain kind of software, or for a certain kind of application and system used within the IT Group. It provides common high-level architecture models and patterns described in the language of General Software Service Reference Model. It has a strong focus on conceptual and logical architectures. It proposes high-level decision criteria for appropriate selection of architecture patterns, infrastructure services, technology or component types. It gives advice of elements of technical governance.

With the help of ITRefCards, the IT Group will have a better control upon interoperability, consistency and durability of BNP Paribas's Technical Architectures. It is a reference for both infrastructure architecture and application software architecture. ITRefCard enhances visibility and adoption of BNP Paribas's technical standards (standards or shared services, standards

architectures, standards components). Thanks to it, project managers and technical architects can accelerate and ease architecture decisions and reviews.

For this specific Business Intelligence, it is one of the ITRefCards that RTA will finish in the following days. We have together seven application software domains mentioned above in which we will establish seven ITRefCards.

In this specific Business Intelligence ITRefCard, I will propose the definition of Business Intelligence in the context of IT Group of BNP Paribas. And then I talk about the benefits brought to the enterprise by Business Intelligence systems. I can not forget the paragraph of attention points in the design and development of a Business Intelligence system. The ecosystem and governance elements are also included in this BI ITRefCard document. Owing to EAGLE, the language of General Software Service Reference Model and platform Instantiations in the language of EAGLE Patterns will be designed and defined with great care and ordinance with the IT Group. The last part of BI ITRefCard is rules and best practices which is the main task of the document.

### 3.5 The Plan

As we said in the passage 4.3 The Product, the team of RTA will finish the documents of ITRefCards in the seven domains above. This year we will finish that of Business Intelligence and Core Application Platform.

For the task of the BI ITRefCard, we can see the general plan from the following BPMN illustration:

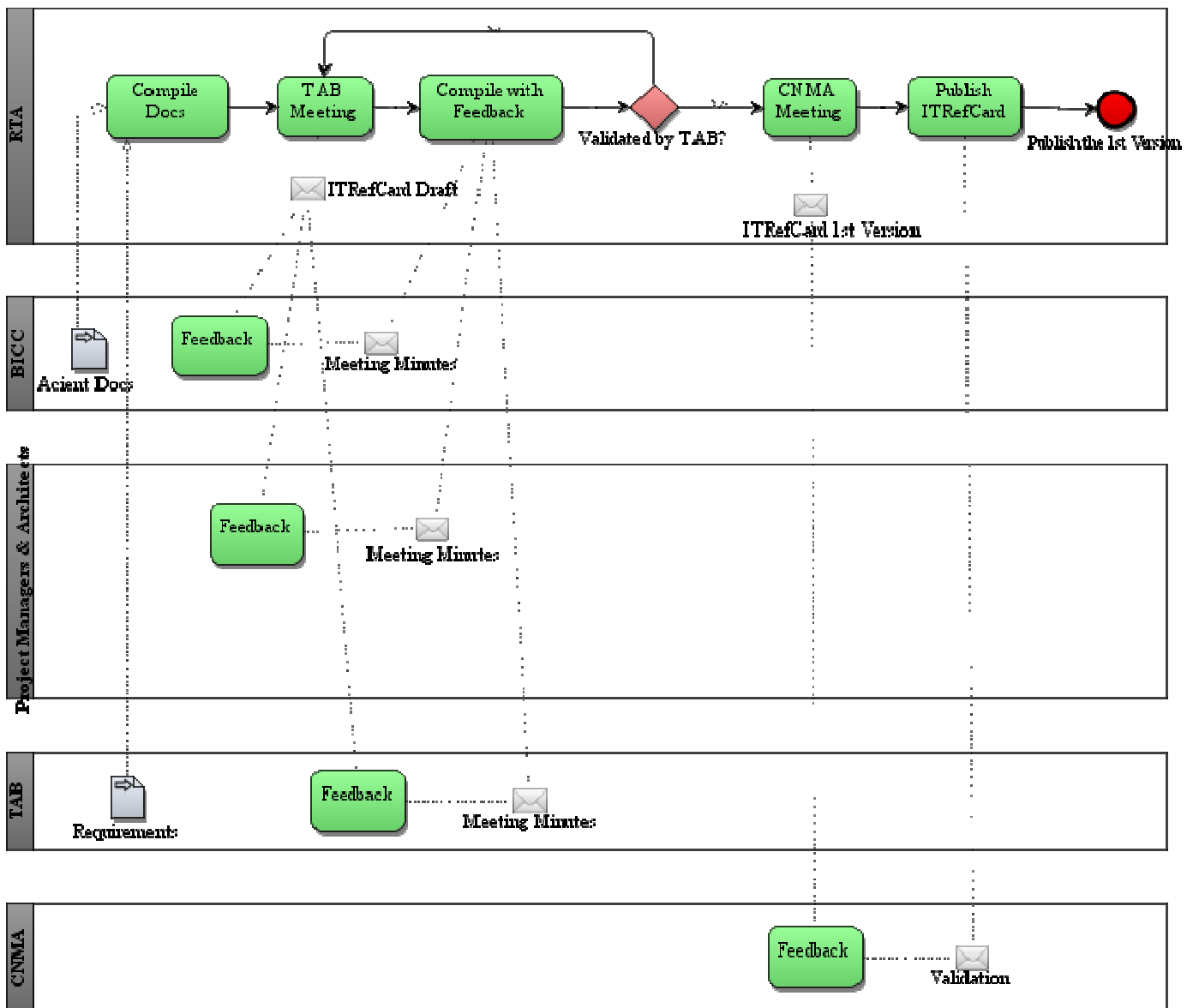


Figure 3-4 Process of the Project

From this diagram, we can notice the main participants of the project.

The main participants include the team of RTA, the BICC, the representatives of technique architects or project manager, the TAB and the CNMA.

We have introduced the function of the BICC in the paragraph of 1.1.6 Business Intelligence Competency Center.

The Technical Architecture Board (TAB) pre-validates the norms and best practices of technical architecture considering the requirement of CNMA. It orients the technical architecture on the level of the IT Group. At the same time, it introduces and makes the propositions about the processes, methods and strategies of animation of the technique architecture within the IT Group, as well as the innovation.

The Comité des Normes, Méthodes et Architecture (CNMA) validates the IT norms and Standards of the Group. It introduces the evolutions in order to put forwards them concerning the opinions of IT experts. It is responsible for framing

and monitoring the commanded work carried on by IT experts and finally validates the results.

The following image shows the roadmap of the team RTA with regarding the Business Intelligence ITRefCard.

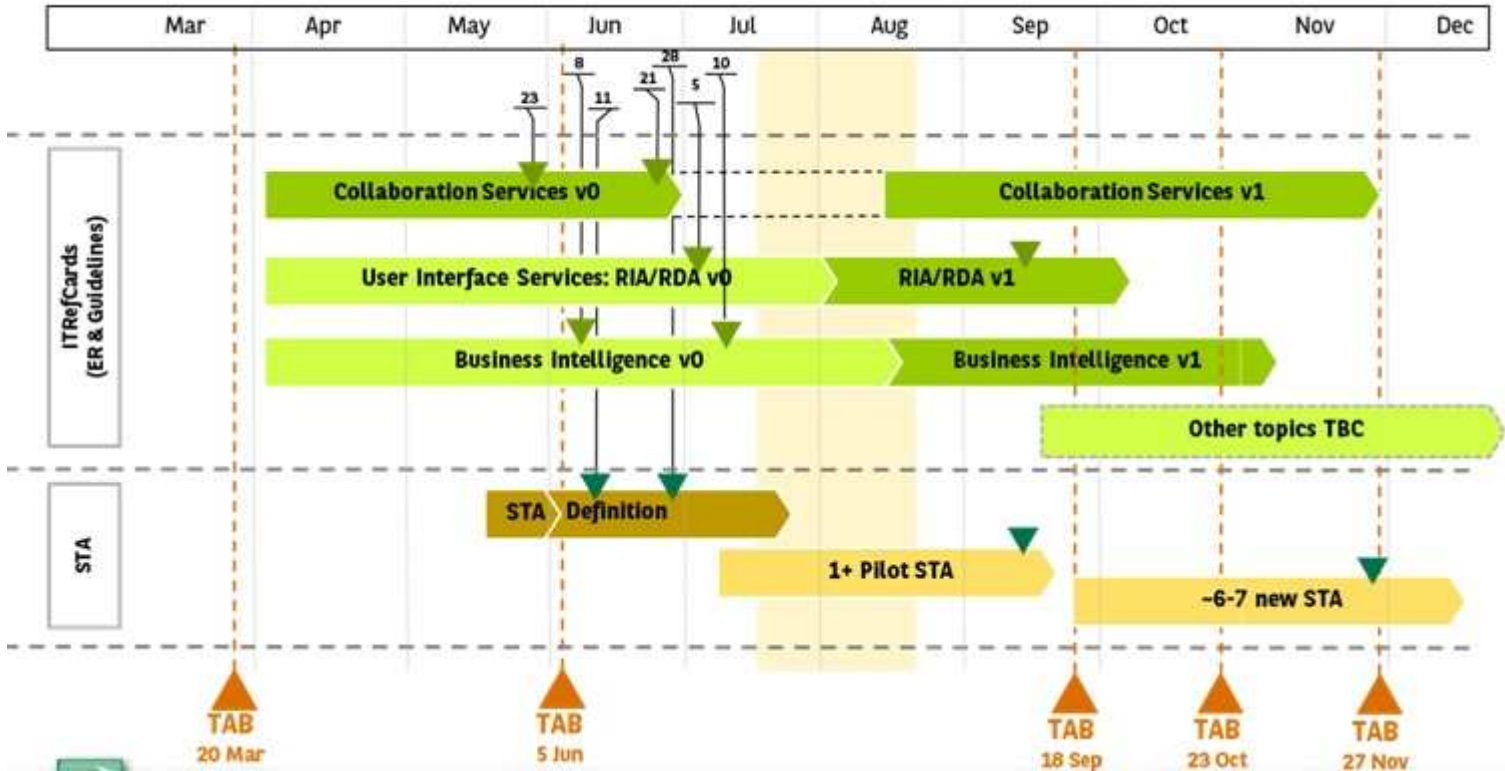


Figure 3-5 Roadmap of the Team Reference Technical Architecture

The whole task for the Business Intelligence ITRefCard should be proposed by Referential Technical Architecture and valid by the Technical Architecture Board in the middle of October 2012.

### 3.5.1 The Beginning of the Project

As for the ITRefCard of Business Intelligence, we had a kick-off meeting on the 20<sup>th</sup> March 2012, which meant the approval of this topic by BNP Paribas Technical Architecture Board and the official beginning of the project.

### 3.5.2 The Process of the Project

During April and May, I collected the contents of the existing documents in BNP Paribas IT Group. For example, the documents from Business Intelligence Competence Center and Master Data Management Community.

At the same time I attended the meeting of other organizations, such as Master Data Management Community and BICC.

In the month of May, I proposed the first version of Business Intelligence Software Reference Model, using EAGLE V4, as well as the first set of rules and guidelines based on already existing BNP Paribas documents. With the two parts we had the first version of BI ITRefCard.

#### 3.5.2.1 The First TAB Workshop

On 8th June 2012, I had the first workshop meeting with the Technical Architecture Board. At the meeting, I, with the guide of the Technical Architecture Board, did a selection of rules and guidelines extracted from existing documents. And then I got new suggestions for technical rules organized in different topics. At last I identified other existing technical documents.

#### 3.5.2.2 The Second TAB Workshop

On 10<sup>th</sup> July 2012, I had the second Workshop meeting with the Technical Architecture Board. At this meeting I proposed the second version of rules and guidelines concerning the feedbacks from the members of the Technical Architecture Board during the previous meeting.

After that I would expect the attendees of the meeting providing their feedbacks, comments, suggestions, and existing documents to us. The deadline was the 15<sup>th</sup> July.

RTA was expected to formalize and send the ITRefCard V0, the first draft to the public by the end of August and get review from the TAB for the V0 on the 18th September.

The Version 1 of the Business Intelligence ITRefCard was supposed to be provided at the beginning of October. And it will get its final validation by the Technical Architecture Board on the 23rd October.

### 3.6 The Situation before BI ITRefCard

Before BI - ITRefCard with general architecture prepared for BI projects, as well as rules and best practices, we only had the architecture and guidelines for each specific project. They rested on each single project and did not have much value of reutilization. It means that for each single specific BI project, BICC would help the development team members to figure out a series of rules and best practices. This would cost much time and money.



The work of BI – ITRefCard aims to put forward a technique architecture based on the GSSRM and EAGLE. The document will be a proper document for all the IT teams and departments within BNP Paribas Group instead of the project – oriented by BICC and project team. The final document will facilitate the development of the development of BI projects.

### 3.7 Working Environment

Our task of Business Intelligence ITRefCard is carried out under the routine work of the Referential Technical Architecture. We were working at the site of Montreuil of Ile-de-France, which is one of the most important sites of the IT Group of BNP Paribas.

## 4 Project: ITRefCard

With the work I did above and the communication with the architects and managers presenting in the TAB meeting, I finally got the conclusion for my task. In order to follow the structure of the ITRefCard, I needed to synthesis and analyze the content I received and re- describe them in a new method, namely in the language of EAGLE. And I must make sure that the new content should coordinate with the new background of the new organization of BNP Paribas IT Group.

### 4.1 The Introduction of Business Intelligence in BNP Paribas Group

For this part of the ITRefCard, I consulted the general introductions of BI systems as well as the internal documents concerning BI. After discussion with the two architects Mr. Avet and Mr. Mullet, I made the syntheses of all the contents and got the conclusion.

#### 4.1.1 Benefits

Business Intelligence is used by managers, analysts, and more broadly, by all employees with responsibilities and decisions to make.

This technology of reporting and analysis, which helps to turn into knowledge the gathered operational data, has become essential to monitor and to drive enterprise, to support decision-making process, to evaluate and to master risks, to control the team activities, to measure commercial performance, to understand market trends, to learn more about customers, etc.

- In Personal Finance, BI helps to reduce risk exposure through more accurate financial credit scoring of their customers.
- In Marketing, BI is used to define which combinations of products and service lines customers are likely to purchase and when, to analyze why customers leave for competitors and thus to model predictive churn and to calculate customer attrition, and to identify most profitable customers and underlying reasons for those customers' loyalty, as well as identify future customers with comparable if not greater potential.
- In Customer Relationship Management, BI assists to further improve customers' experience, with timely and appropriate response to their problems and priorities, and to analyze click-stream data and trends that enhance-commerce strategies.
- In Insurance, BI sets more profitable rates for premiums.

- In Risk management, BI can discover money-laundering criminal activities, detect and deter fraudulent behavior, such as from usage spikes when credit or telephone cards are stolen.
- In Corporate and Business Performance Management, BI is used to analyze activity and thus to determine cost, profitability and drivers, in order to realize periodic budgeting, improve processes and reduce costs across the organization.
- In business activities requiring simulation, BI enables end-users to access data and perform math-based simulations or what-if analyses and to study numerous scenarios to identify a solution or carry out advanced calculations such as cost allocation.

#### 4.1.2 Attention Points

Feedbacks from BNP Paribas BI experiences gain attention to the following points:

- **Organization:** The implementation of a BI system is not a conventional application-based IT project (such as an operational or transactional system), but rather a transversal program. After a new BI solution is deployed, new users are added to it, new indicators are asked for, more and more data are gathered... Therefore, as time goes by, the solution needs continuous investment, and work of BI does not finish like that of a common project.
- **Expertise:** A project team lacking BI application implementation experience will most likely fail to deliver desired results in the first iteration. Since most BI projects have aggressive timelines and short delivery cycles, an inexperienced and unskilled team is a risk that must be avoided.
- **Data quality:** BI services are very sensitive to information quality, higher / more than the operational systems. Organization must face this problem of quality of data available from sources, and operational data need to be checked and filtered/cleansed.
- **Silo:** Sharing data of BI in the organization demands that different units and entities of the organization should overcome the internal data silos that exist within the organization.
- **Reference data:** Master data are reference data shared by several operational systems, and become axes and dimensions of analysis in a decisional system. The requirement of Master Data Management (MDM) often appears with a new BI project, but the solution is necessary for all operational systems that share those reference data.
- **Meta-data Management (MM):** In order to implement a new BI service, it is necessary to define simply and accurately the questions that need to be

answered. Defining business terms and questions accurately, helps the BI analyst and developer to answer end-users' needs more easily.

- **Standardization:** With too many BI products to choose among, it is difficult to share experiences and to develop synergies. To avoid complex, multipurpose or multi - service projects and product solutions, an up-to-date catalog with standardized products makes IT actors more efficient, reduces risk, and accelerates BI implementation.
- **Data access authorization:** Access rights at data level must be defined and taken into account at the earliest steps of a BI project, since they may have an impact on all aspects of the solution then should be anticipated to avoid complexity and performance problems during implementation.

These attention points are useful when a Métier intends to acquire a new BI application, and lead to the following recommendations regarding governance, process and methodology:

- BI projects must be associated within a BI program.
- Only one transverse BI program must be considered per Métier.
- BI program must have identified sponsor and management.
- BI and MDM must remain two distinct systems even if they are under the responsibility of the same governance.

#### 4.1.3 BNPParibas Context and Strategy

To continuously enhance BI systems, BNPParibas has to match BI technologies with decisional, steering and analytical usages of Métiers, promoting proven architectural solutions able to smoothly integrate new technologies and satisfy future user's needs.

In this context, identified stakes for BI in the BNPParibas IT Strategy include:

- **Data quality & governance:** increasing regulations and steering requirements mean increased need for data consolidation and cross-silo analyses.
- **Performance improvement:** rapidly growing volumes of data and complexity of queries.
- **Rationalization of BI systems and associated infrastructures.**
- **Sourcing, with strong vendor consolidation:** suites versus best-of-breed, niche solutions, open-source alternatives.
- **Innovation:** predictive modeling, real-time decision, in-memory discovery, BI as a Service, DB appliances, convergence of operational and decisional data architectures.

For that, BI technical architecture rules and guidelines aim to favor consistent and perennial choices in project, and to maintain this technology under control, in conformance with those stakes.

Emerging technologies like real-time decision, operational intelligence, big data, and unstructured data support are not yet included in the current perimeter of the BI rules and guidelines. However, they are already identified and addressed by BNP Paribas through other Reference Technical Architecture processes (HITS, IT roadmaps...).

## 4.2 BI - SSRM

After consulting the basic architecture of Business Intelligence systems, I tried to classify the services involved in a BI system. And then I mapped these services on the modules of Software Service Reference Model of EAGLE.

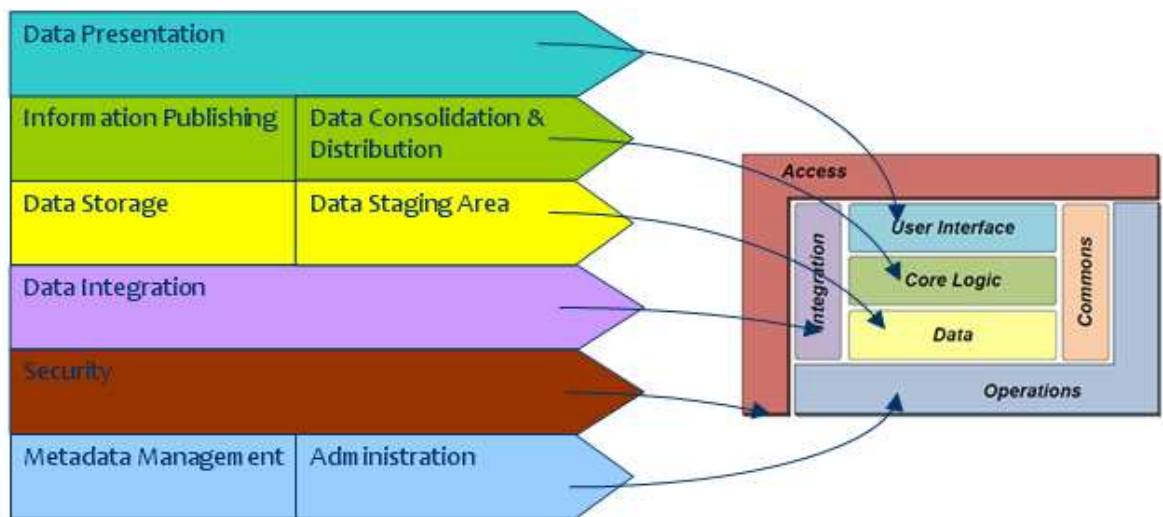


Figure 4-1 Mapping from BI Services to GSSRM

### 4.2.1 Business Intelligence Software Service Reference Model

Concerning of the definition and description of GSSRM of EAGLE, we try to classify all the services into seven sets:

- Access Services
- Operation Services
- Common Services
- Integration Services
- User Interface Services
- Core Logic Services
- Data Services

Accordingly, we give the definition of seven sets of services specifically designed for Business Intelligence systems of the BNP Paribas Group.

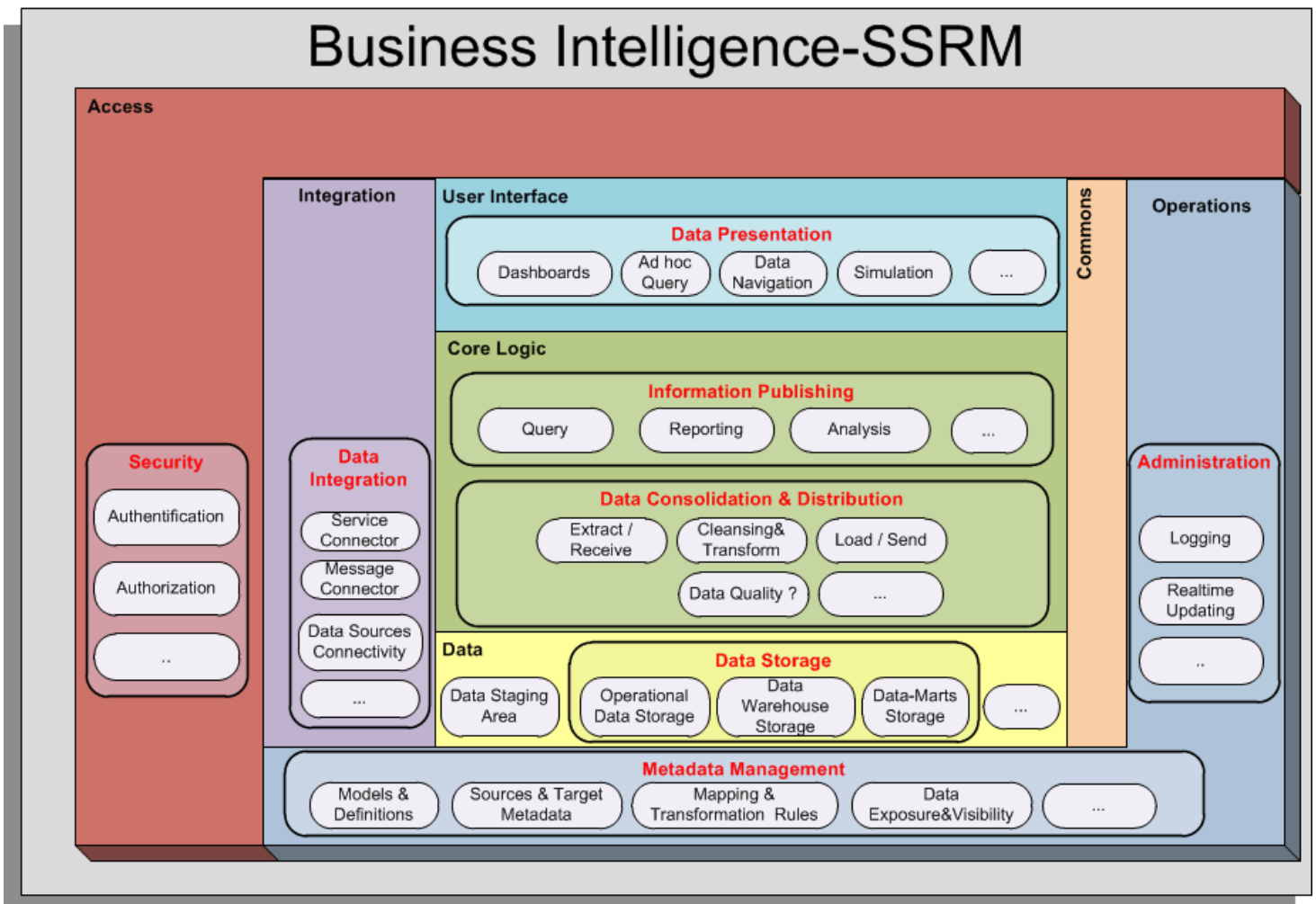


Figure 4-2 Business Intelligence Software Service Reference Model

- User Interface - Data Presentation:** This block of services contains the user interfaces which permit to exploit information of BI in diverse forms. It is the most visible part of a BI system and provides access to the system. This block of services contains the functions to display the data in various ways. It provides straight and direct way for users to understand Métier data and to figure out the condition of business execution.
- Core Logic - Information Publishing:** This block of services contains functions to prepare data and reports for specific applications with the help of analyzing and reporting tools. They implement the interfaces access to information contained in the Data Storage Services.
- Core Logic - Data Consolidation & Distribution:** This block of services contains functions to extract data from data sources and transforms raw data into the organized and reconciled form coherent with the data warehouse where the data will be loaded and consolidated. After that, it distributes data into different data marts corresponding to particular schema and business. It implements the information of a Métier or a group

of particular users. Moreover, it prevents direct accesses to the data warehouse from direct analysis.

- **Data – Data Storage:** This block of services proposes a unique and global data model prepared for the group of the components, which takes into consideration different life circles and information restitution. Operational Data Storage contains current and integrated data consulted for reporting and analysis near the operation, or the operational support for the Métier applications. Data Warehouse contains the historical, integrated and neutral data for the strategic requirements of reporting and analysis prepared for a group of Métier users. Data Mart contains historical and integrated data for the requirements of reporting and analysis specific for users of a certain Métier.
- **Data – Data Staging Area:** This block of services implements part of the services of data collection and exchange. The source data are treated here for further use of Data Storage Services.
- **Integration – Data Integration:** This block of services satisfies the requirements of collection of referential data from Métier applications, supporting applications and external applications. It is responsible for the acquisition of data flow of multiple sources and manages the reconciliation and the synchronization of these flows in order to produce target data.
- **Access - Security:** this block of services refers to all the measures that are taken to protect access to services from the people outside Métier. At the same time, it makes sure that those who want to get access to certain services do have the right authorization and right identification.
- **Operation – Metadata Management:** This block of services contains the functions to assure interoperability among all the information sources and inter data storage.
- **Operations - Administration:** this block of services offers the capabilities to maintain and operate the BI (starting/halting, modifying exec. parameters...) and its supported exchanges (subscription management...)

After the first version of the BI – SSRM, we did some modification before the first validation. And the BI – SSRM was edited as the following:

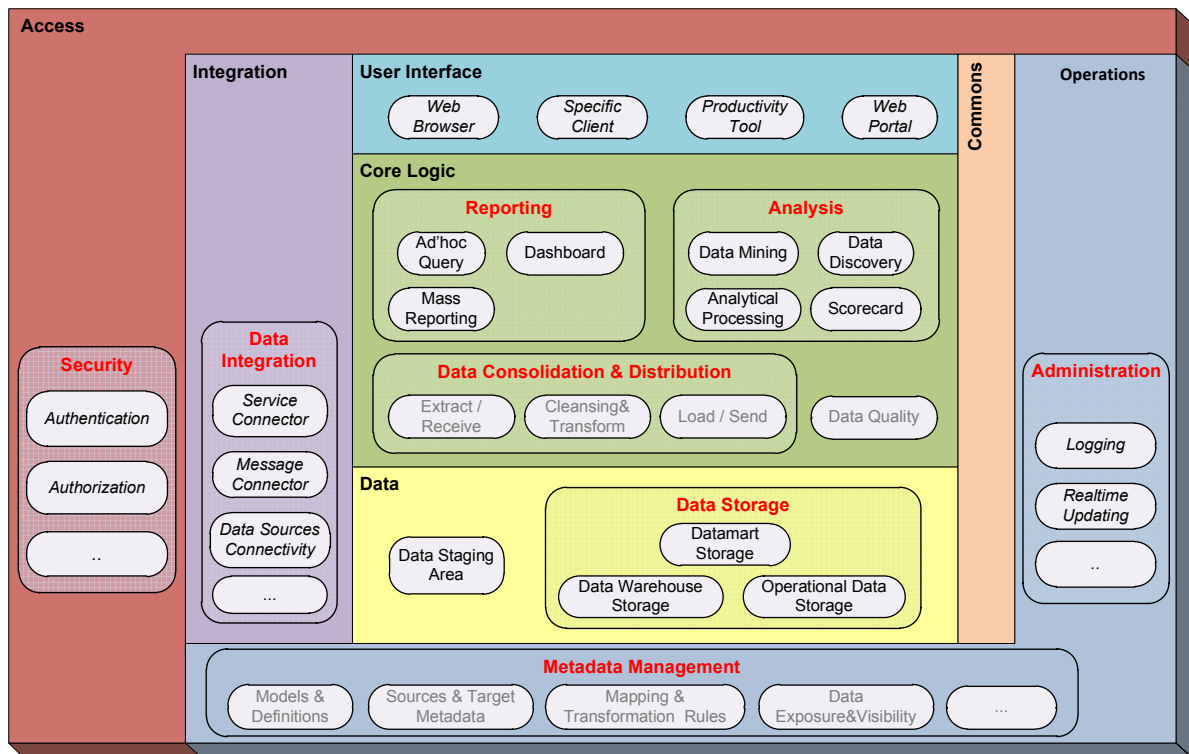


Figure 4-3 Business Intelligence Software Service Reference Model (Modified)

And we have more precise definitions for the block of Reporting, Analysis and Data Storage Services.

	Service	Definition
<b>Reporting</b>	<b>Mass Reporting</b>	<p>Routine creation of reports, with strong capacities of scheduling and processing.</p> <ul style="list-style-type: none"> <li>• Static reporting: Pre-defined and static reports, published to a large number of users.</li> <li>• Interactive reporting: Formatted and interactive reports, with capabilities of high quality generation (pixel-perfect) and user exploration (click to details).</li> </ul> <p>Moreover, the reports support a broad variety of styles for information delivery.</p>
	<b>Ad hoc Query</b>	<p>Allows end-users to create their own data-focused requests and presentation-focused reports, without the help of IT people. In particular, a semantic layer helps them to access valuable information. The service provides to end-users the capabilities to manage and reuse the requests.</p>



	<b>Dashboard</b>	Provides to end-users a quick and easy access to an overview of current business activities, with a time overview (trends of indicators), using a graphical presentation (charts, graphs, gauges and maps). Includes the possibility for end-users (at-a-glance) to structure and publish (in HTML) intuitive reports.
<b>Analysis</b>	<b>Analytical Processing</b>	Ability to exploit multidimensional databases and carrying out advanced online analyses of historical data (OLAP): slice & dice, seed-planting drill through, seed-planting drill anywhere, asymmetric report.
	<b>Data mining</b>	Advanced mathematical / statistical techniques used for pattern discovery (profiling, scoring) and predictive modeling/analysis.
	<b>Scorecard</b>	Ability to manage strategic goals, including the ability to link the objectives with performance indicators (KPI), causality models, initiatives. Supports several methods such as BSC or Six-Sigma.
	<b>Data Discovery</b>	Allows end-users to search, extract, consolidate, and visualize various data sources, in order to investigate data in real-time and find nuances, trends, and outliers “in a flash”.
	<b>Data Staging Area</b>	Collect and temporarily store operational and reference data for further use and integration by the permanent Data Storage services. For timing reasons (all data can't be gathered at the same time) and various levels of data quality (heterogeneous data sources), DSA is the place for synchronization, consolidation, reconciliation, cleansing, standardization, augmentation and enrichment operations, making the data ready for loading into the Data Storage.
<b>Data Storage</b>	<b>Operational Data Storage</b>	Contains the current value of some operational data, captured from multiple sources in near real-time, and used in very short-term decision making, or operational support of the Métier applications. Its data model is very similar to operational data sources.
	<b>Data Warehouse Storage</b>	Contains in a consistent whole, all the current and historical data periodically gathered from the operational and referential systems, and required in support of the Métier's decision-making processes. It is a single point repository of data for a combination of various Datamarts, and its data model is generic for all BI usages.

	<b>Datamart Storage</b>	<p>Contains subset of data from the warehouse for the specific requirements of reporting and/or analysis of certain Métier usages. Its data model is specifically aligned and optimized for the decisional applications to support.</p> <p>DM can be implemented in classical relational database (using "star" and "snowflake" schemas), but in order to be better dedicated to complex ad hoc queries and highly analytical processing, it can use specific technologies such as multi-dimensional database (hyper Cube, OLAP cube) and hybrid database (column-based, in-memory...).</p>
--	-------------------------	---

#### 4.2.2 Reference Architecture

In each ITRefCard, we design a Reference Architecture for the systems to be described and controlled by this ITRefCard. In this Reference Architecture, we try to accomplish a mapping to match the services defined in the model of GSSRM and services to be developed in the future work.

In this part, we demonstrate the main communication and data flows between the different services. The security Services and the Operation Services support the whole system as a general platform. The security Services serve the all process of data processing and decision making. The Operation Services manipulate the management of meta data of the whole system during the whole process of the data management.

The services defined in the GSSRM are mapped into different areas where we will consider developing concerning components or using existing software suits to build the final Business Intelligence system. The red arrows in the picture point the direction of data which go from the outside systems and data sources. And then the data get through the each part of system. The data come into the system through integration services and then they are extracted, transacted, cleaned and loaded. At last they are demonstrated by the User Interface Services.

The first vision of Reference Architecture image is as the following:

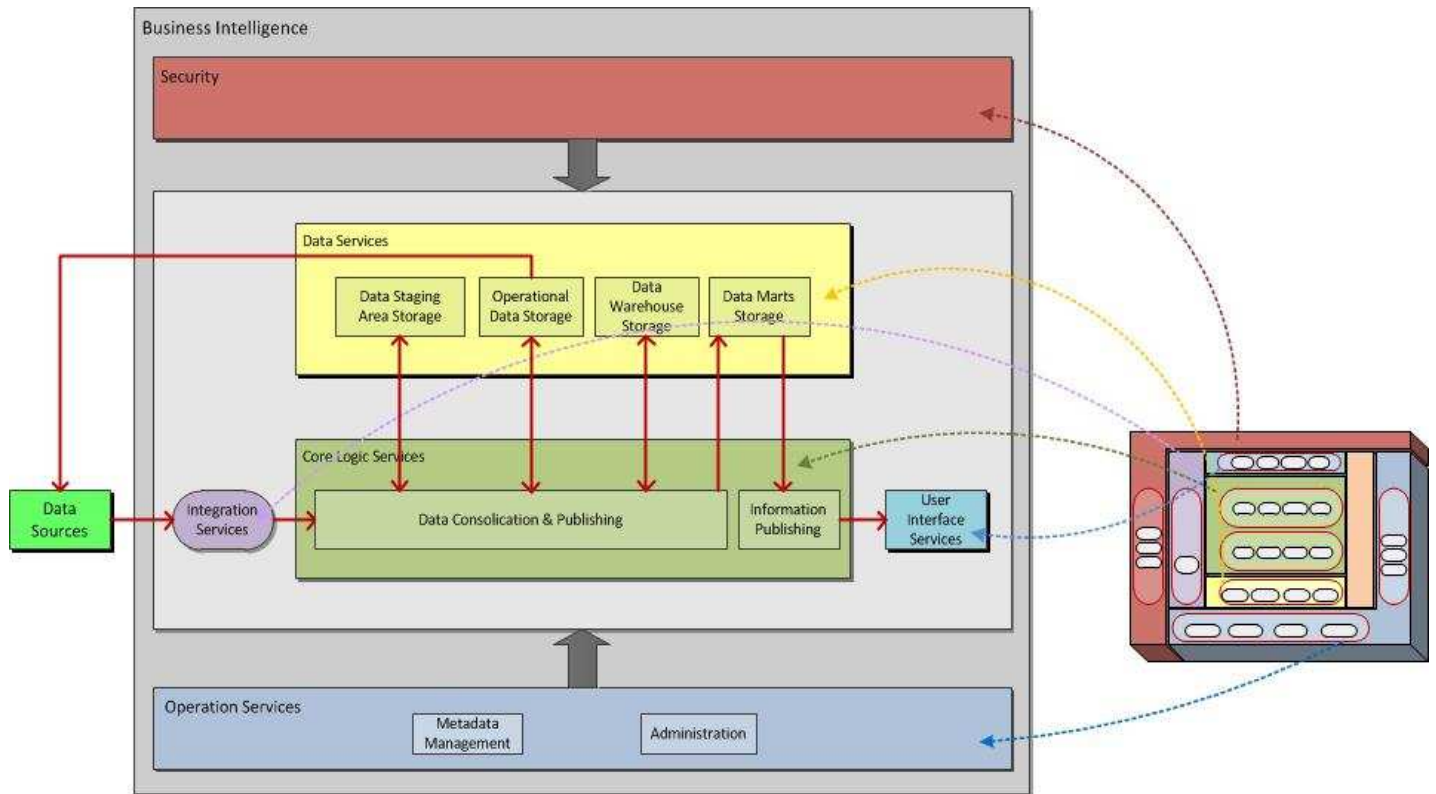


Figure 4-4 Business Intelligence Reference Architecture

However, in the Business Intelligence TAB Workshop meeting on June 8<sup>th</sup>, 2012, the owner of the Business Intelligence Competence Center pointed out that there were two problems in the Data Services Set of the Reference Architecture which had not been appropriately considered. The Data Staging Area Storage is different from the other operation in the Data Services Set. Actually in a Business Intelligence system, there is not any real physical area for the storage of Data Staging Area Storage. It is just an inner storage of data for further use in a short instance. At the same time, the Operational Data Storage, in the Data Services Set, does not transfer data back to the data sources. Instead, it gives the data to the real time business services. This condition exists only in the Business Intelligence systems dealing with operational data. Therefore, we modified the illustration of the Reference Architecture like the following:

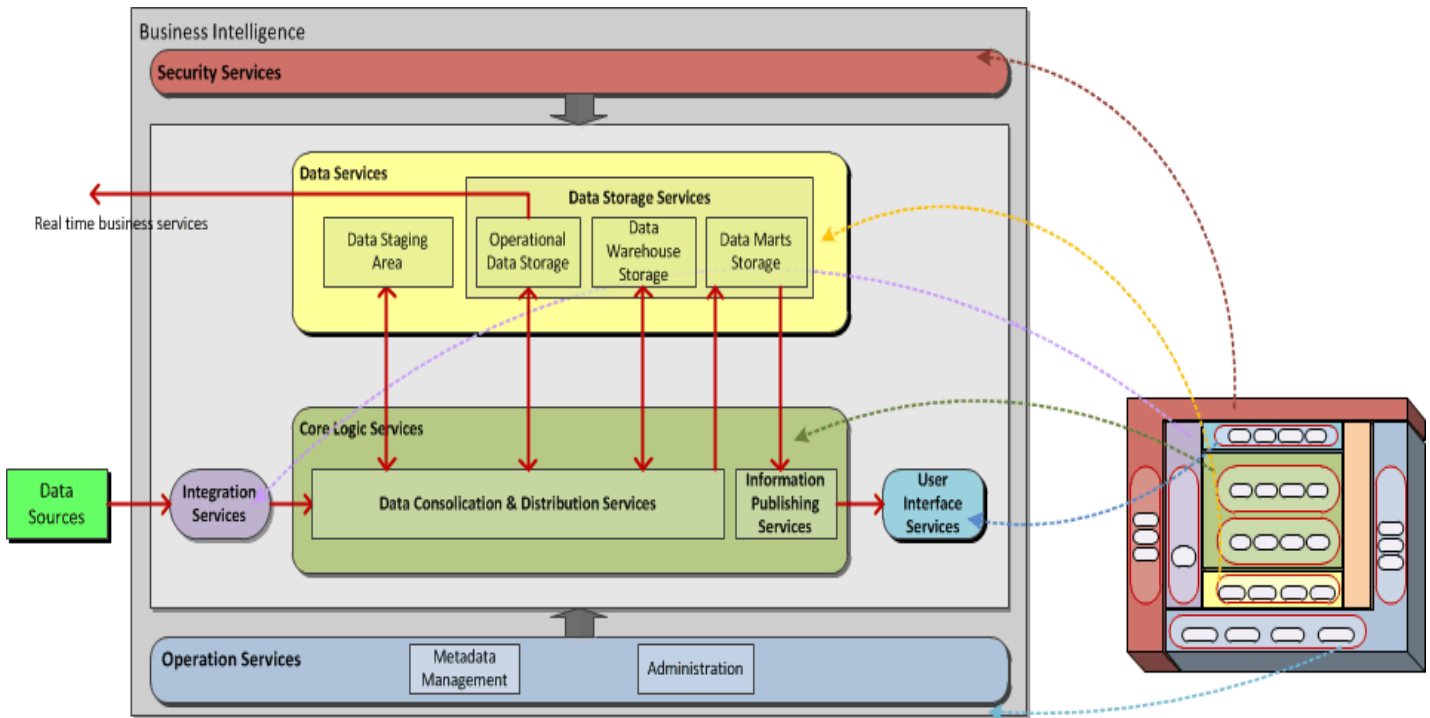


Figure 4-5 Business Intelligence Reference Architecture (Modified)

In the first version of BI – ITRefCard validated by the TAB, we transferred the image into a more visible format.

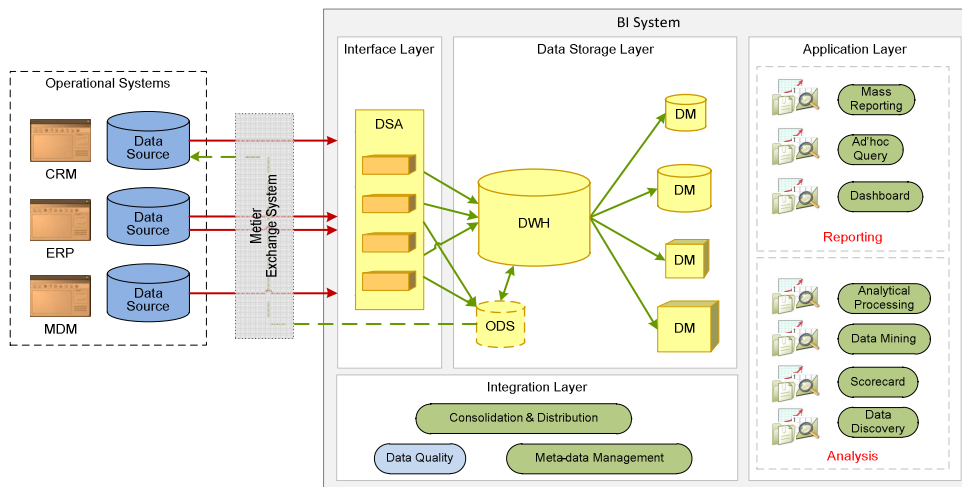


Figure 4-6 Reference Architecture for BI system

This architecture of BI system is a framework for organizing data stores, information processing components and end-users services. The following schema describes the reference architecture of a classical BI system, and its mapping with the software services of the BI-SSRM. This logical and generic model reflects a consensus within BNP Paribas (and the BI market) for BI systems and their components and services.

A BI system is a set of reporting and analysis applications federated by a data storage layer based on a data warehouse. This data warehouse (DWH), at the heart of the BI architecture, consolidates and centralizes a copy of operational

data periodically gathered from multiple Métier sources (customer relationship, sales, marketing, referentials, etc.), and feeds different data marts (DM or Cube) dedicated to specific usage of these data by those BI applications.

The process of data integration therefore plays a significant role in typical BI architecture.

- Data sourcing begins with Data Staging Area service (DSA) that receives data from operational applications (various platforms in different formats) via Métier Exchange System (external to BI system). This DSA is like a “sas” for the BI system and contributes to decoupling it with operational systems.
- Next, the data integration process, upstream and downstream of the DWH, is implemented by consolidation and distribution services (internal Integration Layer). These services extract data from DSA and transform raw data into the organized and reconciled form, consistent with the DWH where the data are loaded and consolidated. After that, they distribute data from the DWH toward different targets (DM) corresponding to their particular schema to meet dedicated business requirements.

Transformations of collected data in DSA can be strong, with control, cleansing, synchronization and structuring, consistent with the generic data model of the DWH. And some data can pass through an operational data store (ODS is an optional<sup>3</sup> component of BI architecture) for additional operations or short-term usages before loading in DWH. This complex integration process between all data components and services of the BI system is usually built with an ETL tool (cf. Exchange Systems IT Reference Card): some of these tools also provide Meta-data Management and Data Quality Management services.

Lastly, BI applications present indicators to business users and enable them to analyze historical data.

- The information can be delivered in a static way (accessed/calculated by an execution during the constitution of a document) or a dynamic way (the data are accessed/calculated during the access to the document or put out beforehand by a batch execution).
- Restitutions are delivered in two ways: push (the user gets access to the application) or pull (a link on a document is transferred by mail).

These end-users' applications only access to data exposed by DM; there is no direct and concurrent access to DWH in order to prevent problems with performance and reliability.

### 4.2.3 Platform Instantiations

---

<sup>3</sup> Among the various roles that an ODS can play, the workgroup retained “standard ODS” as defined by BICC (in opposite or complement of “General ODS” and “Support ODS”). This ITRefCard will evolve in the future in order to formally specify other possible positions of this component (with associated rules and guidelines) by taking account of experience feedbacks and future usages.

As we said in the paragraph 3.1.2 Software Integration Patterns, the IT Group of BNP Paribas has its own patterns to symbolize the design of a distributed system. It offers some basic patterns to symbolize an IT system. And also EAGLE Patterns have some pre-designed models for the future use of development, which somehow are famous in the domain of system development.

In the Business Intelligence systems within the IT Group of BNP Paribas, manipulate the issues of two different kinds of data: Master Data and Transactional Data.

#### 4.2.3.1 ETL implementation

In the case of master data, we take up the implementation of Extract - Transaction - Loading. It is the classic method in Business Intelligence systems.

The illustration shows the simple channel of data processing. We get data from external source systems or internal source systems. The data can be data of database and files. The connectors take in the data and put them forward into ETL tools, which implements the Data Consolidation & Publishing Service defined in the BI-SSRM.

The service of extracting data has interaction with the Data Staging Area. And the Loading Service has interaction with Data Warehouse Service in two ways. One is to load data into the data warehouse. The other is to load the data from the data warehouse into data marts for the use of each subject and area. At last the data will be forward to the User Interface Service.

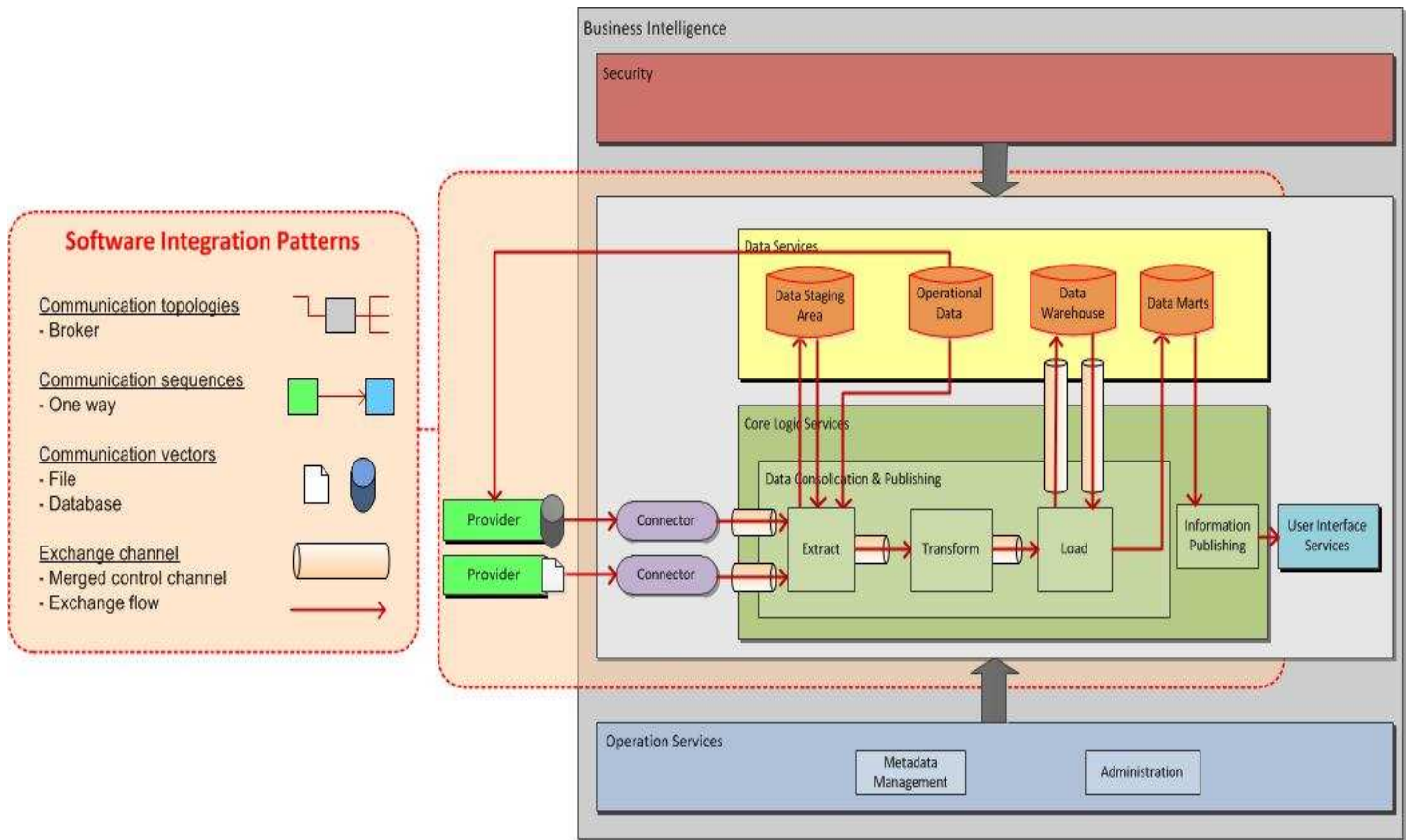


Figure 4-7 Business Intelligence Patterns Implemented by ETL

However, as we said in the former paragraph, BICC gave us the feedback of our wrong understanding about the Operational Data Service. Therefore we did the following modification. The Operational Data Service can not put back data to the source system.

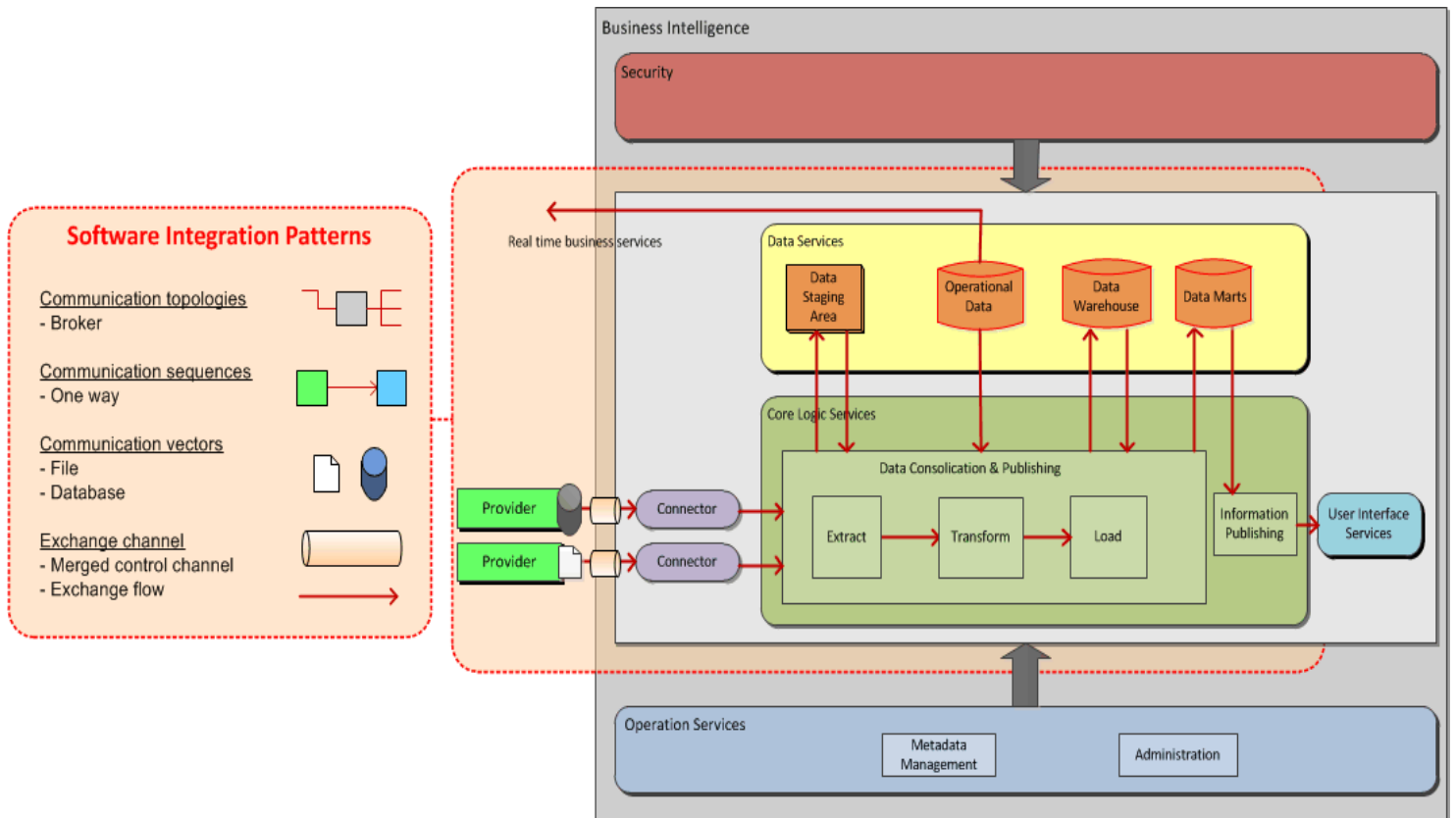


Figure 4-8 Business Intelligence Patterns Implemented by ETL (Modified)

#### 4.2.3.2 Transactional Data implementation

In the case of dealing with transactional data, we design the implementation of using MOM architecture. All the data transferred into the system will be packed like a message and sent to the target components within the system. And we use applications of message adapter and publisher to receive and send messages. And a message bus to coordinate messages.



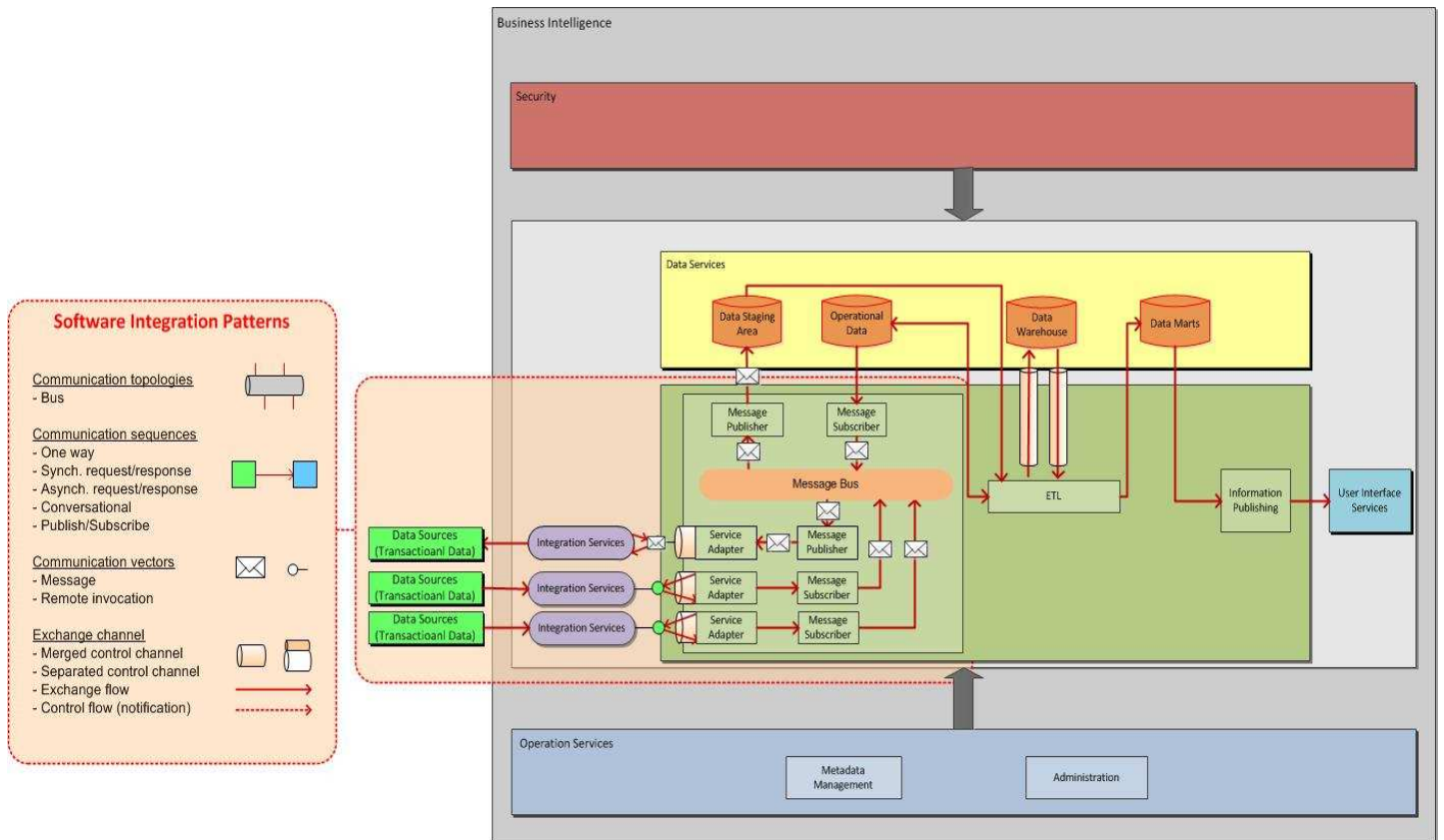


Figure 4-9 Business Intelligence Patterns Implemented by Message

Of course, we had the same problem of Data Staging Area Service. So we modified the pattern and did modifications. It is not the same storage media as the other data services and we can not use the same icon to symbolize it in the pattern.

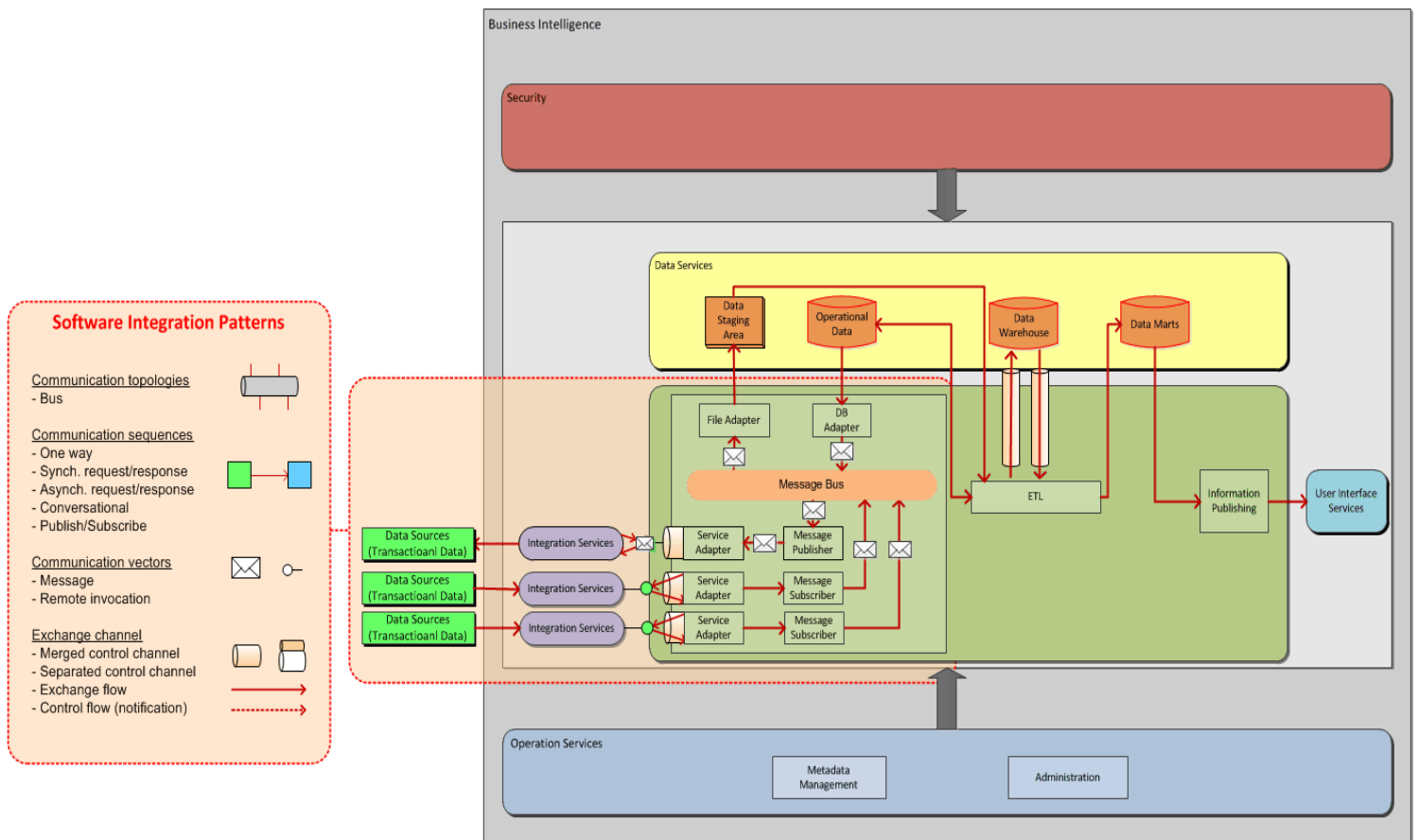


Figure 4-10 Business Intelligence Patterns Implemented by Message (Modified)

### 4.3 Governance Elements

Due to the time limit, we failed to define the governance elements for the Business Intelligence ITRefCard by the end of the August.

### 4.4 Rules and Guidelines

For this part, I define the rules and best practices in designing, developing a Business Intelligence system. I concluded all these rules and best practices from the existing documents of BICC. In the two meetings of Technical Architecture Board, I first proposed our results. The architects and managers present at the meetings offered their remarks to modify the contents of our proposals. I had a second version of the rules and best practices, and I do this job in this iterative way.

In the middle of the month of July, the latest version of the rules and best practices were classified into these eight categories:

- Modularity and Autonomy
- Data Responsibility & Access to Information

- Data Integration
- Data Storage
- Development and Group services integration
- Data Restitution Architecture
- Standardization, Industrialization & Reuse
- MDM system Integration

At the same time, in each of the great classes below, we have several topics to make the rules and best practices more precise and specific.

To present all the information clearly and readable, we use tables to illustrate all the rules and best practices.

( R = Rule, B = Best Practice, D = Definition )

#### 4.4.1 Modularity and Autonomy

Topic	No	Rule / Best Practice	Type
Modularity and Autonomy	1.1	The construction of a single BI system for information system must be a goal for every Métier. (Exception: in some countries, data can be stored locally.)	R
	1.2	A technical infrastructure should be divided into two main parts, the first contributing to the data integration and the second to the restitution of information.	B
	1.3	The zoning security rules must be used to define the BI system solution.	R
	1.4	Regarding to its particular functions, the OLAP sub-system should be placed in the application security zone to improve performances and simplify data access architecture.	B

#### 4.4.2 Data Responsibility & Access to Information

Topic	No	Rule / Best Practice	Type
Data Responsibility & Access to Information	2.1	A BI function must not directly access the information of another BI system.	R
	2.2	A BI user must be authenticated on each needed BI system.	R
	2.3	Data access control and user authorization management should be made only at data service level	B

#### 4.4.3 Data Integration

Topic	No	Rule / Best Practice	Type
Data Integration Service - General	3.1	The Data Integration Service should be implemented through the Data Staging Area Service and Data Consolidation & Distribution Services.	B
	3.2	For Data Integration Service, the duration of the storage is volatile. The collected data should be archived and/or deleted once the provision of Data Storage Service is validated.	B
	3.3	For Data Integration Service, the transformation of collected data in DSA should be strong and it consists of control, clean, synchronize and reform the data.	B
	3.4	Data integration must be forbidden / impossible at BI user desktop level.	R
	3.5	Direct user access and concurrent access to Data Storage bypassing Information Publishing Services must be forbidden.	R

Topic	No	Rule / Best Practice	Type
	3.6	“Data Consolidation & Distribution Services” should use an internal Exchange System (ES) within the BI system (usually ETL-based).	B
	3.7	When existing, the Métier intra-application should not be used as the internal BI ES.	B
Data Staging Area	3.8	A Data Staging Area must be put in place in order to collect, capture and integrate data.	R
	3.9	The Data Staging Area is a “sas”, from which the data sources are treated for the purpose of Data Storage Service by an integration platform.	D
	3.10	The Data Staging Area should deliver the source data in the format transferred by the integration services.	B
	3.11	The data store temporary in the DSA before Data Mart transmission must be un-useful.	R
Connection & Data Collection	3.12	Data duplication should be managed and controlled through an exchange system (external and/or internal).	B
	3.13	The collected data must not be modified at the BI level. If necessary, the changes must be carried out in the source systems by the duly identified data owners.	R
	3.14	Business rules must be implemented in a unique ETL integration process.	R
	3.15	The services of Extract/Receive, Cleansing/Transform and Load/Send should not be mutualized within one single task or domain.	B
	3.16	Volume of data which are transferred by the interfaces in the process of Extract/Receiving should be minimized as possible.	B
	3.17	At deployment level, it should be useful to have the same technical implementation for ODS and Data Warehouse. This situation is defined as an 'Active Data Warehouse.	B

#### 4.4.4 Data Storage

Topic	No	Rule / Best Practice	Type
Data Storage Service - General	4.1	The Data Storage Service should manage a unified model of information, which takes into account the different lifecycles and different restitution modes.	B
	4.2	The Data Storage Service must not be the main source to feed the operational systems.	R
	4.3	The Data Storage Services propose a unified model of data, globally adopted in all the components, which takes into consideration different life circles and information restitutions.	D
	4.4	Update of Data storage schema must be traceable. Traceability must be involved in the evolution process of the data schema.	R
Operational Data Storage	4.5	The Operational Data Storage service data model should be subject-oriented and namely organized by information domains.	B
	4.6	Three topologies of Operational Data Service can be considered when we designed our own architecture: Standard ODS, Global ODS and Operational Support ODS.	D
Data Warehouse Storage	4.7	If (exceptionally) the data of DW become a reference for the other systems, this information must be available for other information systems through ODS services or through the data consolidation & distribution services. Data warehouse data must not become a reference for other BI systems.	R
	4.8	The Data warehouse data should be at the same detailed level as the data source applications.	B
	4.9	Data warehouse data must never be modified or deleted. (Only format and model can be updated.) If a modification is really necessary, it is done by a DBA in the form of adding corrective data.	R
	4.10	Data warehouse historical data must be dated and cannot be modified or deleted. (Only format and model can be updated.)	R

Topic	No	Rule / Best Practice	Type
	4.11	<p>A strategy of access may be defined to redistribute the accesses to DW in an adequate way with the accesses to the data marts:</p> <ul style="list-style-type: none"> <li>- Explicit access to the data through the tools of restitution for questioning detailed and synthesized data.</li> <li>- Implicit accesses to the detailed data through the consultation of the data of data marts when "drill" is needed.</li> </ul>	B
Data Marts Storage	4.12	At deployment level, it should be useful to have the same technical implementation for Data Marts and Data Warehouse. This situation is defined as a 'Virtual Data Mart'.	B
	4.13	The data model of Data Mart Storage should be defined according to the functions to be implemented.	B
	4.14	<p>Data Mart data must never be modified or deleted. (Only format and model can be updated.)</p> <p>The data of Data Marts Storage cannot be updated except very rare occasions when the data can be modified by inserting new occurrences</p>	R
	4.15	<p>Data marts historical data must be dated and cannot be modified or deleted.</p> <p>(Only format and model can be updated.)</p>	R

#### 4.4.5 Development and Group services integration

Topic	No	Rule / Best Practice	Type
Services	5.1	Usage of internal BI software development kit should be limited to specific Métier missing functions.	B
	5,1	To manage user access, BI solutions must use the Group directory and associated authentication mechanisms.	R
Data Structure	5,3	A dimension of the data is composed by one or several tables concerned.	D
	5,4	A dimension table should be identified by a non-significant key. The key is managed independently from the original key and table, which comes from a Métier application or a	B

Topic	No	Rule / Best Practice	Type
		referential application.	
	5,5	The attributes of the dimension table should be defined with textual attributes.	B
	5,6	Elementary data should be memorized at the same data source detail level.	B

#### 4.4.6 Data Restitution Architecture

Topic	No	Rule / Best Practice	Type
Data Restitution Architecture	6.1	The data from the Data Storage Services should be accessed through a client-server application, a web application or a portal.	B
	6.2	The information can be delivered in a static way (accessed/calculated by an execution during the constitution of a document) or a dynamic way (the data are accessed/calculated during the access to the document or put out beforehand by a batch execution).	D
	6.3	The data are delivered in two ways: push (the user gets access to the application) or pull (a link or a document is transferred by mail).	D

#### 4.4.7 Standardization, Industrialization & Reuse

Topic	No	Rule / Best Practice	Type
Standardization, Industrialization & Reuse	7.1	Favor the reutilization of Métier platforms and shared infrastructure platforms.	B
	7.2	Products standardization policy at Métier level should be aligned with referenced technical architectures.	B
	7.3	Components and transversal developments must be reused between Métier BI projects.	R
	7.4	The adopted solutions should respect products status in the BI family of Norms et Standard catalog.	B



Topic	No	Rule / Best Practice	Type
	7.5	A map of data integration should be defined in order to assure a unified management of data architecture and rules of management among different BI system for guaranteeing the consistence and the reconciliation of their data. For the reason of well memorizing, the reconciliation should be assured the definition table of common dimension.	B
	7.6	Sub BI system architecture must permit a reasonable user response time for a common usage.	R

#### 4.4.8 MDM system Integration

Topic	No	Rule / Best Practice	Type
MDM system Integration	8.1	Data responsibility should be assumed by the MDM system before any transfer into a BI system.	B
	8.2	In a BI system dedicated to client reporting, the consolidation by category should be managed in upstream.	B
	8.3	Common governance for data dictionary and interfaces should be implemented when data consolidation is done by BI systems.	B
	8.4	BI systems must not bypass MDM system by directly accessing data at data source level. The BI system must be the client of the MDM system.	R
	8.5	Reference data should be made available to BI systems through standardized data flows (See Data Integration chapter and recommended usage of an exchange system).	B
	8.6	The data quality should be addressed by MDM system; it should be a part of the MDM system SLA.	B

## **5 Conclusion and Future Task**

### **5.1 Conclusion**

From the project demonstrated in this report, we know that generally speaking, a Business Intelligence system is composed of source system, collecting system, data layer and presentation system.

From another point of view, we can define a Business Intelligence system with following services: Data Integration Service, Data Service, Data Consolidation & Distribution Service, Information Publishing Service and Data Presentation Service.

Accordingly, we can propose the rules and best practices ( or guidelines ) for each of the services.

### **5.2 Future Task**

End the end of August 2012, we did not finish the final version of the rules and best practices. Meanwhile, we did not have a clear definition of the governance elements of Business Intelligence.

For this part, we can firstly consult the general governance elements for the process of developing a Business Intelligence system. At the same time we review the governance elements in the IT Group of BNP Paribas. Depending the different phases of the designing and developing, we put the roles into the positions.

## Glossary

BI	Business Intelligence
BICC	Business Intelligence Competency Center
BNPP	BNP Paribas Group
BPM	Business Process Management
CGP	Comité Grand Projet
CNMA	Comité des Normes, Méthodes et Architecture
DM	Data Mart
DSA	Data Staging Area
DW	Data Warehouse
EA	Enterprise Architecture
EAGLE	Enterprise Architecture Group Language
ESB	Enterprise Service Bus
ETL	Extract Transform and Loading
FAB	Functional Architecture Board
GSSRM	General Software Service Reference Model
IPS	Infrastructure & Production Services
ITG	Group IT Function
ITP	Information Technology and Process
ITR	ITRefCard
MDM	Master Data Management
MOM	Message-Oriented Middleware
ODS	Operational Data Storage
OLAP	On-Line Analytical Processing
RTA	Reference Technical Architecture
SSRM	Software Service Reference Model

STA	Standard Technical Architecture
TAB	Technical Architecture Board

## Bibliography

- 1) Z. Chen. *Data Warehouse and Data Mining*. Tsinghua University Press, Beijing, May 2009
- 2) K. Cios, W. Pedrycz, R. Swiniarski. *Data Mining Method for Knowledge Discovery*. Kluwer Academic Publishers, Norwell, Massachusetts, USA, 1998
- 3) A. Berson, S. J. Smith. *Data Warehouse, Data Mining, & OLAP*. McGraw-Hill, 1997
- 4) G. Xia, W. JIN, G. Zhang. Research about Current Situation and Development of Business Intelligence in China. *Science & Technology progress and Policy*, PP. 173-176, January 2006
- 5) X. Zhu. Review of Business Intelligence Theory and Practice. *Computer Systems & Applications*, PP. 114-117, No. 1, 2007
- 6) A. Bara, I. Botha, V. Diaconita, I. Lungu, A. Velocanu, M. Velicanu. A Model for Business Intelligence System's Development. *Informatica Economică*, PP. 99-108, vol. 13, No. 4/2009
- 7) J. Ranjan. Business Intelligence: Concepts, Components, Techniques and Benefits. *Journal of Theoretical and Applied Information Technology*, PP. 60-70
- 8) S. Negash. Business Intelligence. *Communications of the Association for Information Systems*, PP. 177-195, Volume 13, 2004
- 9) J. Su, Y. Huang, Research of Data Warehouse Architecture.
- 10) C. M. Olszak, E. Ziemba. Approach to Building and Implementing Business Intelligences Systems. *Interdisciplinary Journal of Information, Knowledge and Management*, Volume 2, 2007
- 11) J. Nenortaite, R. Butleris. Improving Business rules Management through the Application of adaptive Business Intelligence Technique. *Information Technology and Control*, Vol. 38, No. 1, 2009
- 12) A. Bernstein, B. Groszof, F. Provost. Business Intelligence: the Next Frontier for Information Systems Research? *Proceeding of the Workshop on Information Technologies and Systems in conjunction with the International Conference on Information Systems*, New Orleans, LA, USA, Dec. 15-16, 2001
- 13) Q. Zhang. Analysis on the Current Situation and trend of Business Intelligence. *Securities & Futures of China*, PP. 14-17, No. 2, 2009
- 14) Better Information through Master Data Management – MDM as a Foundation for BI. *An Oracle White Paper*, July, 2010
- 15) B. Rotibi. Development Intelligence: Business Intelligence for Software Development. April 2007
- 16) Multi-Dimensional Modeling with BI. *An SAP White Paper*, May 16, 2006
- 17) D. M. Walker. Overview Architecture for Enterprise Data Warehouse. *A White Paper from Data Management & Warehouse*, March 1, 2006
- 18) EAGLE: Functional Architecture, BNP Paribas internal document
- 19) EAGLE: Generic Software Service Reference Model, BNP Paribas internal document
- 20) EAGEL Meta Model, BNP Paribas internal document
- 21) EAGLE: Dictionary of Software Integration Patterns, BNP Paribas internal document
- 22) EAGLE: The 4 Levels of Representation of the Technical Architecture, BNP Paribas internal document

- 23) EAGLE: How to Build a Functional Architecture, BNP Paribas internal document
- 24) EAGLE: Functional to Application Transition, BNP Paribas internal document
- 25) EAGLE: Business to Functional Transition, BNP Paribas internal document
- 26) EAGLE: Using the Generic Software Service Reference Model, BNP Paribas internal document
- 27) Fondamentaux architecture technique décisionnelle du groupe BNP Paribas, BNP Paribas internal document
- 28) Architecture fonctionnelle des systèmes d'information décisionnelles du groupe BNP Paribas, BNP Paribas internal document

**Appendix – Product: Business Intelligence ITRefCard**

# Business Intelligence

<b>Owner:</b>	ITG - ATR
<b>Publication date:</b>	2012/07/10
<b>Release:</b>	0.0.0
<b>Status:</b>	Draft

<input type="checkbox"/> Lead description	<input type="checkbox"/> Glossary	<input checked="" type="checkbox"/> For immediate implementation
<input checked="" type="checkbox"/> Reference Technical Architecture	<input checked="" type="checkbox"/> Rule	<input type="checkbox"/> At definite date
<input type="checkbox"/> Standard Technical Solution	<input type="checkbox"/> Guideline	<input type="checkbox"/> According to the Métiers needs
<input type="checkbox"/> Technical Component		

**Executive summary**

**Summary of rules**

**Technology overview**

**Definition**

Business Intelligence is a broad category of applications and technologies for gathering, providing access to, and analysing data for the purpose of helping enterprise users to make better business decisions and to acquire a global view of the Métier activities.

Business Intelligence symbolizes a process including collecting data from multiple functions of an enterprise, formalizing and cleansing information, consolidating and diffusing data according to the usage of each specific domain or subject. It consists of a series of interactive activities of exploring and analysing structured domain-specific information to discern business trends or patterns, thereby deriving insights and drawing conclusions.

At the same time, Business Intelligence which usually includes a portfolio of decisional applications relying on a specific dedicated Information System segregated from operational IS, permits piloting the enterprise by providing a detailed or synthetic view, by facilitating the process of making decision and pilotage of activities, and by contributing the performance of commercial activities.

With Business Intelligence superior tools, now employers can also easily convert their business intelligence via the analytical intelligence to solve many business issues. At the same time, BNPP can identify its most profitable customers and underlying reasons for those customers' loyalty, as well as identify future customers with comparable if not greater potential.

And also, Business Intelligence analyses click-stream data to improve ecommerce strategies.

It can discover money-laundering criminal activities, detect and deter fraudulent behavior, such as from usage spikes when credit or telephone cards are stolen.

Meanwhile, Business Intelligence helps to reduce risk exposure through more accurate financial credit scoring of their customers.

A Business Intelligence system determines what combinations of products and service lines customers are likely to purchase and when and determines with attribution and churn analysis why customers leave for competitors and /or become the customers.

It also sets more profitable rates for insurance premiums.

**Benefits**

Business Intelligence provides many benefits to departments and functions which use it. It can eliminate a lot of uncertainty of Métier data within BNPP; enhance the communication among the departments while coordinating activities, and enable organizations to respond quickly to changes in all the areas and domains of the enterprise.

Business Intelligence improves the overall performance of BNPP in that BI assures the quality and accuracy of the data and information. Another reason why BI leads to the enterprise's competitively superior performance is expediting decision-making through acting quickly and correctly on information before competing rivals.

Last but not least, customers' experience is improved owing to timely and appropriate response to their problems and priorities with the assistance of BI.

#### **Attention Points**

When an entity of the Group decides to launch a Business Intelligence solution, we often face the following problems:

- **Initiative:** Business Intelligence is not a project, but rather a transversal program. After a BI solution is launched, we will find that more and more users and data will be added to it. Therefore, as time goes by, we need to invest continuously in BI in the organization. Work of BI does not finish like that of a common project. This concept needs to be dealt with and clarified in organization building BI.
- **Silo:** Sharing data of Business Intelligence in the organization demands that different units and entities of the organizations should overcome the internal data silos that exist within the organization.
- **Proliferation:** With too much Business Intelligence software and product to choose from, the organization runs BI from too many different vendors.
- **Data governance:** The Organization must face the problem of data quality, spatially in the dimension of fixing and managing data. Data need to be cleansed and checked.
- **Uncontrolled Data:** Sometimes, we can find that the organization stores its data in many excel files or even plain text files, so managing data, understanding them and connecting them become too difficult and complicated.
- **Meta Data:** In order to launch a Business Intelligence solution, we need to define simply and accurately the questions that need to be answered. For example, defining business terms and questions accurately will help the BI analyst and developer and answer customers' needs more easily.

#### **Ecosystem Overview**

#### **Governance Elements**

### **BNPP context and strategy**

In order to enhance BI systems, BNP Paribas follows the orientation which it gets from matching BI and strategy Métier.

BI contributes actively to the strategy areas defined by BNP Paribas:

- The client orientation
- The cost control
- The openness to partnership
- The effectiveness of piloting

The five principals of Information System Enterprise Architecture applied in BI:

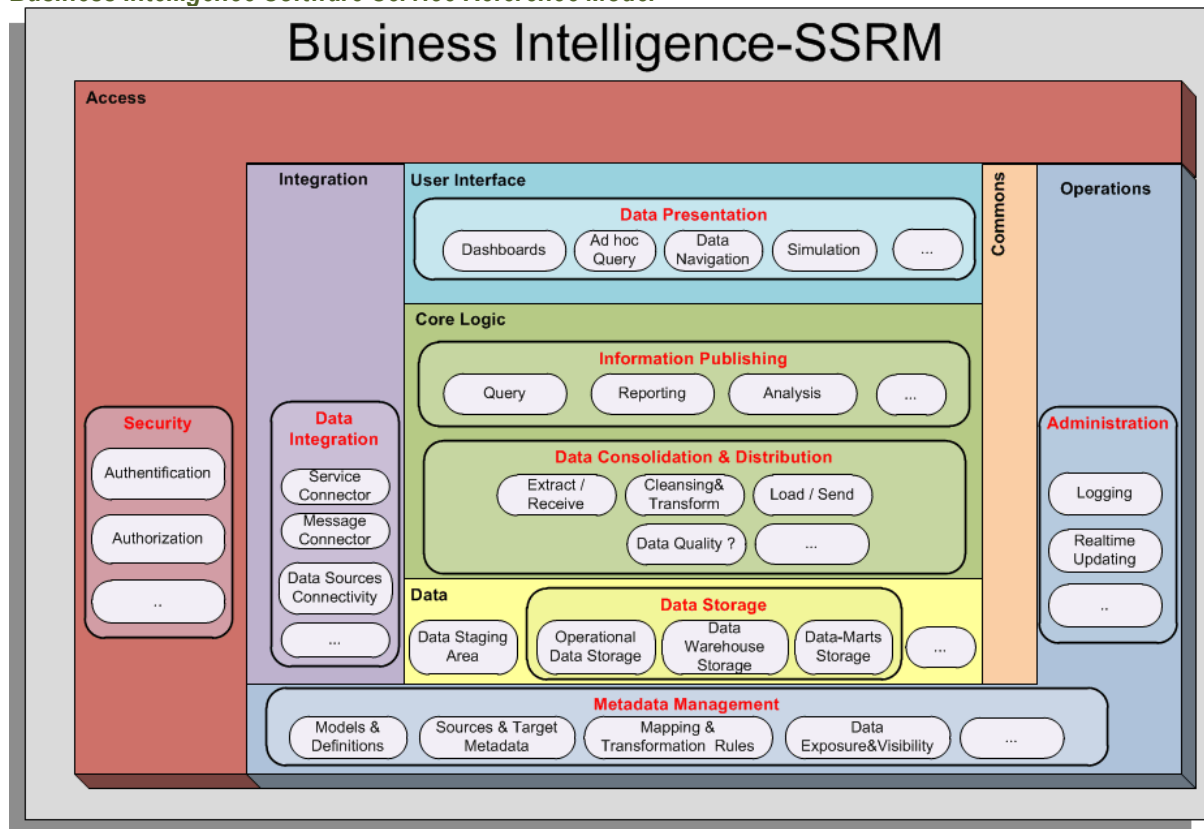
- BI is multiple layered.
- BI is a Group level system.
- Each BI is specific to certain Métier.
- BI follows the standard structure in the specific functional area.
- BI follows the concept norm in IS.

The Métier decisional functions which constituted a BI are included in the Métier Information system. So we can conclude that the BI is a part of the Métier information system. Moreover, there is no single BI which is not a part of an information system. As result, there is no a BIG BI for whole BNP Paribas.

In every level of urbanization we can find functions of a BI.



*Business Intelligence Software Service Reference Model*



**User Interface - Data Presentation:** This block of services contains the user interfaces which permit to exploit information of BI in diverse forms.

It is the most visible part of a BI system and provides access to the system. This block of services contains the functions to display the data in various ways. It provides straight and direct way for users to understand Métier data and to figure out the condition of business execution.

**Core Logic - Information Publishing:** This block of services contains functions to prepare data and reports for specific applications with the help of analyzing and reporting tools. They implement the interfaces access to information contained in the Data Storage Services.

**Core Logic – Data Consolidation & Distribution:** This block of services contains functions to extract data from data sources and transforms raw data into the organized and reconciled form coherent with the data warehouse where the data will be loaded and consolidated. After that, it distributes data into different data marts corresponding to particular schema and business. It implements the information of a Métier or a group of particular users. Moreover, it prevents direct accesses to the data warehouse from direct analysis.

**Data – Data Storage:** This block of services proposes a unique and global data model prepared for the group of the components, which takes into consideration different life circles and information restitution. Operational Data Storage contains current and integrated data consulted for reporting and analysis near the operation, or the operational support for the Métier applications. Data Warehouse contains the historical, integrated and neutral data for the strategic requirements of reporting and analysis prepared for a group of Métier users. Data Mart contains historical and integrated data for the requirements of reporting and analysis specific for users of a certain Métier.

**Data – Data Staging Area:** This block of services implements part of the services of data collection and exchange. The source data are treated here for further use of Data Storage Services.

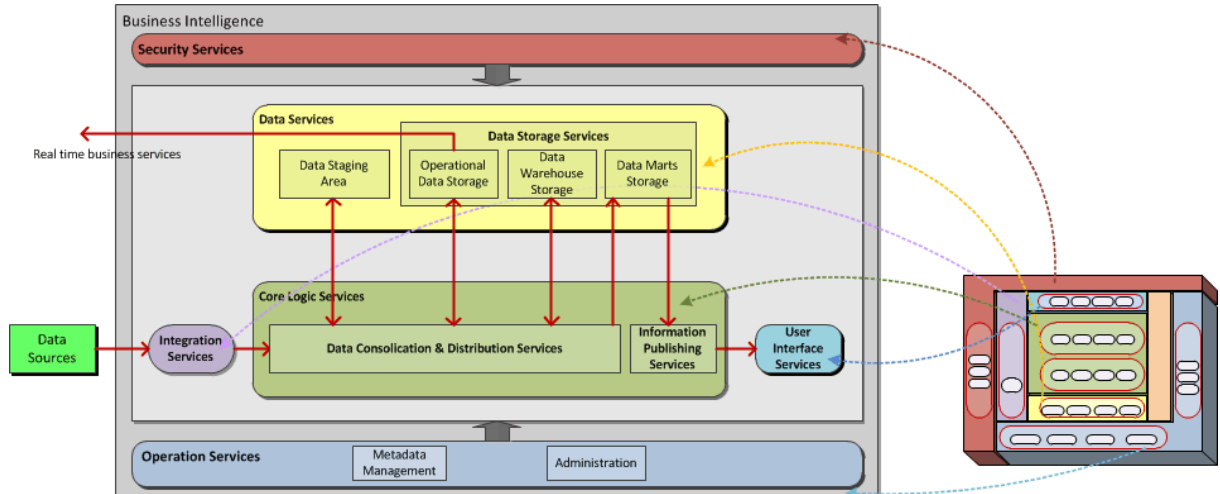
**Integration – Data Integration:** This block of services satisfies the requirements of collection of referential data from Métier applications, supporting applications and external applications. It is responsible for the acquisition of data flow of multiple sources and manages the reconciliation and the synchronization of these flows in order to produce target data.

**Access - Security:** this block of services refers to all the measures that are taken to protect access to services from the people outside Métier. At the same time, it makes sure that those who want to get access to certain services do have the right authorization and right identification.

**Operation – Metadata Management:** This block of services contains the functions to assure interoperability among all the information sources and inter data storage.

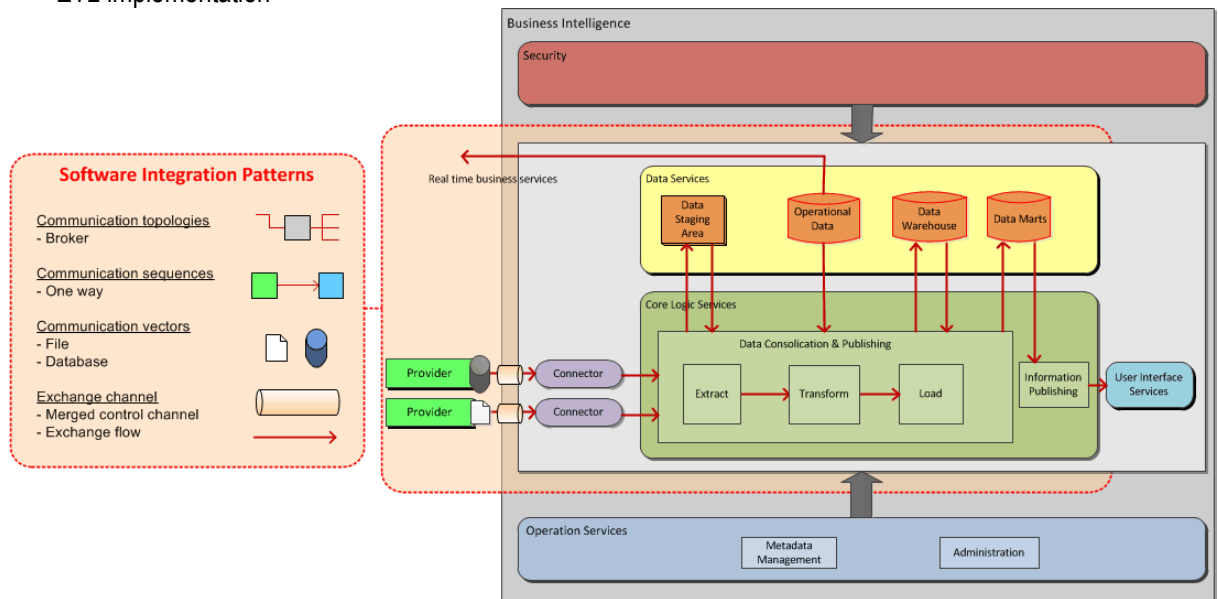
**Operations - Administration:** this block of services offers the capabilities to maintain and operate the BI (starting/halting, modifying exec. parameters...) and its supported exchanges (subscription management...)

### Reference Architecture

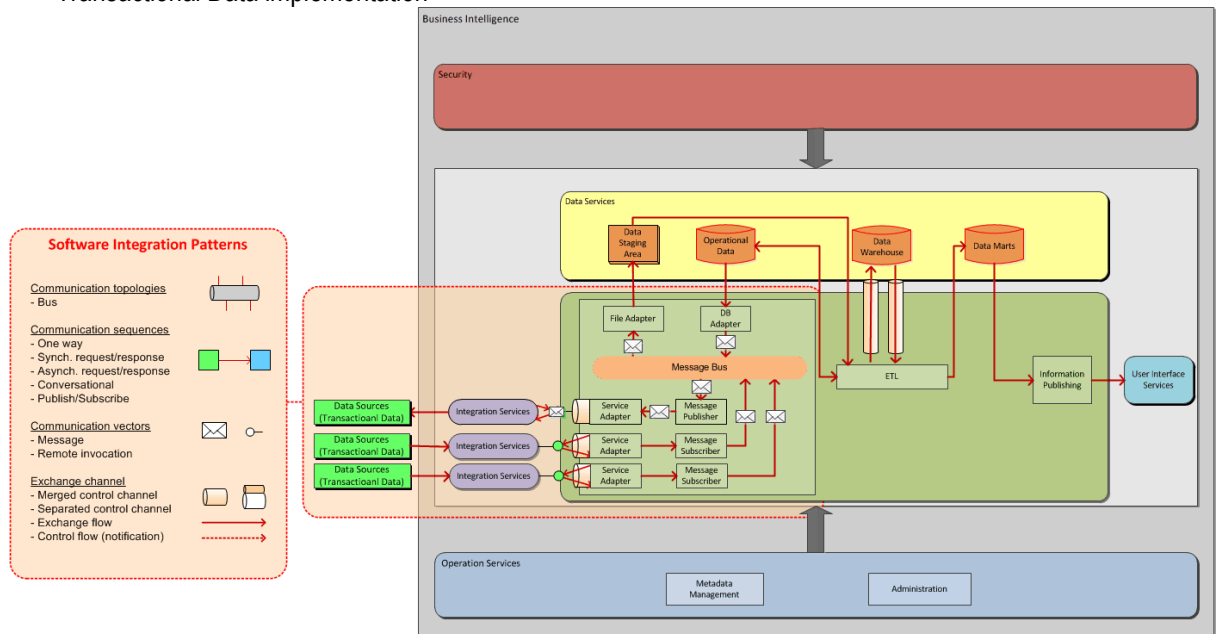


## Platform Instantiations

- ETL implementation



- Transactional Data implementation



**Modularity and Autonomy**

Modularity and Autonomy	1	The construction of a single BI system for information system must be a goal for every Métier. (Exception: in some countries, data can be stored locally.)	R
	2	A technical infrastructure should be divided into two main parts, the first contributing to the data integration and the second to the restitution of information.	G
	3	The zoning security rules must be used to define the BI system architecture.	R
	4	Regarding to its particular functions, the OLAP sub-system should be placed in the application security zone to improve performances and simplify data access architecture.	G

**Data Responsibility & Access to Information**

Data Responsibility & Access to Information	5	A BI function must not directly access the information of another BI system.	R
	6	A BI user must be authenticated on each needed BI system.	R
	7	Data access control and user authorization management should be made at least at data level	G

## Data Integration

Data Integration Service - General	8	The Data Integration Service is implemented through the Data Staging Area Service and Data Consolidation & Distribution Services.	DefouPatt
	9	For Data Integration Service, the duration of the storage is volatile. The collected data should be archived and/or deleted once the provision of Data Storage Service is validated.	G
	10	For Data Integration Service, the transformation of collected data in DSA should be strong and it consists of control, clean, synchronize and reform the data.	G
	11	Data integration must be forbidden / impossible at BI user desktop level.	R
	12	Direct user access and concurrent access to Data Storage bypassing Information Publishing Services must be forbidden.	R
	13	"Data Consolidation & Distribution Services" should use an internal Exchange System (ES) within the BI system (usually ETL-based).	G
	14	When existing, the Métier intra-application ES should not be used as the internal BI ES.	G
Data Staging Area	15	A Data Staging Area must be put in place in order to collect, capture and integrate data.	R
	16	The Data Staging Area is a "sas", from which the data sources are treated for the purpose of Data Storage Service by an integration platform.	Definition
Connection & Data Collection	17	Data duplication should be managed and controlled through an exchange system (external and/or internal).	G
	18	The collected data must not be modified at the BI level. If necessary, the changes must be carried out in the source systems by the duly identified data owners.	R
	19	Business rules must be implemented in a unique ETL integration process.	R
	20	The services of Extract/Receive, Cleansing/Transform and Load/Send should not be mutualized within one single task or domain.	G
	21	Volume of data which are transferred by the interfaces in the process of Extract/Receiving should be minimized as possible.	G
	22	At deployment level, it should be useful to have the same technical implementation for ODS and Data warehouse. This situation is defined as an 'Active Data warehouse'. The ODS being simply made up of dedicated tables in the Data Warehouse.	G Def (pour Active DW)

## Data Storage

Data Storage Service - General	23	The Data Storage Service should manage a unified model of information, which takes into account the different lifecycles and different restitution modes.	G
	24	The Data Storage Service must not be the main source to feed the operational systems.	R
	25	The Data Storage Services propose a unified model of data, globally adopted in all the components, which takes into consideration different life circles and information restitutions.	Definition

	26	Evolutions of BI Data schemas must be traceable.	R
Operational Data Storage	27	The Operational Data Storage service data model should be subject-oriented and namely organized by information domains.	G
	28	Three topologies of Operational Data Service can be considered when we designed our own architecture: Standard ODS, Global ODS and Operational Support ODS.	Definition
Data Warehouse Storage	29	Data warehouse data must not become a reference for other BI systems.	R
	30	The Data warehouse data should be at the same detailed level as the data source applications.[G] The data stored in the DW and the data of the source applications should have the same detail level	G
	31	Data warehouse data must never be modified or deleted. (Only format and model can be updated.)	R
	32	Data warehouse historical data must be dated and cannot be modified or deleted. (Only format and model can be updated.)	R
Data Marts Storage	33	At deployment level, it should be useful to have the same technical implementation for Datamarts and Data warehouse. This situation is defined as a "Virtual Datamart".[G] Depending on the performance, security and cost constraints, the Data Mart(s) and the DW can share the same technical infrastructure ("Virtual DM approach")	G
	34	The data model of Data Mart Storage should be defined according to the functions to be implemented.	G
	35	Data marts data must never be modified or deleted. (Only format and model can be updated.)	R
	36	Data marts historical data must be dated and cannot be modified or deleted. (Only format and model can be updated.)	R

### ***Development and Group services integration***

Services	37	Usage of internal BI software development kit should be limited to specific Métier missing functions.	G
	38	To manage user access, BI solutions must use the Group directory and associated authentication mechanisms.	R
Data Structure	39	A dimension of the data is composed by one or several tables concerned.	Definition
	40	A dimension table should be identified by a non-significant key. The key is managed independently from the original key and table, which comes from a Métier application or a referential application.	G
	41	The attributes of the dimension table should be defined with textual attributes.	G
	42	Elementary data should be memorized at the same data source detail level.	G

### ***Data Restitution Architecture***

Data Restitution Architecture	43	The data from the Data Storage Services should be accessed through a client-server application, a web application or a portal.	G
	44	The information can be delivered in a static way (accessed/calculated by an execution during the constitution of a document) or a dynamic way (the data are accessed/calculated during the access to the document or put out beforehand by a batch execution).	Definition
	45	The data are delivered in two ways: push (the user gets access to the application) or pull (a link or a document is transferred by mail).	Definition

### **Standardization, Industrialization & Reuse**

Standardization, Industrialization & Reuse	46	Favor the reutilization of Métier platforms and shared infrastructure platforms.	G
	47	Products standardization policy at Métier level should be aligned with referenced technical architectures.	G
	48	Components and transversal developments must be reused between Métier BI projects.	R
	49	The adopted solutions should respect products status in the BI family of Norms et Standard catalog.	G
	50	A map of data integration should be defined in order to assure a unified management of data architecture and rules of management among different BI system for guaranteeing the consistence and the reconciliation of their data. For the reason of well memorizing, the reconciliation should be assured the definition table of common dimension.	G
	51	Sub BI system architecture must permit a reasonable user response time for a common usage.	R

### **MDM system Integration**

MDM system Integration	52	Data responsibility should be assumed by the MDM system before any transfer into a BI system.	G
	53	In a BI system dedicated to client reporting, the consolidation by category should be managed in upstream.	G
	54	Common governance for data dictionary and interfaces should be implemented when data consolidation is done by BI systems. ("when data consolidation is done by BI systems?" to be discussed with HME)	G
	55	BI systems must not bypass MDM system by directly accessing data at data source level. The BI system must be the client of the MDM system.	R
	56	Reference data should be made available to BI systems through standardized data flows	G
	57	The data quality should be addressed by MDM system; it should be a part of the MDM system SLA.	G

Some topics to be discussed:

Security and zoning

Product components layout to respect zoning security rules

→eg: Data or Application security zone for OLAP ?

Data access control and user authorization management

- Currently at data services level
  - User authorization cross BI systems?
- Integration with Group services

- Group directory & associated authentication mechanisms, Neos and MS Office, group portal...

#### Development

- In-House development & BI product customization → Limited usage of SDK to extend BI functions
- To avoid migration issues

#### “Big Data”

- Unstructured data support
- Convergence of operational and decisional data architectures

#### Quality of Service

- No direct & concurrent access to Data Storage bypassing Information Publishing Services

#### Data Integration Architecture: MDM, Exchange Systems...

- Recommended usage of an internal ES within the BI system (usually ETL-based) for “Data Consolidation & Distribution Services”
- Clarify role & responsibilities between the intra-BI ES and any Métier inter-application ES

#### Data Integration Architecture: MDM Integration (recommendations based on ‘club MDM’ workshops)

- Data responsibility should be assumed by the MDM system before any transfer into a BI system.
- In a BI system dedicated to client reporting, the consolidation by category should be managed in upstream.
- Common governance for data dictionary and interfaces should be implemented when data consolidation is done by BI systems.
- BI systems must not bypass MDM by directly accessing data at data source level
- the BI must be the client of the MDM.
- Reference data should be made available to BI systems through standardized data flows (Recommended usage of an exchange system).
- The data quality should be addressed by MDM; it should be a part of the MDM’ SLA.

#### Components

- Promotion of BI Suites & criteria for niche products
- Open-source policy
- Data marts: OLAP vs. RDBMS
- BI Appliance usage
- ETL (and/or ESB) within BI
- In-memory

### Operation overview

### Topic management

#### Main contacts

#### Reference documents