

Data Sources Integrator for Management by Indicators

Gustavo Illescas, Iris Figini, Mariano Martinez.

Research Institute in Advanced Information Technology, Computer Management- Centre Associate CIC Group. Faculty of Exact Sciences. National University in the centre of Buenos Aires Province. Paraje Arroyo Seco, Tandil (7000), Argentina.
illescas@exa.unicen.edu.ar, imfigini@slab.exa.unicen.edu.ar,
mmartinez@slab.exa.unicen.edu.ar.

Abstract. This current work exposes the research and development of a tool for the integration of data sources in the context of management by indicators. The main benefits of this research are the administrators of databases that will work with a powerful tool that will allow them to effectively manage the updating of indicators previously defined by information managers. These information managers are the most interested in that the indicators are updated in the databases for the taking of decisions which will directly benefit them. This work gives rise to a thesis in Systems Engineering and the result could be applied in the TIC direction of the Faculty of Exact Sciences as it is expressed in the case study, taking part of the Strategic Plan of Engineers Formation promoted by the Secretary of University Policies.

Keywords: Database, ETL process, Indicators, Management Control.

1 Introduction

Frequently the managerial sectors of organizations diagnose the need to improve some aspect of their management, which establishes a possible hypothesis that should be able to validate and execute the strategies and policies defined. To perform this task formally, indicators and targets will be settled which will allow a measure of expected improvement.

In this sense, we can define an indicator as a compact description of a particular observation. These may relate to a specific theme or can be summarized as a certain number of linked issues [1]. More specifically, an indicator is a variable or factor that provides a simple and reliable means to measure achievement, reflect changes associated with an intervention or help you evaluate the results in an organization. It is a sign, signal, or value that allows establishing differences, behaviors, and trends, its measurement can be quantitative or qualitative, and in a certain period of time.

The results obtained by the measurement allow improved planning, since it is possible to observe events in real time, making decisions with greater certainty and reliability.

To collaborate in this regard, a prototype was developed within the framework of a thesis in systems engineering (thesis 1)¹ called "Tool to Support the Management by Indicators" [2], which manages the creation and modification of indicators (Figure 1a) and present the results through control panels (Figure 1b) by registering them at an own database, the database of indicators (DBI).



Figure 1 a - Display of an indicator.

Figure 1 b - Control panel.

Figure 1 – Display of an indicator and control panel. The data shown are simulated [2].

This DBI does not have any sense on its own, but it is “fed” by the systems of organization that are in their transactional databases (TDB) as it is from the primary process as well as another sources, for example, from external organisms, credit bank entities, organisms of state control, and others.

As an example, we can mention some management systems (accountants, ERP-Enterprise Resource Planning, CRM-Customer Relation Management, and so on) and as a particular case in our home study the Guarani, Kolla, Araucano, Kune, Majen, among other systems.

Due to the complexity of the integration processes from different data sources, the necessity of counting with a tool arouses which will allow implementing an orderly sequence of activities over these processes.

In [3] it is defined the lifecycle of indicators by business intelligence tools implementation:

- definition and computerization of indicators
- conformation of specific DBI from the available databases
- informatization of indicators in intelligent tools of businesses: Balanced ScoreCard- BSC [4], Control Tables [1] On-line Analytical Processing (OLAP)
- implementation of tools of control and follow-up, construction of usage profiles and responsibility
- maintenance of indicators, add/remove/modify to selected tools, monitor the use and
- the decisions of the organization related to the function provided by the tools.

¹ Thesis in preparation: Federico Etchepare and Agustín Servat under direction of Dr. Gustavo Illescas

In Figure 2, we can appreciate a simplified scheme of the mentioned cycle that allows a better visualization of the prototype to develop in which specifications are included [5].

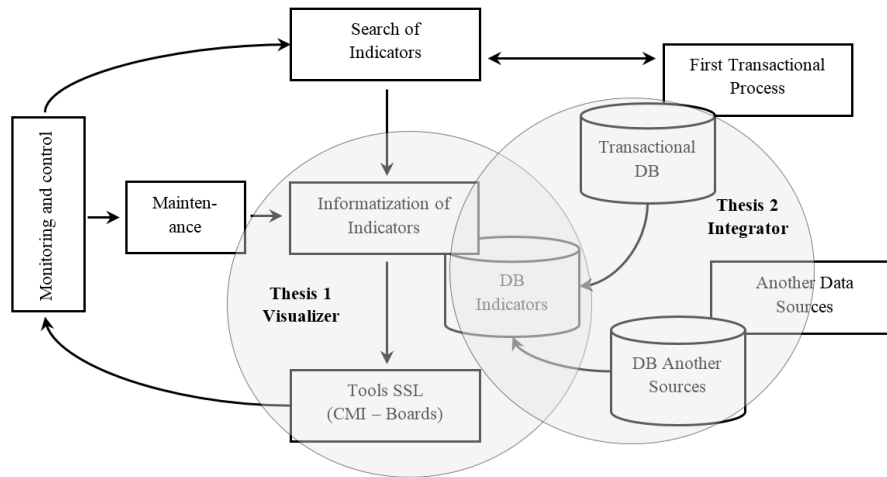


Figure 2 - Simplified scheme with the lifecycle indicators in business intelligent tools [3].

In order to arrive at an integral solution, it is proposed to add an integrative dispositive (thesis 2) to allow the actualization of values of the indicators defined in the DBI. The dispositive will be in charge of providing standard formats to the previous data validation so that the values can be added for update or insertion. For data actualization of an indicator, the information to incorporate is based on the date and the value of the indicator. From the TDB or the DB of other sources can come data from the indicator of different periods. On the other hand, we proceed to classify the states of registered values according to the already defined thresholds in DBI for a certain period.

Once validated the data and incorporated to the DBI, it proceeds to register the value of the last date of actualization to be used when the decision-maker visualizes the indicators.

To summarize, the objective of this work can be observed in the scheme presented in Figure 3.

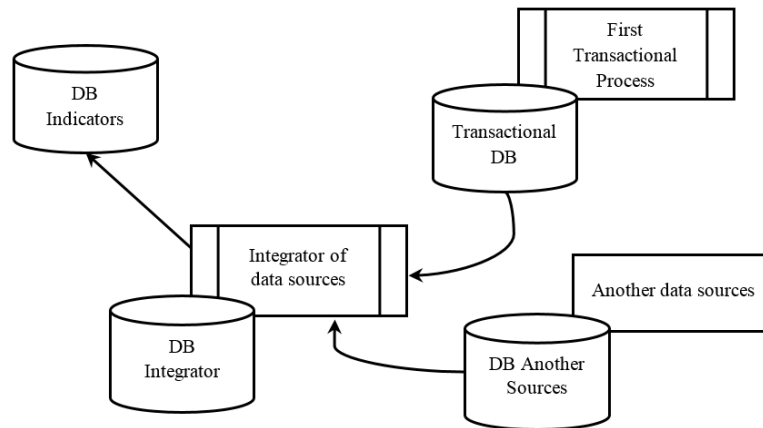


Figure 3 - Scheme of the design of the data sources integrator.

2 Development

The motivation of the current work is a proposal for future works of [3], together with the necessity and institutional interest of the Faculty of Exact Sciences that analyses great volumes of data coming from different sources.

Furthermore, the development of this tool is viewed in the Strategic Plan of Engineers Formation (PEFI – Plan Estratégico de Formación de Ingenieros), which deals with a plan of improving academic indicators for the formation of engineers. The PEFI is driven by the Secretariat of University Policies with the purpose of incrementing the number of engineer graduates to guarantee the quantity and quality of necessary human resources and to make Argentina a developed country [6].

The prototype presented in this work has the characteristics of what we know as the process of Extraction, Transformation, and Load (ETL) [7]. We find different tools as IBM Websphere DataStage [8], Pentaho Data Integration [9], Oracle warehouse Builder [10], Microsoft SQL Server Integration Services [11], among others but that analyzed in detail does not satisfy the specific needs that this work tries to accomplish.

Although there exists at the market many tools that visualize indicators, in our composition of the state-of-art we have not found a tool that allows the extraction of indicators from different and varied sources of data and the integration of them with a visualized solution.

In general, all of them require a manual process to extract the transactional registered information and the supply of such systems, besides, with the possibility that the visualized information will not be find synchronized in the management area, and with the inconvenient of manual tasks and direct manipulation over the data. Thus, it was raised the necessity to build a prototype that automates those tasks.

Then, the current work is based in developing a prototype that acts as an interface between the transactional systems in the organization and the tool of visualization of indicators (mentioned in 1), and through different functions, implement the programming of automated tasks of actualization of indicators (Figure 4).

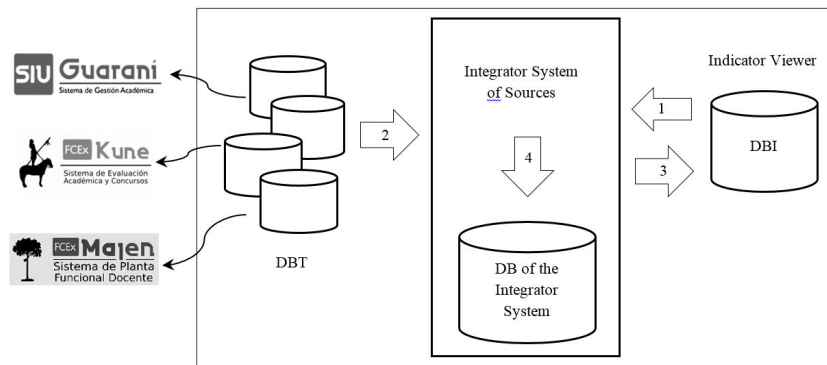


Figure 4 - Reference flow and actualizations to DBI

This prototype may be used in every organization that has multiple data sources and the necessity of evaluating the information from indicators. As a principal case of study, of application and demonstration, the current work is devoted to extract and integrate values for indicators of the different systems that are used in the Faculty of Exact Sciences of the UNCPBA as being the system of SIU-Guarani, of Teaching Plant (Majen) or systems where the teaching contests are managed (Kune).

That is to say, the case study will be first applied in an educative entity but it can be applied without any problems in every field that has recollected data through the database and has the interest to make an analysis to improve some kind of aspect in their organization or business.

The sequence or basic procedure for the use of the tool is shown in Figure 4 and it includes: 1) the lecture of the indicator to actualize from the DBI, 2) the configuration of the connection to the TDB where the values of the indicator are extracted (Figure 5) and the generation to a query to the TDB in an editor of source code (Figure 7) together with the programming of the calendar in the execution of the query, 3) the actualization of the DBI with the obtained values and 4) the register for the audit of the result of the execution in the DB of the integrator.

Figure 5 - Configuration of a connection to a TDB

Fecha	Id Indicador	Fecha Ultima Actualizacion	Usuario	BD Indicadores	Grupo BD Consultas	Resultado	Resultado Descripción	Tiempo Ejecución
18/02/2018	52	18/02/2018 14:02:08	ejecucion_automatica	vm-ifxw-a-bdi_exa	guarani_ingreso	1	Ok	00:00:07
18/02/2018	52	18/02/2018 16:02:54	toba	vm-ifxw-a-bdi_exa	guarani_ingreso	0	Ya existe la tupla (dIndicador, fecha) = (52, 2016-08-15) en la BDI	00:00:01
18/02/2018	52	18/02/2018 16:02:02	toba	vm-ifxw-a-bdi_exa	guarani_ingreso	1	Ok	00:00:01
11/03/2018	172	11/03/2018 15:03:07	ejecucion_automatica	vm-ifxw-a-bdi_exa	guarani_grado	1	Ok	00:00:04
11/03/2018	62	11/03/2018 15:03:10	ejecucion_automatica	vm-ifxw-a-bdi_exa	guarani_grado	1	Ok	00:00:02
21/03/2018	62	21/03/2018 09:03:09	ejecucion_automatica	vm-ifxw-a-bdi_exa	guarani_grado	0	No había datos a almacenar para el indicador: 62	00:00:08

Figure 6 - Audit of the result of the execution in the DB of the integrator

2.1 Code Editor

As a basic requirement of the tool, it was proposed a code editor that will allow writing the necessary queries in SQL enriched with vocabulary programming as PHP. Those queries can be stored in flat files in the server to be later processed in a manual or automatic way according to the calendar, with the frequency of the defined actualization in each indicator.

When a query is defined for the calculation of an indicator for the first time, it is charged to a predefined code to a template, which includes comments according to the DBI guide fields: Name, Methodological Tab, Formula, Frequency of Actualization and Unit of Measure. This is done to orientate the user when writing the necessary code, as it is shown in figure 7.

```

<?php
class indicador_232
{
    /**
     * Datos del indicador
     * Nombre: Estudiantes Sistemas
     */
    /**
     * Esta función consulta la base de indicadores -- NO CAMBIARLE EL NOMBRE
     * Debe retornar un arreglo con el par (fecha, valor) en $valores
     */
    function get_valores()
    {
        $sid_indicador = 232;
        //En $sql se debe almacenar la consulta necesaria para leer de la base de datos de consulta.
        //Debe seleccionar el par (fecha, valor)
        // $sql = "SELECT fecha, valor FROM tabla";
        $valores = Connect::ejecutar_query($sid_indicador, $sql);
        return $valores;
    }
}
?>

```

Figure 7 - Basic template for the code editor

Furthermore, it is now necessary to define the query source; that is to say what type of database will be connected to effectively make the query to obtain the values of the indicator we are talking about.

The template includes an obligatory function called “get values” which must be defined and should return in every case with the field date and value of the indicator. If this is not the case, the prototype includes the emission of exceptions to register the failure and notify it to the user.

2.2 Task modeling

To administrate and execute these queries automatically, the concept of task is applied so that the user can manage them in a simple way, allowing the creation of new ones, modifying the previous ones, consulting on the execution of some of them.

2.3 Organizer

A task organizer (scheduler) brings the possibility to coordinate the tasks, to revise and select them. The tasks are shown on a priority list. Once the organizer selects the tasks, it takes them from the priority list.

Besides, every task has a possible state of execution:

- Pending
- Active
- Successfully executed
- Successfully executed without storage of new data
- Executed with a failure

Each of these states has a different colour that represents them intuitively and visually in the calendar.

2.4 Tasks calendar

Another relevant aspect when dealing with the actualization of tasks is the possibility of using a calendar with visual and detailed information about the state of the execution of done queries as programming tasks (Figure 8).

If the task is not successfully executed, it will be red. If the task is pending or should be done in the near future, it will be yellow. If the task is successfully executed, it will be green. If the task is an actualization of data, it will be dark green. If the actualization is successful without an actualization of values, the task will be light green. Last but not least, if the task is being executed at the moment, it will be blue. So, with a simple visualization at the calendar, the user will have a clear panorama of the state of the executions.

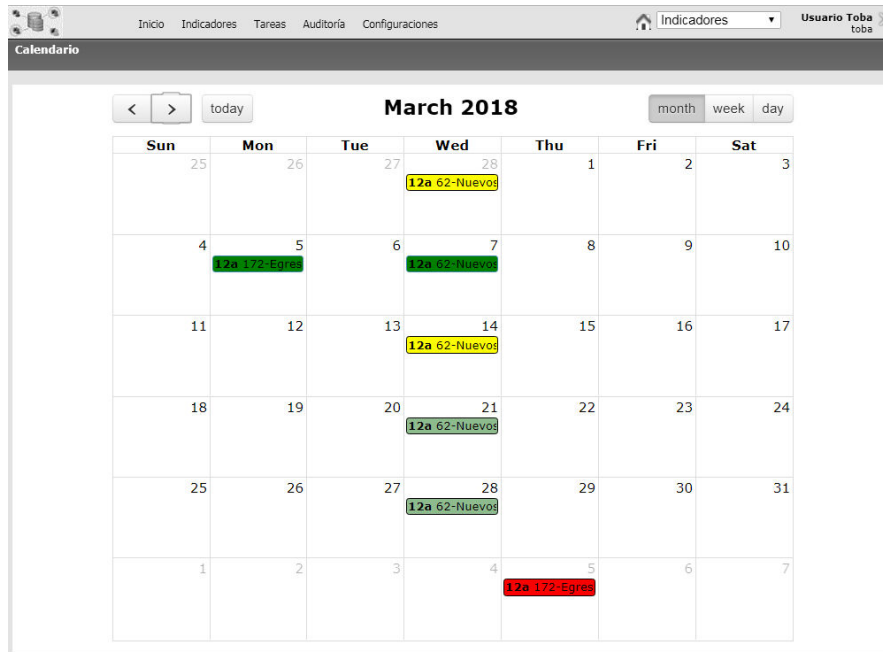


Figure 8 - Tasks calendar

3.3. Conclusions and Future Work

The main idea of this project has been the development of an intuitive prototype simple to use that helps to integrate the different data sources for the management of indicators.

It is achieved the building of a tool which allows the extraction of information and that acts as an interface between the transactional systems of the institution and the prototype that supports the management of indicators (DBI) together with functions that allow the implementation and follow up of automated tasks and register of done operations.

It was possible to install and test the prototype in the Faculty of Exact Sciences of UNCPBA achieving the values of the different systems used in the academic unit to integrate and actualize them in the DBI.

As a test drive, it will be selected (applicants, enrolled pupils, re-enrolled pupils, graduates) for the basic indicators and graduation of each career. On the other hand, the efficiency of indicators (achievement level), efficiency (effort level to get the desired result), efficiency and effectiveness by cohort, breakdown, among others have been defined [2].

Then, it is presented some future work which could be incorporated into the tool after knowing the result of this investigation:

- It allows creating, modifying, moving and deleting tasks from the calendar in an intuitive way.

- It extends the function of connecting with a new database apart from Informix, Postgres, and MySQL.
- It notifies by mail in case of a failure in the execution phase of any automatic task.
- It reports the state of the tasks.
- It continues with the development to improve the use and performance of the prototype which will be the principal matter for a future work.
- It will have a coded access to a REST type web-service for the actualizations in the DBI.

References

1. Ballvé A.: Tablero de control. Ed. Macchi. Buenos Aires. (2000)
2. Illescas G, Santiago M, Martínez M.: Procedimientos y herramientas para la mejora de indicadores académicos en facultades de ingeniería. 20vo Workshop de Investigadores en Ciencias de la Computación (WICC 2018). 1130--1134 (2018)
3. Illescas G. (2014), Tesis Doctoral: Aplicación de métodos matemáticos en el control de gestión por indicadores”. En repositorio digital de la biblioteca central UNICEN. (2014)
4. Kaplan R., Norton D.: Cuadro de mando integral. 2da edición. Ediciones GESTION 2000. (2000)
5. AENOR. Guía para la implantación de indicadores. Norma Española UNE66175. Asociación Española de Normalización y Certificación. Madrid, España. (2003)
6. Plan Estratégico de Formación de Ingenieros (PEFI) <http://pefi.siu.edu.ar>, downloaded 2018-02-19
7. Procesos ETL, <https://www.datasciencecentral.com/profiles/blogs/10-open-source-etl-tools>, downloaded 2017-07-06
8. IBM Websphere DataStage, https://www.ibm.com/support/knowledgecenter/es/SSZJPZ_8.1.0/com.ibm.swg.im.iis.production.iisinfsv.overview.doc/topics/cisodsoverview.html, downloaded 2017-07-21
9. Pentaho Data Integration (Kettle ETL), <http://www.pentaho.com/product/data-integration#data-integration->, downloaded 2017-05-21
10. Oracle Warehouse Builder, https://docs.oracle.com/cd/B28359_01/owb.111/b31278/concept_overview.htm#BABJGDAF, downloaded 2017-05-11
11. Microsoft SQL Server Integration Services (SSIS), <https://docs.microsoft.com/en-us/sql/integration-services/sql-server-integration-services>, downloaded 2017-06-09