

Actions for data warehouse success

Aziza CHAKIR

Systems Architecture Team,
Laboratory of Informatics, System
and Renewable Energy
Hassan II University - ENSEM
Casablanca, Morocco

Hicham MEDROMI

Systems Architecture Team,
Laboratory of Informatics, System
and Renewable Energy
Hassan II University - ENSEM
Casablanca, Morocco

Adil SAYOUTI

Systems Architecture Team,
Laboratory of Informatics, System
and Renewable Energy
Hassan II University - ENSEM
Casablanca, Morocco

Abstract—Problem statement: The Data Warehouse is a database dedicated to the storage of all data used in the decision analysis, it meets the customer requirements, to ensure, in time, that a data warehouse complies with the rules of construction and manages the evolutions necessary of the information system (IS).

Results: According to the studies carried out, we see that a system based on a data warehouse governed by the best practices of The Information Technology Infrastructure Library (ITIL) and equipped with a multi-agent system will make it possible our direction to ensure governance tending towards the optimization of the exploitation of the data warehouse.

Keywords—Information Technology Infrastructure Library (ITIL); data warehouse; governance; insufficiencies of the data warehouse; multi-agent system.

I. INTRODUCTION

The data warehouse is not merely a new practice, it is found in all the fields and corporation having data. It is a true resolution in the computing world. All the conditions are well used for a good decision making.

The raw data extracted or transformed into information are transformed by an ETL (Extraction, Transformation and Load).

The information is stored in a data warehouse to be analyzed by tools for analysis transforming this information into knowing (Figure 1).

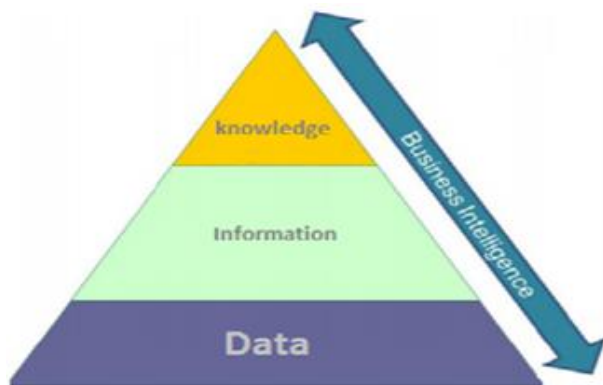


Fig. 1. Business intelligence [1]

A shallow analysis or erroneous input data may cause a wrong decision. A decisional project is prone sometimes of bugs or of dysfunction related to the bad human handling and it directly impacts the response time of a whole decisional chain.

To overcome these insufficiencies, we proposed a governance of the data warehouse based on the best practices of ITIL and of the multi-agent system.

It would be useful before presenting the method used to show the benefit of choosing ITIL and multi-agent system.

A. The benefit of choosing ITIL with data warehouse:

ITIL is a rather complete framework of reference which treats all the fields of the IT governance. . Its continuous updating and harmonization with other standards such as COBIT, ISO 27000, PMBOK and regulations such as SOX, Basel II solvency, and the possibility of using it in the data warehouses guided our choice.

B. The benefits of the governance of a data warehouse with the multi-agent system:

The governance of an information system present common point with the multi-agent system at knowing management by process [7]. This management is ideal for IT governance, which governance is not other than a set of processes in interaction between them for a better management of information technology.

II. DATA WAREHOUSE

The data used in decision making or the decisional analysis are stored to constitute a database, it is this same database which is called after data warehouse.

The tools of ETL provide the power of the data warehouse from production bases. It is only one simple copy of these data since the data warehouse transforms these last into information. Which information is transformed into knowledge through other algorithms business intelligence (Figure 2).

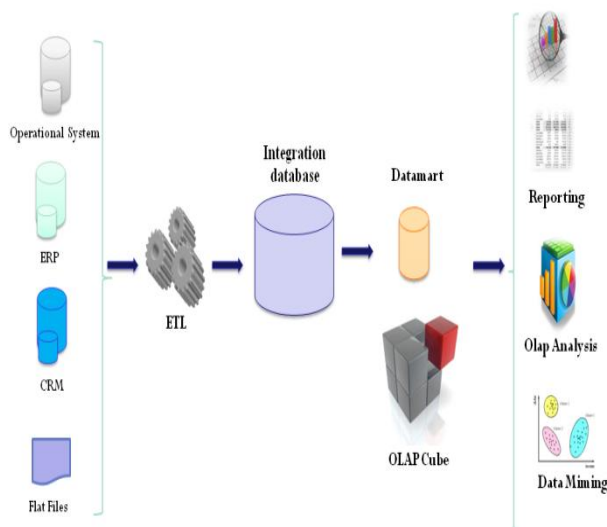


Fig. 2. The steps of feeding a data warehouse

The knowledge generated by the data warehouse is used for:

- To manage and/or predict (for example sales).
- To evaluate the risks (for example the risk customer for an insurance).
- To make a study of the behaviors of the customers in order to allow the companies to define strategies to target their customer.

A. Features of data warehouse

A data warehouse is a collection of « Subject oriented, integrated, nonvolatile, time variant collection of data in support of management decisions » [12] [13].

- Subject-Oriented:

The database is built according to the subject area, for example (customers, products, risk,...).

- Integrated:

The data comes from different production applications, can exist in all different forms. It must be integrated in order to standardize and give them a way, understandable by all users. In a data warehouse, the data must have a coding and a unique description.

- Non-volatile:

In the data warehouse, the data will not disappear and will not change with treatment over time.

- Time-Variant:

The historical data is kept in a data warehouse, The data is also stamped. We can see the evolution in time of a given value.

B. Insufficiencies of the data warehouse

The data warehouse is exposed at the risks related to [1]:

- An improper setting up, this is the case, for example, want to use this knowledge at all costs without checking the validity of data and even their volume.
- A poor quality of the data or a badly made analysis will involve erroneous results and bad decisions by the executive body of a company.
- The excellent opportunities given by a data warehouse are likely not to be exploited more, if the data warehouse set up causes changes for the users or of the bugs during their use.
- The need to change the decision in the case of development of a company (for example creation of new services).

It is clear that a data warehouse is not easy to implement and to maintain in operational condition.

III. GOVERNANCE

The governance of information systems [2] [3] [11], is the procedure that defines the way organizations are able to align IT strategy with business strategy, and to ensure that companies remain on track to achieve their goals, and implement good ways to measure performance.

IT governance provides effective, efficient and compliant computer to enable an organization to achieve its objectives use.

A data-processing framework of governance answers some key questions, such as the way in which the computer department functions as a whole and that which the management of the key indicators needs.

A. ITIL is a normative reference frame

ITIL (Information Technology Infrastructure Library) [4] [5] is a framework of best practices for the delivery of IT services. It helps to improve efficiency and reduce risk.

ITIL provides a methodological approach consisting of a series of modules to help companies and organizations to improve the use of IT resources.

ITIL consists of five modules, all modules will manage an IT service and align IT services with objectives of a company. The five modules are:

- Strategy of the services.
- Design of the services.
- Transition from the services.
- Exploitation of the services.
- Continuous improvement of the services.

B. ITIL and IT service support

Support of IT services using ITIL all aspects to ensure that the field of information technology can support the IT applications that provide business functions and can guarantee the continuity, availability and quality of service to users. [15]

Support service defined by ITIL is provided through five key processes and function and are used as following:

- Service Desk (function)
- Configuration Management (process)
- Change Management (process)
- Release Management (process)
- Incident Management (process)
- Problem Management (process)

The diagram below summarizes the key aspects of the methodology support service defined by ITIL (Figure 3).

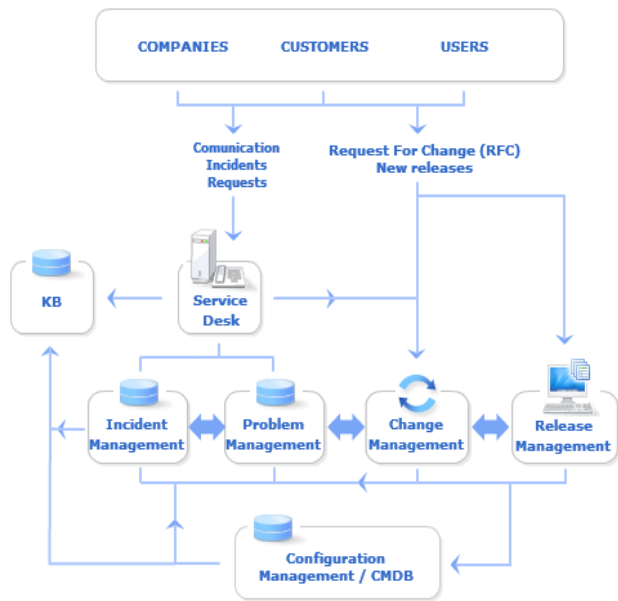


Fig. 3. The methodology of support of service defined by ITIL [1]

C. ITIL and IT service delivery

In order to monitor and improve the quality of the applications of the information system on levels of customer service, ITIL offers the following five processes:

- Management of service levels
- Capacity Management
- Availability Management
- IT service continuity management
- Financial management service

D. The role of ITIL in data warehouse

To keep the data warehouse in good condition, we must make good needs analysis, and conceive well the data which will be to use by it, and maintain the data warehouse provide a high quality service to customers. And ensure good management problems encountered, the infrastructure management, management of information dissemination, performance tuning database and the level of service provided.

ITIL is the best practice one to implement a data warehouse to work properly [6][8][9][10].

The following table (Table 1) shows the usefulness of ITIL processes to keep the data warehouse in good condition.

TABLE I. Role of ITIL in data warehouse

ITIL processes	Role of ITIL in data warehouse
Service Centre	Process to ensure the processing of all user expectations that these are simple requests or malfunctions caused by the data warehouse.
Incident management	The aim of incident management is to restore data warehouse in the shortest time, with minimal impact on users.
Problem Management	Process to optimize the level of service by analyzing the real causes of malfunctions and there by anticipating corrective action to address the shortcomings in organizing and controlling the use of resources.
Change management	Process describing the activities to quickly and efficiently conduct all changes to minimize the risk of negative impact of these changes on the quality of service.
Management put into production	Process to coordinate all activities related to the storage, management, distribution and implementation of data warehouse
Availability management	This process ensures a level of availability of the data warehouse to customers in accordance with the contract services and remaining financially viable [6].
IT service continuity management	This process ensures the continuity of the data warehouse in the event of incident [6].
Service level management	This process maintains the planning, contracting, implementation and monitoring of services and service levels, working with clients responsible for this activity and providers responsible for providing the service [6].

IV. MULTI-AGENT SYSTEM

A multi-agent system is a distributed software system consisting of several autonomous entities with different interests - agents, occurring at the same time, sharing common resources and communicating them.

Multi-agent systems can reduce the complexity of solving a problem by dividing the required subsets namely, combining independent to each of these subsets intelligent agent and coordinating the activity of these agents.

Modeling of the proposed solution is based on the principle of Multi-agent systems is: "Everyone must cooperate to achieve the same goal."

The proposed architecture is composed of the following agents:

- User Agent

Cognitive agent can communicate, intervene and monitor service-center Agent and process agent.

- Service-Center Agent

Reactive agent that reacts when the action is required.

- Process Agent

Hybrid agent, by stimulating service-center agent and cooperation with the user agent and the business agent, it takes a decision.

- Business Agent

Reactive agent, depending on the situation, it interacts with the process agent.

- Knowledge Base Agent

Reactive agent whose task is the retrieval of information from the knowledge base.

A. Functional specifications

A service is triggered by an event which is detected by the sensor of a process.

When the process started, it draws data from the knowledge base for possibly trigger other processes and generates an audit report.

Monitoring services will be measured in% for carrying out checks of all active processes.

The historical services provide a basis for an intelligent system.

The following diagram shows the operation of a process (for example "change indicator") (figure 4):

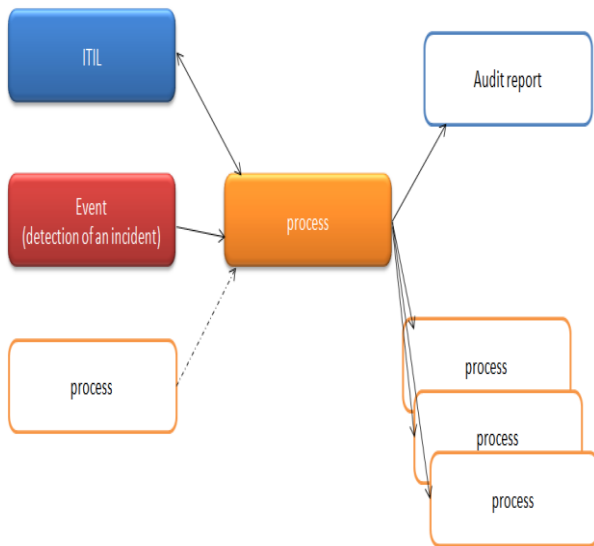


Fig. 4. Example-change indicator

V. CONCLUSION

According to the study there are gains to be made through better governance of a data warehouse. Using our method, we ensure customers satisfaction, quality service, and maintenance of a data warehouse in good condition.

Using the best practices of ITIL by professionals allows to highlight the operations that lead to improvements that can recognize weaknesses in our control.

Hence the need for a decision-making system, data warehouse, most reliable way to get to identify its problems seen to reduce the disturbances due to a malfunction in an IT services company, and this can only be done by governance of decision support system, data warehouse.

The proposed method suggests tactics for each blocking position to maintain the operability of a data warehouse.

We expect the development of a generic platform to represent the proposed method.

References

- [1] Maxime Poletto, " L'informatique décisionnelle-Thèse Professionnelle ", 01 juin 2012.
- [2] Morley+Al, " Processus métiers et S.I. - Gouvernance, management, modélisation - 3e édition".
- [3] Jamal Skiti et Hicham Medroumi, "La Gouvernance des Technologies de l'Information à base du Système Multiagent et le référentiel COBIT".
- [4] Pascal Delbrayelle, <http://www.itilfrance.com>.
- [5] Nicolas Dewaele, "L'ITIL: Un référentiel pour la qualité des systèmes d'information", 23 mars 2011.
- [6] Tariq Rahim Soomro et Hasan Yousef Wahba, "Role of Information Technology Infrastructure Library in Data Warehouses", 2011.
- [7] Jamal Skiti et Hicham Medroumi, "Nouvelle Méthodologie de la Gouvernance des Technologies de l'information à base du Système Multi-agent".
- [8] Yves B.Desfossés , Claude Y.Laporte , Alain April et Nabil Berrhouma, " Méthode d'amélioration des services de TI, basée sur ITIL, dans les entreprises québécoises ", septembre 2008.
- [9] Aziza Chakir, Hicham Medromi et Adil Sayouti, "La gouvernance du système d'information à base des bonnes pratiques d'ITIL V3", novembre 2012.
- [10] STEIGMEIER Alexandre, "Comment articuler les différentes normes et méthodes".
- [11] Bruno Claudepierre, "Conceptualisation de la Gouvernance des Systèmes d'Information, Structure et Démarche pour la Construction des Systèmes d'Information de Gouvernance", 10 décembre 2010.
- [12] W. H. Inmon, "Building the Data Warehouse", 4 edition (October 7, 2005).
- [13] Elzbieta Malinowski, Esteban Zimanyi, "Advanced Data Warehouse Design: From Conventional to Spatial and Temporal Applications".

AUTHORS

Aziza Chakir is an engineer in computer science from the ENSIAS, Mohammed V – Souissi University in July 2009, Rabat, Morocco. She prepare her thesis at ENSEM, Hassan II University, Casablanca, Morocco.

Hicham Medromi received the PhD in engineering science from the Sophia Antipolis University in 1996, Nice, France. He is responsible of the system architecture team of the ENSEM Hassan II University, Casablanca, Morocco. His actual main research interest concern Control Architecture, Architecture of System and Software Architecture Based on Multi Agents Systems and Distributed Systems. Since 2003 he is a full professor for Control systems and computer sciences at the ENSEM, Hassan II University, Casablanca. He managed eight Research projects and he has published several Patents and publications in international journals and conferences.

Adil Sayouti received the PhD in computer science from the ENSEM, Hassan II University in July 2009, Casablanca, Morocco. In the same year he received the price of excellence of the best sustained thesis in 2009. In 2003 he obtained the Microsoft Certified Systems Engineer (MCSE). In 2005 he joined the system architecture team of the ENSEM, Casablanca, Morocco. His actual main research interests concern Remote Control over Internet Based on Multi agents Systems.