Data Warehousing and Data Mining for the Management of Scientific Research

Mohamed El Mohadab
Dept MCS of SMSU
Beni Mellal, Morocco
m.elmohadab@gmail.com

Belaid Bouikhaleine
Dept. MCS of SMSU
Beni Mellal, Morocco
b.bouikhalen@usms.ma

Fahd Ouatik
Dept LPHA of CAU
Marrakesh, Morocco
fahd.ouatike@gmail.com

Said Safi
Dept. MCS of SMSU
Beni Mellal, Morocco
safi.said@gmail.com

ABSTRACT

Scientific research [1] represents a very important axis for the university, because it ensures its innovation and its productivity and develops the competencies of their researchers and the reputation and the glow of their research laboratory.

But the management and automation of this sector represents a great challenge for universities either for managers, directors, or the researchers from which comes the need of find relevant and effective solution.

To manage this sector, we have to study several computer solutions which ensure good management of the information system, in this direction we find that the Data Warehouses and the Data Mining are relatively well mastered when it comes to scientific research data because it is increasingly complex, diverse.

The Data Warehouse and the Data Mining is recognized as the core of the decision-making system: it integrates and stores data from the different functional areas of an organization for make it easily accessible to decision-making processes on order to ensure adaptation to the new change that can be brought to the system.

Keywords

Data warehouse, data mining, Data mart, OLAP, Modeling Process, Scientific Research, Research Management, Decision Support Systems.

1. INTRODUCTION

Lately the world has entered into an era that is called the societies and knowledge in which science, knowledge innovation its source of wealth. In this vision the university represents a pillar of development of each country.

In the last decade, the use of information communication technologies became important in all sectors, this technological revolution enhance establishments to integrate the information communication technologies in their information system in order to implement governance and democratization of having information.

In this case, the Moroccan governments through the E-Governance strategy [2] move forward in the field of information technology and encourage all establishments to adopt new tools for developing their information system.

This new strategy encourages public university to adopt information and communication technology tools like Data Warehouse to facilitate communicating information around the functional processes in the public university and to improve their performance especially in scientific research.

An information system, including a data warehouse system, is user-interfaced and designed to provide information useful to support strategy, operations, management analysis, and decision-making functions in an organization.

In this work, we contribute to the modeling and the implementation of management information decision support system in a public university especially in Scientific Research.

The paper has five parts. First section we propose a short recall of the scientific research in public university. Second section shows the relevant literature review of Data Warehouse. In the third section we introduce the context of study. The fourth section shows the case study. The fifth section of the paper shows the Pairing between Data Warehouse and Data Mining.

2. Scientific Research in Public University

Research is therefore not conducive to control and management. However, the rapid evolution of the highly competitive world of higher education today imposes constraints that require the establishment of a minimum management framework.

For the university nowadays to take risks is an essential aspect of the dynamism of the institution, but the risk must also be understood and managed.

Five ways in which management integrates culture into the university have been identified [3]:

- Strengthen the central steering structure.
- Develop peripheral activities.
- Diversify sources of funding.
- Mobilize academic skills.
- Integrate a corporate culture.

The study departments are the stones on which universities build their success, and the structures that link these departments directly to the center of the university, without intermediaries, shorten lines of communication and accelerate decision-making.

Good governance [4] contributes to the success of the institution when the external elements involved, the administrative staff and the scientific community work closely together. On the other hand, progress will be hampered if one of these elements takes over.

Research Management Studies [5]defines twelve characteristics of a productive research environment:

- Clear objectives that have a coordinating function.
- Emphasis on research.
- A culture of specific research.
- A positive group climate.
- Strong participatory governance.
- A decentralized organization.
- Frequent communication.
- Accessible resources (especially human resources).
- The size, age and diversity of the research group must be sufficient.
- Appropriate rewards.
- High emphasis on recruitment and selection.
- Management with the necessary experience and expertise in the field of research to put in place the appropriate organizational structures and apply participatory management methods.

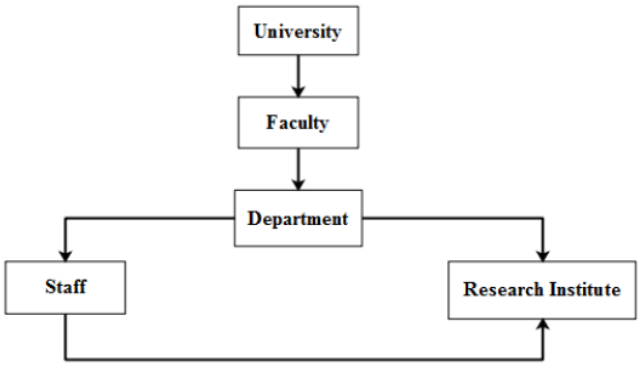

Fig. 1 Scientific Research Architecture

Among the main outcome indicators used in research universities were the following:

- Measures of resources: Financial resources.
- Number of researchers.
- Number and percentage of current researchers.
- Applications for research funding.
- Number of publications (by origin, for example, journals publishing articles submitted for examination by specialists).
- Quotes.
- Directed theses (completed and supported).
- Research applications (patents, licenses).
- Academic awards (editorial positions, Special rewards).

Scientific research competence today is important for many professions and activities: it is necessary not only to creatively apply the obtained knowledge but also to create new knowledge, to carry out the applied researches. Also, scientific research activity is the basic component of developing science education.

The purpose of the research is to describe the current situation of organization and realization of scientific research activity, to define essential factors promoting and hindering students' interest in scientific research activity, to determine lecturers' competence peculiarities in the sphere of organization and realization of scientific research.

In the first study years, the students should create reports, present works, raise problems and propose various problem solution variants; the most important criterion was the competence of experts and current research activities (scientific publications, participation in the national and international projects etc.).

The PhD students have possibilities to participate in seminars, projects, conferences, that lecturers willingly help the students to choose the research subject that interests them.

The participation in conference is great. On the other hand, it is obvious, that the main subject in scientific research activity is the student and sufficient lecturer's contribution promoting this activity; scientific research activity requires consistency, diligence, creativity.

Despite the financial difficulties, the institutions must find the possibilities and form conditions for professional improvement.

Between the most important recommendations:

- Apply modern study methods, promoting critical thinking and new subject search.
- Students should be more involved in the performance of the lecturers' research works, as assistants, putting data in order.
- Develop lecturers and students' team work.
- Include in study programs more subjects for education of research competences.
- Prepare more complex science projects, in which students could participate.

3. LITERATURE REVIEW OF DATA WARHOUSE

Nowadays, decision-makers need a synthetic and global view of the information circulating in their organization in order to guide and adapt their decision-making. To facilitate this process, they use decision support systems. These tools allow decision-makers to have quick and interactive access to a set of data to gain a global view of the activities of a given establishment [6]. The basic concept of a Data Warehouse evolved over the past 20 years.

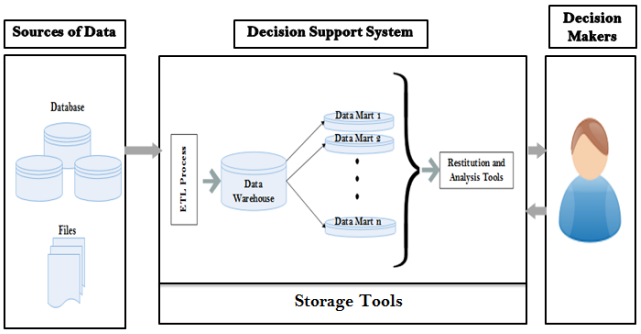

Fig. 2 Architecture of a decision support system

- **Data Warehouse:**

A Data Warehouse is the centralized storage space of an extract of relevant data sources for decision makers. Its organization must facilitate management data in the form of a unified vision and should allow for the conservation of necessary developments for decision-making [7].

- **Data Mart**

A data store is an extract from the warehouse suitable for a class of decision makers or for a particular purpose and organized according to a model adapted to decision-making processes [8].

- **Extraction, Transformation and Loading tools**[9]**:**

Defines a standardized procedure for each Data Warehouse; Extract signifies the connections and rules that the process has to follow to draw data from different sources; validation, completion, and standardization of data need to be done to transform into Data Warehouse compliant structure.

✓ Extraction

Is the first step in the process of providing data to the Data Warehouse. Extracting means reading and interpreting the source data and copying it to the preparation area for further manipulation.

✓ Transformation

It is a series of operations that aims to make the target data homogeneous and can be processed in a consistent way [10].

✓ Loading

This is the operation of loading the cleaned and prepared data into the Data Warehouse[11].

- **Analysis and Restitution of Data:**

Data from a Multi-Dimensional Database is queried using OLAP technology using tools graphs or in a textual language.

- **Online Analytical Processing:**

An OLAP system [12] is defined as a decision-making system in which data stores follow a multidimensional organization of data so provide effective support for OLAP analysis.

- **OLAP Data Modeling:**

The aim of multidimensional modeling is to organize data so that OLAP applications are effective and efficient [13], OLAP analyzes consist of following indicators considered as points observed in a space defined by different axes of analysis. Two approaches exist:

- **Modeling in cube:**

Among the main proposed models that speaks of cube, This vision manifests itself through a separation between the values and the structure of the elements, The data cube[14] is formed by observation axes of indicators placed in the cells, for each observation axis, a graduation is chosen to observe the data at an adequate level of granularity, but among the negative point on this type of modeling, more than three axis of analysis meets representation problems multidimensional spaces.

- **Multidimensional modeling:**

In answering the problem manifested by cube modeling, other approaches try to overcome these limits based on a set of concepts such us: dimension, fact, hierarchy…; but this modeling suffers from the absence of a standardized consensus on this formalism, Axes of analysis represented by the dimensions the analyzed variables represented by the indicators.

- **OLAP manipulation operations**

You cannot find an operation set that provides all of the OLAP manipulation operations, but most of the proposals provide partial support for different categories of operations.

Amongst these different operations, the most important are the rotation and drilling operations which are based directly on the metaphor of the cube. Another operation can be found on a scientific literature and the many existing software.

✓ Drilling operations

Drilling operations [15] make it possible to analyze with more or less precision an indicator based on the hierarchical structure of the analysis axes.

The "roll-up" is to analyze the data according to a level of less detailed granularity as opposed to drilling down ("drill-down"), which allows analyzing the data with a finer level in granularity.

✓ Rotation operations

The rotation operations consist in most of the time to change the axis of analysis in use it is called dimension rotation.

The second case of rotation consists in changing the subject of the analysis into a constellation it is a rotation of the fact.

The third case of the rotation consists in maintaining the same axis of analysis but one change in the graduation it is a rotation in the hierarchy.

✓ Restriction operations

Restriction operations [16] can restrict the set analyzed data. ("slice") is to express a restriction on one of the data of one of the analysis axes by specifying a given cube slice. The specification of a sub-cube ("dice") is to express a restriction on data from an analysis indicator.

✓ Transformation operations

It adds a dimension attribute in as an indicator of analysis ("push") or it converted an indicator of analysis into parameter ("pull").

✓ Scheduling operations

"switch" allows changing the position of the values of the parameters of the dimensions or to reorder the parameters of a hierarchy "nest" makes it possible to reorder the parameters of a hierarchy. We nest an attribute in another hierarchy.

4. THE CONTEXT OF STUDY

In this context, even Moroccan public universities are becoming more dependent on using information communication technologies tools for developing, communicating information around the functional processes, especially in the scientific

research section. The aim of our research study is to develop a Data Warehouse related to scientific research.

Universities use Information and communications technology for academic purposes that differ from other organizations because they have different environments and circumstances. In order to make more informed decisions about a scientific research department, the ultimate goal for our system is to facilitate the management of scientific research decision makers to get and find the relevant information they are looking for.

Although, universities are planning to renew their information systems in the future, this necessitates the call for more research efforts in this area. Because, the study concluded that Data Warehouse potentially improves services offered to decision makers and administrative staff to manage and take the right decision based on real and relevant data.

Change management [17] is a primary concern of many universities in terms of adopting a Data Warehouse, as activities, processes, and methodologies.

The aim is to develop a pertinent information system to support researchers and their scientific activities in the Moroccan public university.

- **Justification for the choice of Microsoft SQL Server**

SQL Server is an incredible database product that offers an excellent mix of performance, reliability, ease of administration, and new architectural options, yet enables the developer or DBA to control minute details. SQL Server is a dream system for a database developer.

The business intelligent SQL Server [18] give us the possibility to create a database that summarizes the OLTP data into new tables. Because it is still a refined set of data, it is often insufficient in performing analytics that are holistic to the entire organization.

The implementation of an automatic system for managing scientific research is based on SQL Server Software. The reasons for this choice are:

- Transform complex data.
- Modernize reporting.
- Flexible.
- Enable hybrid BI.

- **Restitution: Multidimensional OLAP Analysis**

Conceptually, Data Mart relies on multidimensional data modeling from the Data warehouse. The data is viewed in the form of dots in a multi-dimensional space with the data cube or hypercube metaphor. Each data represents a cell of the cube. The edges of the cube represent the axes of analysis of the data and include several graduations to allow observation of the data according to different levels of detail. This modeling allows the expression of online analyzes (OLAP) multidimensional.

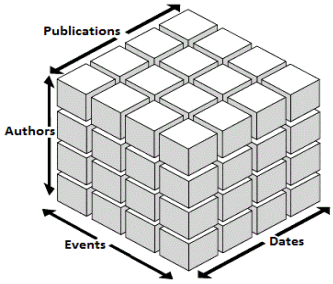

Fig. 3 Sample data cube

The interrogation of a multidimensional database is a succession of exploration operations. Following a multidimensional analysis of the technology used, or by OLAP analysis with reference to the type of data handled [19]. An OLAP analysis runs as follows:

- Selection of a first request
- Data Navigation by (drilling down, drilling up, data restriction, etc.)

By the notion of fact, we designate the object we want to analyze. In fact, one or more measures (indicators) are observed. A set of values is associated with each fact; these values are taken by the fact for each measure. These measurements are generally numerical.

A dimension offers the user points of view differentiated to analyze or observe facts and defines an axis of analysis. It consists of one or more attributes, called hierarchical levels. It corresponds to the levels of detail that can be observed on the facts. A level is composed of elements called members or modalities. The links between levels can have different cardinalities.

An OLAP cube represents the aggregated values of the measure in a multidimensional space defined by the user. he chooses the dimensions according to which he wants to analyze the facts and also for each dimensions the level the hierarchy on which he wants to work. The combination of the modalities of the selected dimensions designates the cells of the cube. The latter contains the value of the measure or measures corresponding to the combination modalities. This value of the measure can be either a detailed value or an aggregated value according to the hierarchical levels chosen.

5. CASE STUDY:

In this paper, we propose the study of publications of the research laboratory belonging to our public university for the purpose of analyzing scientific publications according to the main field of research.

On the other hand we propose the follow-up of the procedure (Harakiya, Ich'àae, Nadwa) carried out by the researchers of the university according to the axes of analysis Dates, the axis of analysis Events which can be of national or international level, the axis of analysis Authors according to the statutes of the authors professor or PhD student.

Ich'àae: This procedure is dedicated to the publication and dissemination of research works, support nature, information on the project leader, project description, etc.

Nadwa: This procedure is dedicated to the organization of scientific events, budgets, organizing committee, coordinator of the event, research entity organizing the event, etc.

Harakiya: This procedure is designed to manage the mobility of researchers and everything concerning their participation in scientific events and research internship.

The example in the figure 4 corresponds to the modeling of scientific research publication based on authors and keywords over time. The axes of the analysis are represented by the authors, the support, the keyword and the time dimensions. The dimension time is characterized by four parameters day, month, quarter, year organized hierarchically: the day parameter represents a finer graduation (to observe the precise date finer) than the graduation month, itself being a finer graduation of year. It is possible to group the month into quarter. The authors dimension contains two hierarchical levels in order to group the authors based on their status. The genre publication dimension contains the id of the work, the journal as well as the number of pages, the volume, … according to its national or international scope.

- Basic concepts:

- F all the facts,

- D all the dimensions,

- H all the hierarchies,

- M all measures,

- I the set of instances of dimensions,

- A the set of dimension attributes.

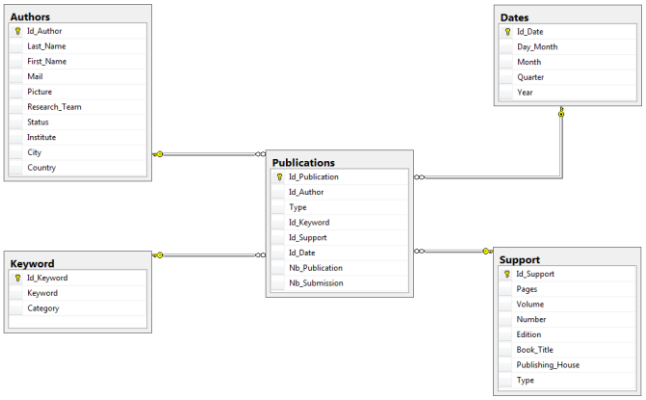

Fig. 4 Multidimensional modeling of publications

For example we can say that the dimension author D_1^{Authors} contains one level H_1^1Author, and the level H_1^1 contains several modalities between themes we find a_1^{11} Professor, a_2^{11} PhD student.

To count the number of publication in fact table, we use the aggregation functions in our case to count the number of publication related to Author x we use COUNT this function counts the number of instances in an aggregate.

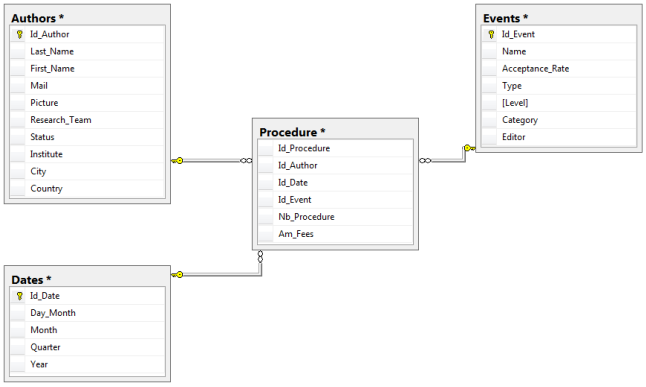

Fig. 5 Multidimensional modeling of procedures

For this case, we can say that the dimension D_3^{Events} contains one level H_3^1 Event; the level H_3^1 contains two modalities a_3^{11} national, a_3^{12} international.

To count the number of procedure in fact table, we use the Aggregation functions in our case to count the number of procedure related to laboratory **x** we use COUNT this function counts the number of instances in an aggregate.

A fact and its associated dimensions make up a schema star. One possible generalization is to describe a "constellation of stars" consisting of several facts and several dimensions possibly shared forming a constellation pattern.

A generalization from the two preceding figure consists in proposing a constellation model [10] which groups together a set of facts associated with dimensions which are provided with multiple hierarchies.

A constellation is defined by $(N^{SR}, F^{SR}, D^{SR}, Star^{SR})$ where:

- N^{SR} is the name of the constellation,

- $F^{SR} = \{F_1, \dots\dots\dots\dots\dots\dots F_n\}$ is a set of facts,

- $D^{SR} = \{D_1, \dots\dots\dots\dots\dots\dots D_m\}$ is a set of dimensions,

- $Star^{SR}: F^{SR} \rightarrow 2^{D^{SR}}$ is a function that combines facts with dimensions.

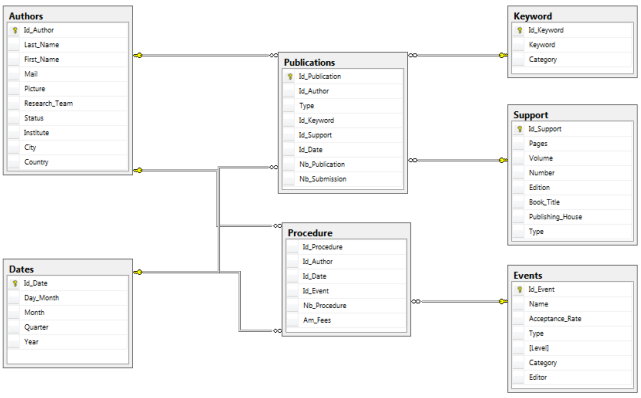

Fig. 6 Constellation schema of the multidimensional database

To simplify we designate the concept of fact F_i, respectively of dimension D_i and Hierarchy H_j, by its name N^{F_i}, respectively N^{D_i} and N^{H_j}, our formal description of the schema in constellation is as follows:

This constellation scientific research (SR) has two facts and five dimensions. It is defined by $(N^{SR}, F^{SR}, D^{SR}, Star^{SR})$ where:

N^{SR}= ' Scientific Research ';

$F^{SR} = \{F^{Publication}, F^{Procedure}\}$;

$D^{SR} = \{D^{Authors}, D^{Keyword}, D^{Dates}, D^{Support}, D^{Events}\}$;

$Star^{SR}(F^{Publication}) \rightarrow \{D^{Authors}, D^{Keyword}, D^{Dates}, D^{Support}\}$;

$Star^{SR}(F^{Procedure}) \rightarrow \{D^{Authors}, D^{Dates}, D^{Events}\}$;

The procedures carried out by researchers of the doctoral center can be studied according to the fact:

$F^{Procedure} =$
$(N^{Procedure}, F^{Procedure}, D^{Procedure}, Star^{Procedure})$ where:

-$N^{F^{Procedure}} = ' Procedure'$;

-$M^{F^{Procedure}} = \{Nb\_Procedure, Am\_Fees\}$;

-$I^{F^{Procedure}} = \{i_1^{Procedure}, \ldots., i_n^{Procedure}\}$;

-$IStar^{F^{Procedure}} = \{i_k^{Procedure} \rightarrow \{i^{Authors}, i^{Events}, i^{Dates}\} | k \in [1..y], I_k{}^{Procedure} \in I^{F^{Procedure}} \wedge \exists i_{k_i}^{Authors} \in I^{D^{Authors}} \wedge \exists i_{k_j}^{Events} \in I^{D^{Events}} \wedge \exists i_{k_p}^{Dates} \in I^{D^{Dates}}\}$.

6. FROM DATA WAREHOUSE TO DATA MINING

In the previous section, we talked about the Data Warehouse its conception and its development, in this section we try to introduce some algorithms of the Data Mining [20] in order to better explore the existing data and to offer to the decision-maker a more realistic view in making predictions based on existent data in the Data Warehouse, for that we had to choose the datamining algorithm that produces results with the smallest possible error rate.

Data mining, contrary to the data warehouse, tries to find the independence and relationships that exist between the data based on a set of methods of adequate mathematical and statistical analysis.

Several models have been implemented in this project. We are introducing some of them in order to be practical, in this example we try to evaluate the number of endowment provided by the university council to the researcher in order to participate in national or international conference (HARAKIYA) according to the following criteria:

-For the presentation of a paper at a scientific event, only one person (student or teacher-researcher) can receive support from the University.

-PhD students will benefit from this support only from the 2nd year for internships and participation in a scientific event in Morocco and abroad.

-Each doctoral student can benefit from an annual support of 9000 DH abroad and 2000 DH in Morocco, after the opinion of commission of scientific research of the USMS.

-Teachers and PhD students who have not submitted the report on the supported action by the deadline cannot benefit from any support from the University the following year.

-Only teachers and doctoral students belonging to research entities (laboratory, team, center) recognized by the university council can benefit from this support. Otherwise, a certificate must be drawn up by the head of the institution showing the opportunity and the repercussions of such an activity for research in the establishment.

The first example is evaluation of the number of endowment provided by the university council to the researchers based on the following criteria: Year Thesis, Supported this Year, Opinion of the Head of the Research Entity, Gender. We choose to work with Naïve Bayes algorithm, the reasons for this choice are:

1. For each hypothesis: We associate a probability observation of one or several instances may change this probability.

2. We can talk about the most hypotheses likely, based on the conditional probabilities and Bayes rule.

3. Forecasting the future from the past, while assume independence attributes.

4. Bayesian probability is the estimation of an event knowing a preliminary hypothesis is verified (knowledge).

Microsoft Naive Bayes [21], chooses Year Thesis as the most significant attribute for the first level. After that the algorithm chooses Opinion of the Head of the Research Entity and Supported this Year as the next attributes by importance and finally the Gender. At the end, there are evaluations of expected endowment depending on the values from analyzed attributes.

Input data will be randomly split into two sets, a training set and a testing set, with percentage of data for testing around 30%. The training set is used to create the mining model. The testing set is used to check model accuracy.

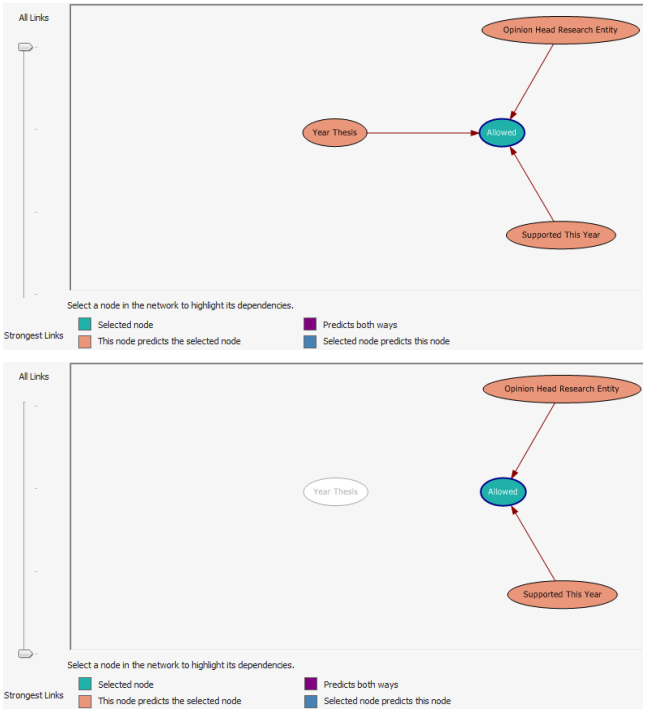

Fig. 7: Analysis of correlation between criterions for evaluation of the number of care provided by the university council.

Figure7 represents correlation between criteria for evaluation of the number of endowment provided by the university council. By moving cursor to the left-hand side of the Figures, the intensity of correlation between observed criterions is displayed. Therefore it is easily perceived that the obtained results depends on Year Thesis the least, the most on both Supported this Year, and Opinion of the Head of the Research Entity.

The second example discusses the grants giving to the laboratories based on the following criteria: Number Registered, Number Publication, and Number Event. Microsoft Naive Bayes chooses Number Publication as the most significant attribute for the first level. After that the algorithm chooses Number Registered as the next attributes by importance and finally the Number Event. At the end, there are evaluations of expected grants depending on the values from analyzed attributes.

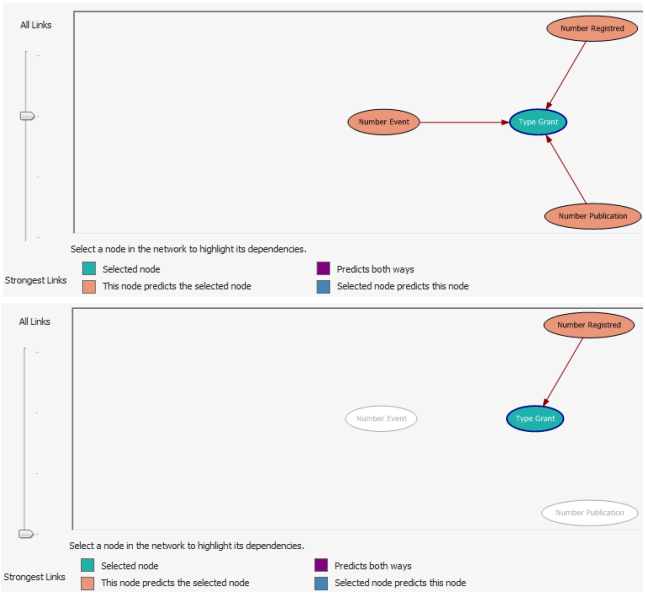

Fig. 8: Analysis of correlation between criterions for evaluation of the grants giving to the laboratories.

Figure8 represents correlation between criteria for evaluation of the grants giving to the laboratories. By moving cursor to the left-hand side of the figures, the intensity of correlation between observed criterions is displayed. Therefore it is easily perceived that the obtained results depends on Number Event the least, a bit more on Number Publication, the most on the Number Registered.

In our case, Now we want to decide situation when the Number of new Registered in the laboratory in the previous year is 5, the Number of Publications in the previous year is 14 and the Number of Events organized is 2 which is not available into our trained dataset.

Our example gets classified as 'Medium' about the type of grant giving to the laboratory.

To conclude, on the one hand the use of the data warehouse is a flexible solution for decision-makers in universities and on the other hand it is adapted to the administrative staff in the university because it can used daily tool for explore Data Warehouse like Microsoft Excel unlike other Olap tool.

Another advantage manifests through research on knowledge and information exciting in the Data Warehouse without having deep knowledge on the complex query languages.

Data Mining allows the user to analyze a large number of data by offering the possibility of the prediction at the requested time; the latter offers the decision-maker the opportunity to explore the current data and to have an idea about the future behavior of the institution. This marriage of existing data explorations and prediction of future behavior helps decision makers in the process of solving the problems.

7. Conclusion

The new orientation by the government by adopting the governance in all aspects of institutional Moroccan life leads to the change of the way of employing information system in this sector because the affect administrative staff, the quality of services, and the time of respond, etc.

Scientific research represents one of the axes that need modernization and automatization by adopting and implementing new solution like Data Warehouse .In this direction, this paper has examined the needs presented by decision-makers and the administrative staff and present as solution the implementation of Data Warehouse and Data Mining.

we demonstrated through the use of the Data Warehouse and Data Mining on a corpus of data relative to the university and especially on the scientific research without preliminary need of the deep knowledge on the algorithm of datamining nor on the language of request in a single and ultimate goal of optimizing the decision-making process for solving the problems encountered.

8. REFERENCES

[1] V. Lamanauskas and D. Augienė, "Development of Scientific Research Activity in University: A Position of the Experts," *Procedia - Soc. Behav. Sci.*, vol. 167, no. Supplement C, pp. 131–140, Jan. 2015.

[2] H. Lamharhar, D. Chiadmi, and L. Benhlima, "Moroccan e-Government Strategy and Semantic Technology," in *Government e-Strategic Planning and Management*, Springer, New York, NY, 2014, pp. 323–343.

[3] R. Adams, S. Martin, and K. Boom, "University culture and sustainability: Designing and implementing an enabling framework," *J. Clean. Prod.*, vol. 171, no. Supplement C, pp. 434–445, Jan. 2018.

[4] M. Gross and R. Źróbek, "Good governance in some public real estate management systems," *Land Use Policy*, vol. 49, no. Supplement C, pp. 352–364, Dec. 2015.

[5] M. N. K. Saunders and F. Bezzina, "Reflections on conceptions of research methodology among management academics," *Eur. Manag. J.*, vol. 33, no. 5, pp. 297–304, Oct. 2015.

[6] G. Colliat, "OLAP, Relational, and Multidimensional Database Systems," SIGMOD Rec, vol. 25, no. 3, pp. 64–69, Sep. 1996.

[7] "Wiley: Building the Data Warehouse, 4th Edition - W. H. Inmon." [Online]. Available: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0764599445.html.

[8] W. H. Inmon and D. Linstedt, "3.4 - Data Marts," in *Data Architecture: a Primer for the Data Scientist*, Boston: Morgan Kaufmann, 2015, pp. 115–119.
[9] A. Reeve, "Chapter 6 - Extract, Transform, and Load," in *Managing Data in Motion*, Boston: Morgan Kaufmann, 2013, pp. 29–35.
[10] S. H. A. El-Sappagh, A. M. A. Hendawi, and A. H. El Bastawissy, "A proposed model for data warehouse ETL processes," J. King Saud Univ. - Comput. Inf. Sci., vol. 23, no. 2, pp. 91–104, Jul. 2011.
[11] M. Bortolini, M. Faccio, E. Ferrari, M. Gamberi, and F. Pilati, "Time and energy optimal unit-load assignment for automatic S/R warehouses," Int. J. Prod. Econ., vol. 190, no. Supplement C, pp. 133–145, Aug. 2017.
[12] A. V. Korobko and T. G. Penkova, "On-line analytical processing based on formal concept analysis," Procedia Comput. Sci., vol. 1, no. 1, pp. 2311–2317, May 2010.
[13] D. Park, J. Yu, J.-S. Park, and M.-S. Kim, "NetCube: a comprehensive network traffic analysis model based on multidimensional OLAP data cube," Int. J. Netw. Manag., vol. 23, no. 2, pp. 101–118, Mar. 2013.
[14] C. Ciferri, R. Ciferri, L. Gómez, M. Schneider, A. Vaisman, and E. Zimányi, "Cube algebra: A generic user-centric model and query language for OLAP cubes," Int. J. Data Warehouse. Min. IJDWM, vol. 9, no. 2, pp. 39–65, 2013.
[15] J. Li and B. Xu, "ETL tool research and implementation based on drilling data warehouse," in 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery, 2010, vol. 6, pp. 2567–2569.
[16] Fuzzy Data Warehousing for Performance Measurement - Concept and | Daniel Fasel | Springer.2014.
[17] A. Matei and C. Antonie, "The Need for Positive Change: Adapting Management in Public Administration," Procedia Econ. Finance, vol. 26, no. Supplement C, pp. 345–350, Jan. 2015.
[18] Adam Jorgensen , Jorge Segarra , Patrick LeBlanc , Jose Chinchilla , Aaron Nelson, Microsoft SQL Server 2012 Bible, Wiley Publishing, 2012.
[19] S. Naouali and S. B. Salem, "Towards reducing the multidimensionality of OLAP cubes using the Evolutionary Algorithms and Factor Analysis Methods," *ArXiv160204613 Cs*, Feb. 2016.
[20] Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques. Elsevier.
[21] C. L. Curotto and N. F. F. Ebecken, Implementing Data Mining Algorithms in Microsoft SQL Server. Southampton: WIT Press / Computational Mechanics, 2005.