Develop The Enterprise Data Warehouse as a New Instrument For a Banking Governance

I.D.E.A. Project
BANCA POPOLARE DI MILANO
Mario Migliori – Gianni Fuolega
12° Italian Banking Group

Employees 6,990
Branches 505
Retail Credits € 15,1 billion
Total Assets € 28,3 billion
Net Profit 2000 € 228 million

Banca Popolare di Milano
Project Goals

I.D.E.A. ⇔ Integrate Data & Applications

“.....the main goal is to reorganize the bank’s data assets in a way to assure the sharing of a unique and integrated matrix of data that can be used by all Management Information Systems modules to track and examine bank events.”

from Bank President speech on Annual General Meeting in April 2001
Project Goals

- Create an Enterprise Data Warehouse as a precursor to:
  - Manage in an **integrated way** the information available in Management Information Systems
  - Have a **complete trace of rules** of data processes
  - **Reconcile** financial and accounting data Vs. book-keeping data

- Realize specific Data Marts to:
  - Increase the User autonomy on operative analysis with data
  - Decrease maintenance costs
The Pyramid Vision

EDS EIS DSS

E.M.I.S.
(ALM, VAR, Marketing, Risk, Accounting, Book-Keeping, etc.)

DATAWAREHOUSE

Legacy Systems

External Data System

Enterprise Management Information System

Executive Information System (TdB)

Decision Support System
E.M.I.S. Modules

Compulsory Fulfilment
Central Bank
Book-Keeping

E.M.I.S.
Enterprise Management Information System

Risk Management
- Credit Risk
- Financial Risk
- Operational Risk
- Internal Auditing

Control Management
- Management Accounting
- Cost Allocation

Market Development
Marketing Management
Planning & Budgeting
Project Requirements

- Any relevant data element collected from a data source is stored in a single place
- Rule Management integrated and managed directly by Users
- Enterprise wide Data Warehouse
- Specialized Data Marts
- Data Model designed and focused on Bank requirement without forcing a standard model inside the Bank specifications
- M.I.S. modules with specific algorithms are outside the Enterprise Data Warehouse
Implementation Process

Tools used through the phases of implementation process

- Logical Model & Physical Model (CA Erwin)
- Program (SAS WA)
- Enterprise Data Warehouse (IBM DB2)
- Data Marts (SAS WA, SAS MDDB, SAS IntrNet)
Technical Architecture

Operational & Legacy System

Guide System

Legacy Systems

Operational & Legacy System

Extract, Mapping & Validation

Legacy 1
Legacy 2
Legacy 3
Legacy 4
Legacy n

Errors Correction

Files

File 1
File 2
File ...
File n

OPERATIONAL DATA DEFINITION

DATAWAREHOUSE

RULE MANAGER

CONTROL FUNCTIONS

Acquisition
Re-classify
Enrichment
Calculation

Enterprise Data Warehouse

TRANSLATOR MANAGER

Subsetting
Mapping
Summarizing

DATA MARTS TOOLS

QUERY & REPORTS
OLAP
DATA MINING
AUDIT TRAIL
EIS

DATA MART
DATA MART
DATA MART
DATA MART

SPECIALIST MODULES

Accounting Management
Budgeting
Marketing Management
Financial Risk Management
Credit Risk Management
Book Keeping
Central Bank

ENTERPRISE DATA WAREHOUSE

DATA MART

METADATA

Recycle

Acquisition
Re-classify
Enrichment
Calculation

Recycle
Technical Architecture
(Equity Management Data)

Host 2064-1C6 (1280 mips)
Host 2064-1C8 (1615 mips)

- Legacy Systems: 150 daily files - 25 million records
  10 monthly files - 1 million records
- E-Data Warehouse (DB2): 250 tables - 450 million records
  (25 months on line and 7 years archived)
- Data Marts (SAS/MDDB): 20 tables - 25 million records

n.2 Server NT – IBM Netfinity 5500
- Rule Manager: 8.600 rules (whose 4.500 Book-Keeping)

P/C Client Windows
- Data Model: 10 Subject Area - 220 Entity – 1.150 Attribute
SAS Tools

SAS ADD-INS USED
- Register DB2 Table
- Load DB2 Table
- Slowly Change Dimension
- Multiple Join Generator
- Generate ALL Code
- Append Data Set
- Data Dictionary

METADATI

SAS/Warehouse Administrator

SAS/Access to DB2
Linguaggio SAS 4GL

SAS/Connect
MDDB Server

Host

P/C Users

Enterprise Miner
Enterprise Reporter

Register DB2 Table
Load DB2 Table
Slowly Change Dimension
Multiple Join Generator
Generate ALL Code
Append Data Set
Data Dictionary
Project Plan

- Phase 1: Technical Architecture, Rule Manager and Equity Management Model
- Phase 2: Financial Intermediation Model
- Phase 3: Lending & Services Model
- Phase 4: Saving & Deposit and Services Model
- Phase 5: Other services

Timeline:
- May 2000
- Apr. 2001
- Dec. 2001
- Jun. 2002
- Dec. 2002
- May 2003
Project Team

- **GIEFFE**
  - Project Management
  - Data Model Design

- **B.P.M.**
  - Project Management
  - User & Technical Support

- **Back Office Analysis**
  - Engineering
  - SOL-TEC

- **Front Office Analysis**
  - Price Waterhouse Coopers

- **Book-Keeping Analysis**
  - Arthur Andersen

- **Tech. Design & Programs**
  - sas
Success Key Factor

- **Management Sponsorship:**
  - Complex project with enterprise wide impact on bank organization

- **Project Team mixed:**
  - Users, Technicians & Consultants “ad hoc” with specialized and professional knowledge

- **Metodology:**
  - *de-coupling of data* between Operational & Legacy Systems versus Management Information Systems
  - *Data Model designed on BPM needs* without copy or cloning a data model from other experience
  - clear comprehension of analysis process
  - technical tools with a covering of the whole developing cycle
  - strategic vision but realization by steps (five phases)
Authors:
Mario Migliori – Banca Popolare di Milano
Gianni Fuolega – GIEFFE Management Consultants

Presentation Title:
“Develop the Enterprise Data Warehouse as a New Instrument for a Banking Governance”

Presentation Type:
Main conference stream (with case study element)

Presentation Content:
Industry Focus: Banking
Business or IT Solution:
Data warehouse development and implementation
“Develop the Enterprise Data Warehouse as a New instrument for a Banking Governance”

Authors:
Mario Migliori – Banca Popolare di Milano
Gianni Fuolega – GIEFFE Management Consultants

Abstract

The rapid pace of the Financial Services marketplace is placing increasing pressure on management within traditional banks to make increasing demands for immediate access to structured information in all sectors.

It was these demands which led to the I.D.E.A. (Integrate Data & Applications) project. IDEA aims to create a single integrated platform containing all relevant data to meet the business needs of management to analyse profit, credit & financial risk and to enable the effective planning of commercial and marketing activity.

The project was divided into five main phases, each addressing a specific area of business as follows: Equity Management, Financial Intermediation, Lending, Savings & Deposits, and Services.

Within each area there are four modules which perform discrete services in the following:

1. Acquisition and collation of each relevant data element in a single unique instance.
2. Centralisation of the rules extraction, extrapolation and correlation which transform the data element into Information.
3. Storage of the information in a single Enterprise Data Warehouse based on IBM DB2® technology.
4. Creation and management of data marts based on specific user needs to distribute the information in specialised local environments.

Other main technological components used in the project include SAS Data Warehouse Administrator®, SAS SCL® Language, SAS MDDB®, SAS IntrNet® and ERwin®.

Company Profile

The Group “Bipiemme - Banca Popolare di Milano”, is a leading Northern Italian banking group and the twelfth largest group in Italy. The Group predominately provides commercial banking service for both retail and small and medium-sized corporate customers and, in addition, offers its customers capital market services, brokerage services, debt and equity underwriting, asset management, insurance underwriting and sales, leasing and factoring services.

At the end of year 2000, the Group had 505 branches (mainly in the Region of Lombardy) with 6,990 employees, with a total assets of € 28,3 billion, a total retail credits of € 15,1 billion and a net profit of € 228 million.

1 mario.migliori@bpm.it - Banca Popolare di Milano – Milan - Italy
2 fuolega@gieffenet.it - GIEFFE Management Consultants – Venice – Italy
Introduction

The acronym goes a long way in explaining the aim of the project: **I.D.E.A.: Integrate Data & Applications**. That is the “idea” which originated the project.

There are many reasons why a project like IDEA should come about, but we think that the dynamic of Management Information Systems with various, different and continued requests from Users, combined with the technical evolution of software tools is one the most important.

The bank developed a similar project more than ten years ago and it was clear that the old instruments and process was not going to be able to continue to support strategic decision and organizational change with the same tools.

This is the back ground within which the bank decided to start with a feasibility study where the main goal was to identify which were the old User requests that were answered incompletely and which were the new Users requests to which the actual systems were not able to produce answer.

The result of feasibility study demonstrated that it was time for a change in the Management Information System process.

Project Goal

At the last Annual General Meeting, the President of the Bank said that the main goal of the IDEA Project is to reorganize the bank’s data asset in such a way as to assure the sharing of a unique and integrated matrix of data that can be used from all Management Information Systems modules for examining bank events.

With these few words we can find all the key information that we need to understand the project.

In particular the goals can be divided in two main areas:

- The first is to create an Enterprise Data Warehouse that can be used for anybody that wishes to process and transform management data to information;
- The second is to design and realize specific Data Marts, managed directly by end users, in a way that they can increase their autonomy in data analysis.

Moreover we think that with this architecture many activities in software maintenance will be concentrated in only one direction with the direct effect of a reduction in maintenance costs.

The Vision

IDEA Project has a main vision about how the Enterprise Data Warehouse is inserted in the Information System overall Architecture, in particular between the Production System and the Management Information System.

The following picture explains in a graphic way the concept expressed:

Looking at the picture we can identify the Enterprise Management Information System (E.M.I.S.) as the main “user” of Enterprise Data Warehouse.

Moreover at the top of pyramid it is possible to identify some other systems that are used to complete the support to the management decision. Such systems are:
- **Executive Information System (EIS):** used to present to Top Management a few specific and critical indicators that explain how the bank is going. These systems could be named as: Tableau de Bord, Balanced Scorecard, etc.

- **Decision Support System (DSS):** used to simulate the impact of alternative hypotheses;

- **External Data System (EDS):** used to managed and compare external with internal data.

To understand the E.M.I.S. modules is very important to have a clear comprehension of which kind of users need to interface with the Enterprise Data Warehouse.

The E.M.I.S. modules can be divide into the following four main area:

- **Control Management:** is the area in which are present all the modules used to control how the bank in going, such as: Management Accounting, Cost Allocation, Planning & Budgeting.

- **Risk Management:** is the area in which are present all the modules used to measure and control the different kind of risks, such as: Credit Risk, Financial Risk, Operational Risk and Internal Auditing.

- **Market Development:** is the area in which are present all the modules used to control and develop the market with commercial and marketing actions, such as: Marketing Management and Planning & Budgeting (this is the same module present in Control Management area).

- **Compulsory Fulfilment:** is the area in which are present all the modules used to produce all the compulsory output requested by law, such as: Book-Keeping and Central Bank reporting.

The following picture present the EMIS module in summary.

Some project requirements have been defined before starting the project.

- **Integrity:** means that any relevant data element collected from a data source is stored uniquely in a single place. Moreover the rules used to process data are integrated within a unique module and managed directly by Users.

- **Multidimension:** means the ability to examine the same data with different dimensions related to the User’s need. It is possible to do that using an Enterprise wide Data Warehouse together at specialized Data Marts.

- **Flexibility:** means the ability to change the model simultaneously with any organization change. It is possible to do this using a Data Model designed on specific bank requirements that is able to make the desired modification. This would not be possible if would be used a model cloned from a standard approach.

On technical architecture design, we obtained system flexibility by moving specifics and dynamics algorithms in a separate area dedicated to specialist modules, outside the
Enterprise Data Warehouse. In this way the Users can change quickly their specialist module with low impact on the Enterprise Data Warehouse System, because the only activity to do would be to mapping the new system interface.

**The technical approach to the solution**

The approach used in the implementation process can be structured into four steps:

- The first step is to transfer the knowledge and the User requirements into a model. The ERwin® software tool was used to design the logical and physical models.
- The second step is to program the process that transform the data in the physical structure designed by the model. In this step the use of the SAS tools has been a key factor success.
- The third step is to put data into the Enterprise Data Warehouse IBM DB2® was used as Data Base Management System (DBMS).
- The last step is to create the Data Mart where SAS Multi Dimensional Data Base (SAS/MDDB®) was used as DBMS and also SAS tools were used for process management.

The following picture describe the process of the main modules used on technical architecture:

On examination of the technical architecture diagram it is easy to identify the main modules which perform discrete services in the following areas:

- Acquisition and collation of each relevant data element into a single unique instance.
- Centralisation of the rules extraction, extrapolation and correlation which transform the data element into Information.
- Storage of the information in a single Enterprise Data Warehouse based on IBM DB2® technology.
Creation and management of data marts based on specific user needs to distribute the information in specialised local environments.

The SAS tools used in that technical architecture are:

- SAS/Access® to DB2 (Host)
- SAS/Connect® (Host)
- SAS/4GL® Language (Host)
- SAS/Warehouse Administrator® (Server)
- SAS/Share® (Server)
- SAS/IntrNet® (Server)
- SAS/Connect MDDB® Server (Server)
- SAS/Enterprise Miner® (P/C client)
- SAS/Enterprise Reporter® (P/C client).

Some statistics regarding Equity Management Data emphasise the efforts that have been required in this first phase of the project:

- **Data Model**: 10 Subject Area – 250 Entity – 1.150 Attribute
- **Legacy Systems**:
  - 150 daily files processed with 25 million records
  - 10 monthly files processed with 1 million records
- **Enterprise Data Warehouse (DB2)**:
  - 250 tables with 450 million records stored on line for 25 months and archived for 7 years (as recommended by the Basel Committee)
- **Data Marts (SAS/MDDB)**:
  - 20 tables with 25 million records.

**Rule Manager**

The Rule Manager is the most important software module realized in the project.

The main function is to help the User to define and manage the functional rules that are needed to put data into the Enterprise Data Warehouse in the way and with the classification required.

The screen pictures presented on enclosures explain with some simple steps an example of how it works. The following paragraph describes the content of the screen pictures.

1. Screen picture 1 presents the introduction on the module with the User identification and its password. After this step the User can choose the product that to be examined.
2. Screen picture 2 presents which kind of product the User has defined for any input data source.
3. Screen picture 3 shows an example of a recursive process that is used when the definition of one product need more than one process.
4. Screen picture 4 shows an example of the Boolean algebra used to identify a product.
5. Screen picture 5 shows the output variables that are generated for the specific product defined in the previous steps.
6. Screen picture 6 shows the various ways that can be used to format the output variable, such as: constant, variable, look-up table and formula.
7. Screen picture 7 shows an example of output format using a formula.
8. Screen picture 8 shows an example of output format using a constant.
9. Screen picture 9 shows an example of output format using a look-up table.
10. Screen picture 10 shows the function used to look for data inside the data dictionary.

Of course the Rule Manager module is much more complex than the example described above.

In any case it is important to remember that the effort used to design and create the Rule Manager module can be shared by all project’s phases, and this is one of the various reasons that justify a generalised instead of a specific approach.

**Project Management**

A complex project like IDEA requires particularly strong project management organization, especially regarding two basic points: resource (personal, technical, and financial) and time.

The bank decided that knowledge is a success factor and decided to proceed with internal personnel together with external consultants in order to obtain an injection of knowledge and methodology.

Because the theme was new and very complex the bank decided to use different consultants specialized in their competence area. So the team project has been structured as follow:

- Banca Popolare di Milano: Project Management and User & Technical support;
- GIEFFE Management Consultants: experts in Data Modelling and project management;
- Engineering and SOL-TEC: experts on Financial Back Office procedures;
- Price WaterhouseCoopers: experts on Financial Front Office procedures;
- Arthur Andersen: experts on Book-Keeping procedures;
- SAS: experts on Technical Design and Programming.

In this way the IDEA Project received a strong external impulse oriented around the goals.

Regarding the project plan, the decision was first, to define what the Users expected from the project, to have a complete and total project vision and second to divide the entire project into five phases organized in a logical and functional way.

The five phases are the following:

- Phase 1: Equity Management;
- Phase 2: Financial Intermediation;
- Phase 3: Lending and its Services;
- Phase 4: Savings & Deposits and its Services;
- Phase 5: Other Services.
At the end of each phase the bank has a piece of Enterprise Data Warehouse complete with all the data related to that specific functional area.

In synthesis we retain that the following point are the really key success factor for a good job:

- Management Sponsorship, necessary because the project is very complex with enterprise-wide impact on bank organization.
- Mixed project team between Users, Technicians ad Consultants with specialized and professional knowledge.
- Methodology, there are few point that need attention:
  - De-coupling of data between Operational & Legacy System versus Management Information System;
  - Data Model designed on specific bank needs without copy or cloning a data model from other experience;
  - Clear comprehension of analysis process, so all component of team project work in the same way and the output can be used to everybody;
  - Technical tools with a covering of the whole developing cycle without technical interruption that need specific integration efforts;
  - To understand the strategic vision but to realize by steps.

In conclusion we can say that create an Enterprise Data Warehouse is a very ambitious project, but today it is the only solution to manage data in a highly integrated, multidimensional and flexible way.

**Trademarks**

- ERwin® is a product of Computer Associate Inc.
- IBM DB2® is a product of IBM®
ENCLOSURE

RULE MANAGER - Screen Picture nr.1

RULE MANAGER - Screen Picture nr.2