MEASURING THE PERFORMANCE OF EDUCATIONAL ENTITIES WITH A DATA WAREHOUSE

Mihai Păunică[1]
Marian Liviu Matac[2]
Alexandru Lucian Manole[3]
Cătălina Motofei[4]

ABSTRACT: This paper attempts to outline the benefits of using a data warehouse for supporting the management decisions in a pre-academic learning institution. One of the aspects is the measurement of the scholar performance of the pupils, correlated with the financial efforts made by the school management. The performance is analyzed from a multi-dimensional perspective, which allows the extraction of high quality information.

Key words: data warehouse, education, analysis, performance, multidimensional

JEL codes: I21, I22, C82, C88

Introduction

The information system used to design, allocate and use the monetary resources for the pre-academic public educational system involves actors such as the educational institutions, the local public administration, and the educational management structures at local and national level (the inspectorates and the Ministry of Education, Research, Youth and Sport).

The measurement of scholar performance is a key preoccupation of the management and of the entire educational personnel also. The next step is the evaluation of the correlations between the results (educational performance) and the financial allocations. One of the main indicator classes that describe the financial efforts made for education is made of the standard cost and its associate indicators. The level of the standard cost is a reference measure for the orientation of any decision-maker, and also a basis of comparison, on time, space, educational domain criteria.

The software system described in the paper offers the possibility of analysis for the results of the educational process. The "output" of the system presents, on one hand, the results achieved by pupils, and, on the other hand, a set of derivate indicators, that aim to correlate the results and financial indicators.

Literature review

The main aspects of database design were outlined by Schneider (2008). The need for quality and performance measurement in the educational world was emphasized by Stanciu, Florin, Rădulescu and Aleca (2009); they have studied the case of the higher education management. The aspects concerning the structure and construction of data warehouse were described by Inmon (2005), Popa, Udrică, Manole Vasilciuc, Gârba (2006). The instruments used for analysis, such as the MDX language, the Reporting Services were presented by Harinath and Quinn (2006) and Mundy, Kimball and Thornthwaite (2006).

---

[1] Academy of Economic Studies, Bucharest, Pta. Romana no.6, mihai63p@yahoo.com
[2] Academy of Economic Studies, Bucharest, Pta. Romana no.6, livius77@yahoo.com
[3] Artifex University, Bucharest, Economu Cezarescu no. 47 Street, alexandru.manole@gmail.com
[4] Academy of Economic Studies, Bucharest, Pta. Romana no.6, cmotofei@yahoo.com

Research methodology

The research started from the premise that the management of the educational entity is interested by the results of the educational process and the efficiency of allocated funds. The complexity of data necessary for analysis suggests that the best solution is a business intelligence system, organized around a data warehouse.

The software system designed includes the following components:

a) the data sources, made of relational databases and worksheets;

b) the data staging area, a SQL Server relational database;

c) the data warehouse, implemented in SQL Server;

d) the reporting component.

The model proposed for data sources has been presented in a previous article (Păunică, Țurlea, Matac, Manole, Motofei 2010). The data sources designed are:

- a relational database, that records the scholar achievements of the pupils (the ones recorded in the catalogue and the results at competitions), the evaluation of the teachers (as a measure of interaction between the main actors of the educational process);

- a set of worksheets, that refer standardized forms used by the educational entity (i.e. the budgetary execution).

The data staging area proposed for the data warehouse is represented by a relational database, under SQL Server. The advantages of using a relational database are evident, the most important of them are:

- SQL Server offers instruments that allow the import of data from the platforms on which the data sources exist;

- stored procedures or action queries can be used to clean and aggregate data;

- the existence of mechanisms for the configuration of primary and foreign keys, with strict compliance towards the restrictions for these fields.

The main functions of the data staging area are the creation and update of the time dimensional tables and the fact tables. These functions are based on make-table (SELECT ... INTO) and append (INSERT ... INTO) queries.

The data are imported from the data sources in the database (data staging area), the transformation operations are performed and then the structure of the data warehouse is prepared. The single data source for the data warehouse has several advantages, amongst them we outline the uniqueness of records, as the tables that will migrate to the data warehouse will include primary and foreign key restrictions to assure the integrity of data.

The structure of the software system is presented in figure no. 1a.
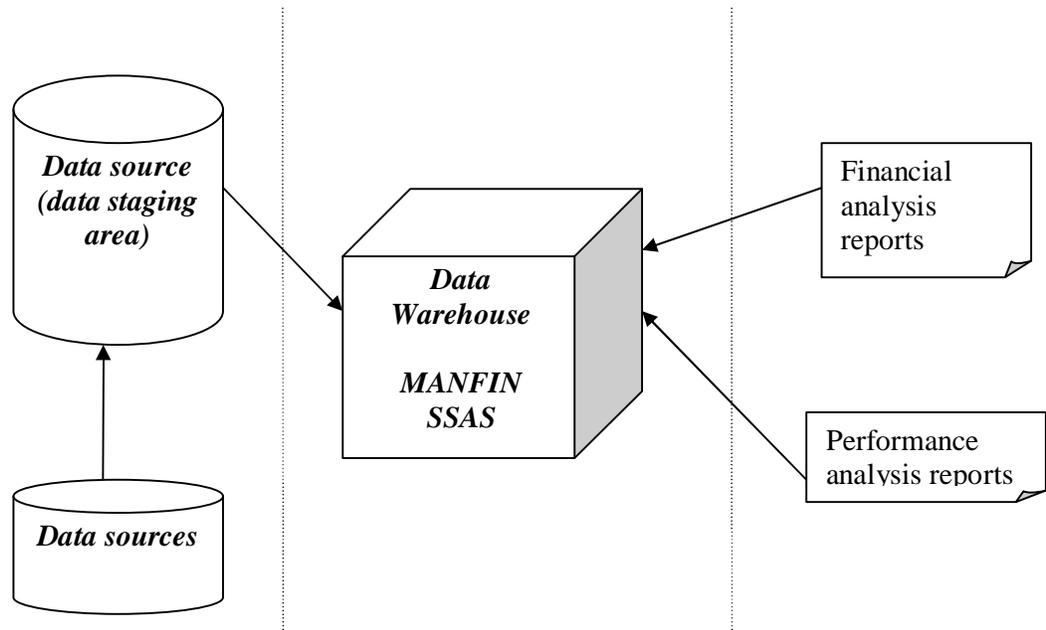
Fig. no. 1a – The MANFIN data warehouse - the core of the BI system

The overall structure of the data warehouse is presented in figure no. 1b. Likewise, the structure was presented and described in a previous work by Păunică, □urlea, Radu, Ciocoiu, Manole (2010).
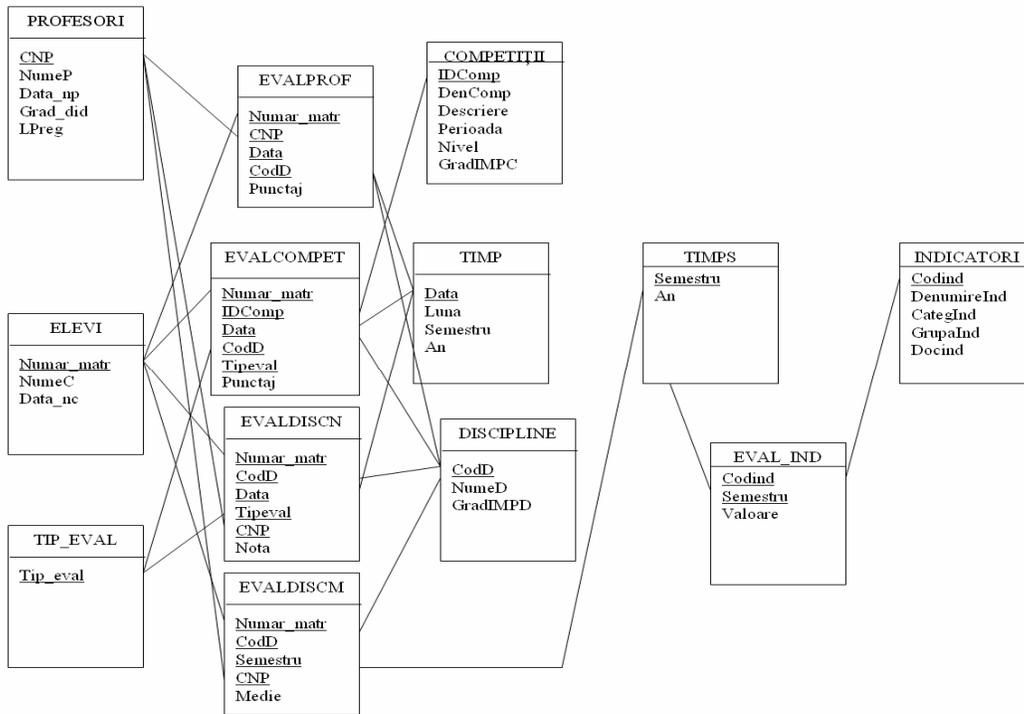
Fig. no. 1b – The MANFIN data warehouse - the conceptual model

The data warehouse reflects the following components:

- a set of dimensions that describe the actors involved in the educational process, the curricula and the scholar competitions;
- a set of fact tables that include the measures desired for analysis (the evaluation of teachers by their subjects, the results at competitions, the total and average scores obtained by pupils);
- a subset of tables describing the financial indicators involved in the analysis.

The creation of the two time tables is justified by the two levels of aggregation for the measures: on one hand, we attempt to include the notes and competition-related results (that are characterized by a certain randomness in time), on the other hand, we aim to offer the possibility to compare aggregate data (semester-based aggregation). This set of functional objectives explains the set of fact tables, along with the different dimensions that describe the measures included in the model.

Therefore, the model of the data warehouse complies with the constellation schema, a model template that, despite complexity, allows the share of dimensions between multiple fact tables and the implementation of all aggregation levels necessary for the users.

The physical model of the data warehouse is described in figure no.2.
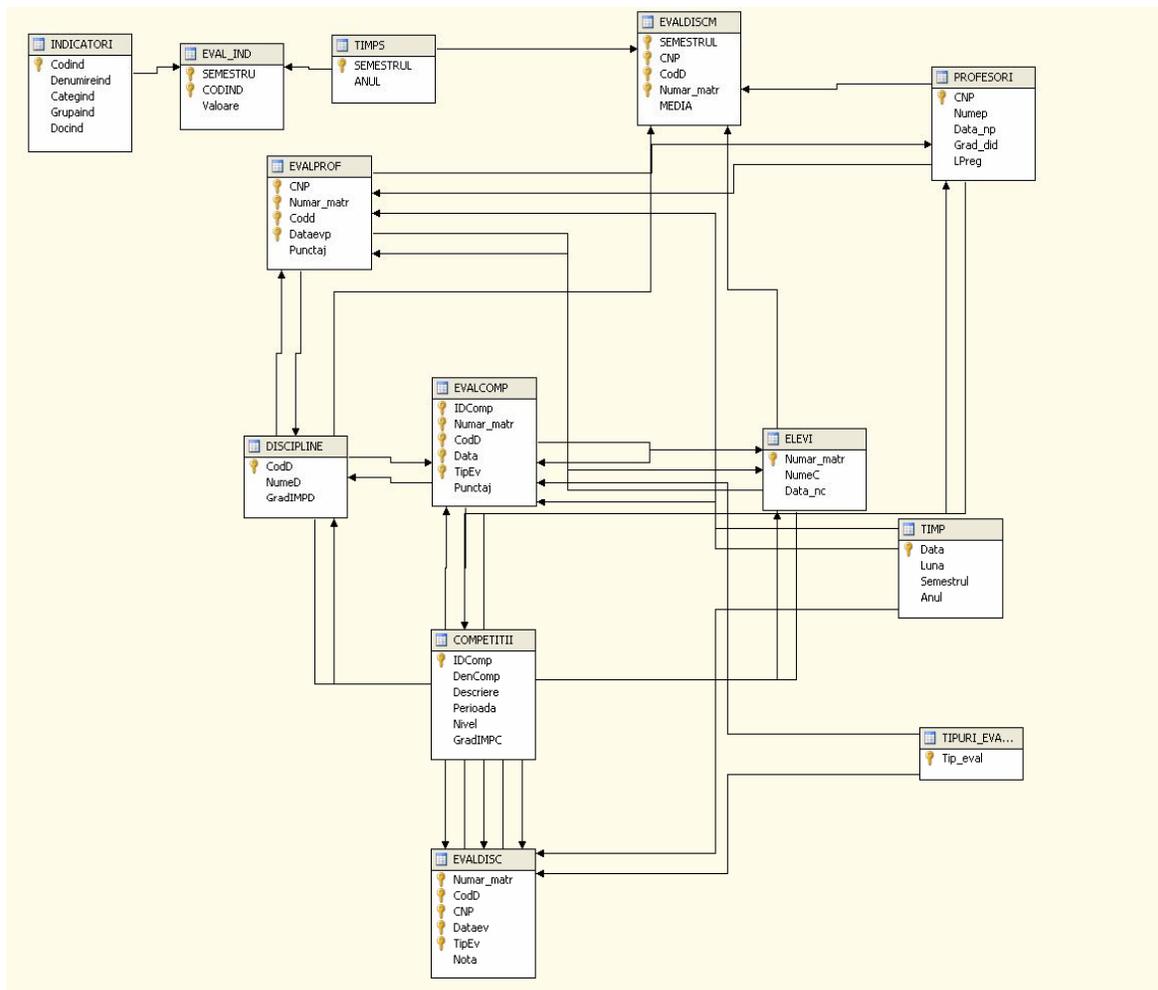
**Fig. no. 2 – The MANFIN data warehouse - the physical model**

The data loading mechanisms are represented by a set of INSERT...INTO queries, that perform the following functions:

- the update of the time dimension: involves the insertion of new members from all sources (scholar evaluation - notes, teacher evaluation, results from competitions). The other time dimension is updated by using an INSERT ... INTO query based on the first dimension. In order to assure the integrity of the operation, all the queries were merged into a stored procedure.

- the update of the fact tables.

The other dimensions extract their data directly from the corresponding tables in the data staging database.

The data processing mechanisms aim to transform the data into useful information for the decision makers. The analysis follows several directions.

First, we analyzed the notes achieved by pupils during the semester. By using the SQL Server Analysis Services, the cube presented in the following figure was created:
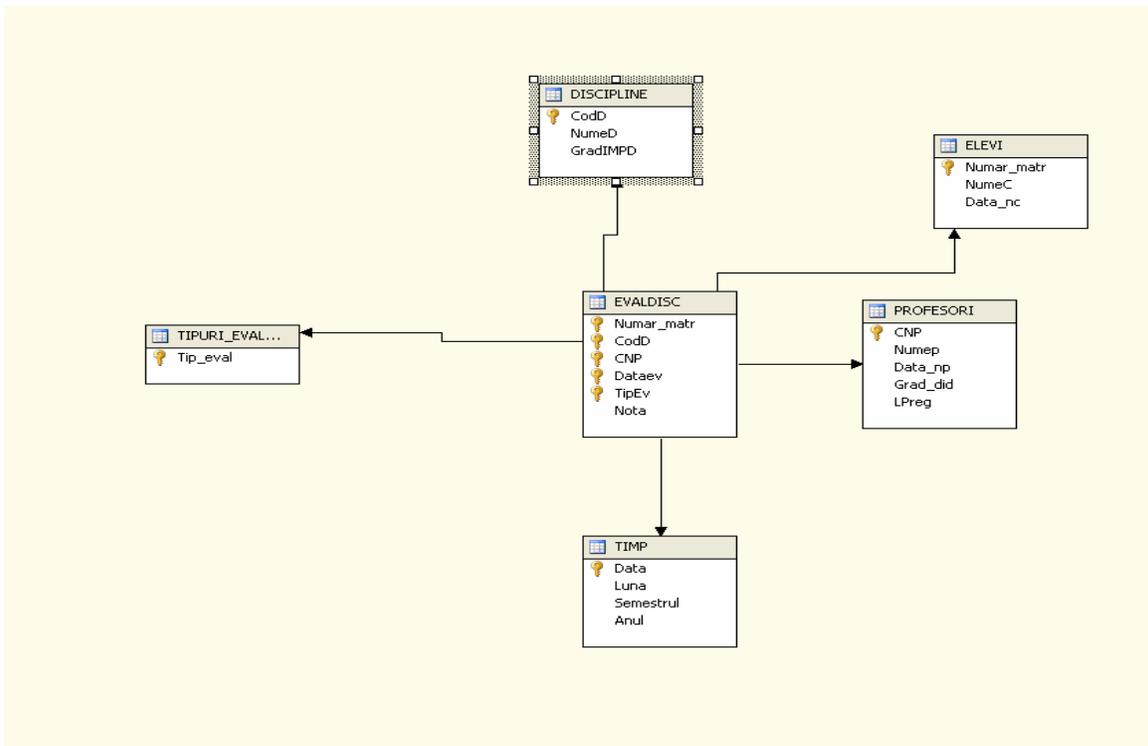

Fig. no. 3 – The *cubEVALUARE_NOTE* - the physical model

The structure of the cube allows for the analysis of the results achieved during regular class time, offering the option to analyze every dimension and aggregation level.

The first report indicates the total score obtained, grouped by disciplines, a score calculated based on the notes achieved. The disciplines can be grouped according to their degree of importance (marked by the GradIMPD attribute in the dimension). This report offers high-level information about the level of knowledge achieved by children at that discipline.

The MDX query for the report is presented in the figure no.4:
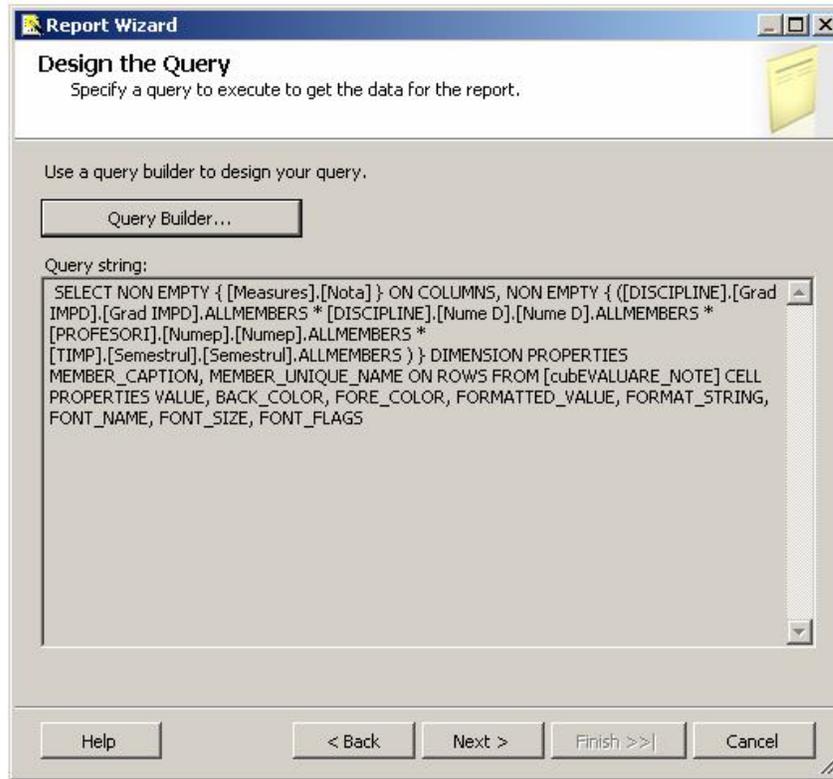
Fig. no. 4 – The MDX Query, taking into account the importance of the discipline

The structure of the report is presented in the following figure:

			Semester I, 2008-2009	Semester II, 2008-2009
Very high	Physics	Voican Amalia	334	329
	Romanian	Petre Camelia	324	325
	Mathematics	Zamfir Antonia	332	335
Average	Painting	Fratila Loredana	332	326
	Geography	Zamfir Antonia	334	329
	Applied Informatics	Petrescu Eliza	330	325
	History	Petrescu Eliza	335	322
		Zamfir Antonia	328	331

Fig. no. 5 – The Report, taking into account the importance of the discipline

The system also offers further analysis, by comparing the results of the pupils in function of the level of scientific achievement of the teachers, analysis of the results for every pupil and making them available for the respective teachers, in order to improve the scholar achievement of the children.

As previously stated, the analysis is not complete without taking into account the correlation between the results and the financial indicators. In order to perform this, we have defined a system of derivate indicators, that include both the financial efforts and educational results.

The main indicator designed is the *"educational efficiency of the standard cost"*, calculated as a report between the total standard cost and the conventional score (points) achieved by the pupils during the reference period (an average standard cost). The conventional score is determined as a sum of the average score from the catalogues.

The calculation formula of the indicator is:

$$Ics = \frac{Ch_s}{Pcm} \tag{1}$$

where:

Ics - index of educational efficiency of the standard cost
Chs - Cost standard, total
Pcm - Conventional score

By taking into account the components of the standard cost, the indicator can be further analyzed according to the formula:

$$Ics = \frac{Ch_p}{Pcm} + \frac{Ch_{ms}}{Pcm} + \frac{Ch_{perf}}{Pcm} + \frac{Ch_e}{Pcm} \tag{2}$$

where:

Ch_p - personnel expenses

Ch_{ms} - material related expenses (utilities)

Ch_{perf} - personnel training expenses

Ch_e - other material expenses

The analysis is performed by defining a new cube, with the structure presented in the following figure:
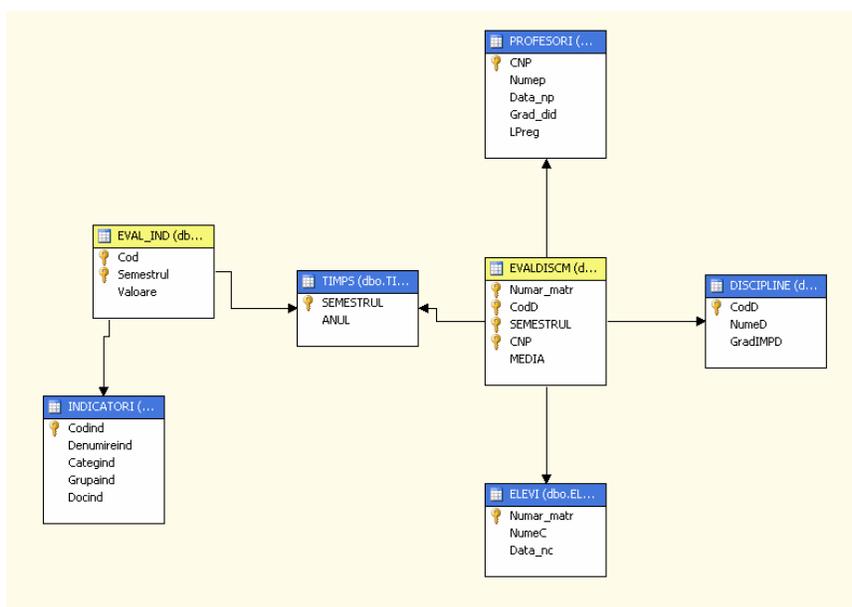
Fig. no. 6 – The *cubEVALUARE MEDII* cube, the physical model

The structure of the cube allows the analysis of the index, as a total and by factors of influence, with the help of an additive analysis model based on formula (2). The drilldown instrument helps in easing the presentation of information. The influence of any component of the standard cost can be observed and measured. By further detailing the structure of the standard cost, more factors of influence are obtained and the granularity level for the analysis increases.

The structure of the report (based on high granularity) is presented in the following figure:

			Semester I, 2008-2009	Semester II, 2008-2009
Index of educational efficiency	⊟ Personnel training cost	Personnel training expenses	11.64	11.27
	⊟ Standard cost for utility expenses	Water, severage, sanitation expenses	4.07	6.54
		Lighting expenses	27.63	39.81
		Heating expenses	43.58	67.79
	⊟ Standard cost for personnel expenses	Personnel expenses for the movement of employees	2.34	6.83
		Salarial rights for the teaching staff	715.78	1034.43
		Salarial rights for the auxiliary teaching staff	159.06	192.11
		Salarial rights for non-teaching personnel	119.3	251.22
		Budget transfers	35.52	54.31
	⊟ Other expenses - materials/services	Other expenses - materials/services	45.62	133.97

Fig. no. 7 – The "Index of educational efficiency of the standard cost" - analysis report

Conclusions

The business intelligence systems are an important component in contemporary decision support environment. The described data warehouse solves some of the issues related to the performance in pre-academic education. The structure of the data warehouse allows for multiple analysis fronts, such as the educational performance (scores, average scores, competitions, performance of the teachers), the efficiency of the costs (measured with the use of a system of indicators derived from the standard cost - an item widely used in educational environment).

References
1. Harinath, S., Quinn, S. R., 2006. Professional SQL Server Analysis Services 2005 with MDX, Wiley Publishing, Inc., Indianapolis
2. Inmon, William H., 2005. Building the Data Warehouse, Fourth Edition, Wiley & Sons
3. Mundy, J., Kimball, R., Thornthwaite, M., 2006. The Microsoft Data Warehouse Toolkit With SQL Server 2005 and the Microsoft Business Intelligence Toolset, Wiley Publishing Inc. Indianapolis
4. Păunică, M., Ţurlea, C., Matac, L.M., Motofei, C., Manole, A., 2010. Software model for the financial management in pre-academic education-data sources, Romanian Statistical Review, Issue 3/2010 - Supplement
5. Păunică, M., Ţurlea, C., Radu, C., Ciocoiu, R., Manole, A., 2010. Model of a data warehouse for the educational management, Romanian Statistical Review, Issue 6 - Supplement
6. Popa, Gh., Udrică M., Manole A., Vasilciuc, B., Gârba, M., 2006. Microsoft SQL Server, Economica Publishing House, Bucharest

7. Schneider, M., 2008. A general model for the design of data warehouses, International Journal of Production Economics, Elsevier, Vol. 112 (2008), Issue 1 (March)
8. Stanciu, A., Florin, M., Rădulescu, C., Aleca,. O, 2009. Solutions for decision support in university management, Economia Journal, Management Series, Faculty of Management, Academy of Economic Studies, Bucharest, Romania, Volume (Year): 12 (2009) Issue (Month): 1 (June)
9. Fact constellation schema architecture, available at http://etl-tools.info/en/bi/datawarehouse_constellation-schema.htm, consulted October 1st, 2010