Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

BUSINESS INTELLIGENCE

John Jairo Vargas, Luis Miguel Velasco, John Alejandro Acosta, Juan Pablo Camelo
jhon.vargasca@cun.edu.co, luis.velasco@cun.edu.co, john.acosta@cun.edu.co, juan.camelo@cun.edu.co
Docent: Jorge Enrique Portella Cleves, jorge.castillop@cun.edu.co

Summary

The objective of the following article is to present some concepts to understand a little more about what business intelligence is, what is its meaning and what it is used for, how it has influenced the life of companies to create their business strategies and why this tool helps us in making decisions in an organization, some information systems such as DSS, GDSS, EIS, EDSS, with the use of tools and to have a secure control of the information in a database either OLAP or OLTP type according to the structure of the company, EDSS, with the use of tools and have a secure control of the information in database either OLAP or OLTP type according to the structure of the company which in turn are represented by schemes (Star, Snowflake, Constellation), which of course must be accompanied by a Database Management System (DBMS) to manage the information and in turn the data warehouse where it will be stored.

Keywords: Business Intelligence, Information systems, database, schemas, SGBD, DSS, GDSS, EIS, EDSS.

I. INTRODUCTION

To begin we will give the definition of what Business Intelligence is, at present and with the technological revolution that we are experiencing, there is a great variety of tools and computer systems that facilitate and streamline processes in a company, their good execution shows us the differences between the competition and the market. In order to know how to make good decisions regarding the market, the future and the existence of a company, information systems capable of collecting, storing, analyzing and interpreting large amounts of data that are had are implemented and converting them into quality information, layers of being analyzed by stakeholders who based on this information are supported in decision-making and even in the creation of great market strategies.

The implementation and use of business intelligence in companies, helps to a large extent in decision-making based on information collected by information systems and the implementation of tools for this purpose, which is established as one of the steps to continue to bring a company to the highest level of competitiveness.

II. BUSINESS INTELLIGENCE

This concept is put before many years ago in the history of computers, it is mentioned for the first time in history in the year 1856 in a work by Richard Miller Devens, called Cyclopedia of Commercial and Business Anecdotes in which this is spoken for the first time. term, and which I use to describe in the text the way in which a banker was successful, who had certain gifts to understand instability, the market and political problems to get ahead of his commercial rivals (competition).

The term Business Intelligence is heard for the first time in 1958 by a German computer scientist named HANS PETER LUHN, who publishes his article "A Business Intelligence System", in which he refers to a system for storing information for textual documents and defines BI as "the ability to apprehend the interactions of the facts presented in such a way as to guide action to the desired goal," Hans Peter Luhn was given the title of the father of business intelligence.

 doi:10.20944/preprints202103.0579.v1

In the 50's, the hard disk appeared, created by IBM, which revolutionized the data storage industry, thus more components such as floppy disks, laser disks among others appeared, as the industry grew, the need was created more data storage, which gave a leap to the appearance of the first enterprise systems, (DSS) for its acronym decision support systems, with this begins the race for the revolution in the organization of data based on technological systems.

In the 70's, the first advances in BI are shown, using tools that could access data and organize it in the same way, but this in turn was a difficult advance to understand.

Then in 1989 when Howard Dresner, who was a Gartner technician and who proposed business intelligence, as a method for decision-making in companies, using tools based on computer systems, and from which processes and strategies are improved of companies.

It was the decade of the 80's and past the 90's, where real advances began to be seen in the business intelligence revolution, as more suppliers entered the era of technological advancement where the improvement of warehouses took place of data which improved the flow of data and in turn decreased the time to access the data.

As data storage improvements advance, more tools emerge, which are linked to Business intelligence, such as tools that extract the data, transform it and load it by its acronym (ETL), it also appears (OLAP) which is the processing software analytical online.

At the end of the 90's and the beginning of the 2000, many more providers appear that reach the market where the production of data and reports are shown, in addition to organizing and displaying them well, they are sent and in an orderly manner, after several studies and transformations of the methodology and the process to apply this new technology that was visible to everyone, it became necessary to train more people in the matter, since in the past these tools were used only for the technology area , and not to ordinary users, with the new advances developed on business intelligence and the speed with which everything is happening, the need was created to turn BI into a tool that everyone can and knows how to use, making it understandable for all users and in this way it began to be implemented in various types of companies, and in various sectors, which became business intelligence a necessity for companies.

The News of Business Intelligence

What is business intelligence: as a first measure, business intelligence, as indicated by its acronym in English, (BI) Business Intelligence shows us how to bring together the infrastructure, applications and tools, mixing them with the best practices that enable the analysis of the information or data to make the best decision and business performance, which guides the managers or business owners to success.

On many occasions it is thought that Business Intelligence is only applied by company CEOs to make important decisions, but this also helps us to see the real operation of the other areas of the company.

It is known that SMEs are small companies and there are very few that take advantage of business intelligence to carry out their ventures with good decision making, on the other hand, most SMEs still use the old model, or the best the most traditional Microsoft Excel permanently for the creation of your financial reports.

Among many of the benefits of this BI tool, which tries to help understand the trend and extract different perspectives from the data so that as CEOs of a company we can make the best business decisions, tactics or strategies.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

This serves to achieve advantages in the commercial environment that an organization is interested in, thus maintaining a profitable and competitive position, a company that does not implement BI is doomed to make repetitive errors which could be avoided if an intelligence system were implemented which allows the use of existing technologies and taking a leap into the future, making the most of the data provided by the company and which will be transformed into valuable information for our strategy.

In daily life, and although it may not seem like every time we go to a shopping supermarket, we create buying habits ourselves, which are used by the stores, which apply business intelligence, to obtain this data from the preferences of the customers. users and carry out a study of consumption habits, this in order to create strategies for consumers and be able to offer them better commercial offers, these being tempting for the user and in this way to be able to generate greater income, this is reached when there is a good applicability of business intelligence in a company.

A. What is BI business intelligence for?

Business intelligence, in addition to giving the company competitive advantages, gives meaning to the data, transforming it into a gold mine for opportunities, making known the past and the company and in this way being able to take it to a better future.

It collects practical information, which is compiled in comprehensive analysis systems to identify weaknesses or strengths, areas of opportunity to develop strategies focused on objectives, tasks can also be automated and processes improved, as indicated above, good management brings a positive accelerated result in profit, in addition to an effective study of consumer buying habits, this in order to anticipate and predict trends, creating consumer profiles to improve marketing strategies, marketing and sales or services.

Advantage

- Data Storage.
- Immediate availability of data for analysis.
- Process monitoring.
- Better visualization of the company.
- Better understanding and learning from past mistakes.
- More accurate reports.
- Improvement in the strategies the company focuses on.

Therefore, Business Intelligence has evolved by leaps and bounds and has made it possible to include many more processes than there were not many bases, it has been complemented over time and thanks to the investigations that have been carried out, such as data mining , metrics, descriptive analysis, statistical analysis, visualization and data preparation.

B. SMEs and use of business intelligence.

SMEs in Colombia are a majority of small companies, created by ideas that come to the fore by a person, a group of friends and even families, these small companies represent a large percentage of the country's economy, of course it does not compare with The financial muscle of a multinational, which has a large percentage of employability of people in different areas of the economy, many run with the luck of generating good income, but the common employee is the one who carries all the operating burden and is the least paid.

When we realize how difficult it is to create a company in the country, entrepreneurs are helped with courses, training or many who are professionals struggle to get their ventures forward fighting to try to comply with all the regulations imposed by the government, they come out afloat many left behind in the attempt. With the technological advances in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

various fields that are shown daily and the applicability for the operation whether or not to apply it to SMEs, the most innovative will be victorious to continue the long road that awaits them.

Currently, SMEs in Colombia are implementing various information systems many times influenced by key customers or the largest business partner that the company has, often becoming a headache, business intelligence is a system designed for need and growth of each company, this is not a standard system that applies to all companies equally, In any case, SMEs in Colombia do not frequent the use of technological tools for their business projection, and this due to generations of managers or owners who are reluctant to change, innovation and modernization of technological structures if they have them.

C. Business Intelligence and strategy in companies

As we well know, Business Intelligence is the combination of technologies, tools and processes that allow us to transform data into information, which serves us after a deep analysis in shaping strategies, making decisions to obtain the best results in the market and get ahead of the competition, some of the BI strategies lead us to the use of cutting edge technology and in this way create a more aggressive approach and redefining more the focus of what you want to obtain.

The markets most targeted by business intelligence are marketing, logistics, finance, sales and manufacturing, these are the fields in which this type of system can be applied in a certain way better, obtaining amazing results, for example in manufacturing. One of the good results in which BI is applied is in inventory rotation, quality analysis and online production, as well as an area where business intelligence is implemented on a large scale is shown in sales, it supports the sales analysis, customer detection, product analysis, forecasts and projections, in addition to analyzing customer buying habits to create better offers making the business have better profits, all these strategies come from the data obtained by the company to along your way.

Structure of Business Intelligence

A simple structure with which a BI system can be started, and with which a small company or SME could make its leap to technological innovation is by capturing the information of its customers and also that of its suppliers, thus starting a reliable and true source of information, they pass this source to an integration with the database, to this we apply a BI tool which will generate a response and which will be analyzed to later carry out strategies and decision-making, something as basic as This comment can become a great business intelligence project, of course it is guided by trained personnel, professionals in the subject who can take from a simple idea to a complex computer system, this is what makes decision-making possible. , the approach to strategies that will lead a company to a prosperous future.

The tools on which business intelligence is based and more than a structure is:

- Data Mining: it is commonly known as data mining and this automatically identifies trends and behaviors of the data, in order to identify patterns that explain the behavior of the data.
- Agents: they are non-human intelligence programs, which perform basic tasks.
- Data Warehouse: It is a collection of stored data, which remains safe, but at the same time easy to manage and consult.
- Classic BL: stores Molap or Rolap data
- BL operational: Analyze the information taken from the data source.

Business intelligence allows us to optimize resources, verify the fulfillment of goals and objectives, as well as the ability to make good decisions in implementing business strategies.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

On the other hand, we cannot forget about Big Data when referring to this tool today widely used by specialized data centers, multinationals, government and other entities that require a tool of this magnitude to store the large flow of data that they obtain every day after day and which takes advantage of the large volumes of information to make business decisions as well as improve and optimize their processes, Big Data and Business Intelligence are complementary to each other, they work in many ways in similarity, but if we succeed the integrality of the two concepts, we will achieve a very powerful tool.

Types of tools used in BI

Knowing that it is an extensive topic for these times and despite the excessive demand that there is by suppliers, the following types of BI tools can be taken as a reference:

1. Data Management: extract, transform and transfer the data.
2. Data Discovery Applications: collects and evaluates information and then performs predictive analysis that projects it into the future.
3. Reporting: shows the information collected in a graph and in an orderly way to be viewed by the company.

Most used Business Intelligence tools:

Figure 1: Business Intelligence Tools 2019

III. INFORMATION SYSTEMS FOR DECISION MAKING

Information systems for decision making are all those alternatives or data sets that allow us to solve a problem in an optimal and effective way, for this we can see a clear example such as that of a leader who always seeks to be original, thinks about the development of the company, raises strategies in advance so when different situations or problems arise, it is clear how to deal with them. On the other hand, we have the Warehouse that allows us to store a large amount of collaborative information that advises companies through lived or fictitious experiences for a very promising development, where its main objective is to help in decision-making.

Steps for decision making.

Initially we will present the steps that we must take to have a successful decision-making result with the acquisition of new data for the company:

- *Definition of the problem*

You have to focus and define the problem from the beginning, in addition to investigating how relevant it is and the impact it will cause to give its respective priority.

- *Analysis of the problem*

Call meetings in which the exchange of knowledge is sought and generate critical thinking about the problem.

- *Evaluate alternatives*

Propose several alternatives and not only settle for what is already structured, over time the interests of the market change, resulting in continuous adaptation.

- *Choice of alternatives*

Select the one that is ambitious and puts to the test the company that will surely bring a great financial and mental reward.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

• *Apply decision*

Finally, the alternative is applied based on the previous steps, which helped to focus more on the goal and fulfill it properly.

Tools

In order to carry out these steps, it is necessary to implement tools, to better obtain a complete information system:

- *Strategic decisions*

These must be questioned for a long time and will have long-term consequences, that is, we must not generate too much pressure so that emotions are not against us.

- *Forecasts*

Acquire sufficient knowledge that allows anticipating problems and the use of mathematical operations accompanied by statistics that are quite precise, whose interpretation is in the applications that manage to obtain a series of records that will help reduce economic risks, sales growth and generate credibility is suggested compared to other SMEs that even when requesting a loan with one or more financial entities the probabilities of acquiring it since the guarantees of the company will be shown by probabilities and number of monthly sales.

- *Budget*

It gives a more assertive approach to the monetary asset of the company that will give you an overview of what you have, what you can invest in and the allocation of prices according to the market.

- *Programming of activities*

Execute a series of protocols that allow traceability of the service and avoid any type of misunderstanding due to lack of control.

- *Low risk decisions*

They are those whose function is not to anticipate the result but rather to look for all the possible risks that a task carries to determine how feasible it is.

• *Control decisions*

It proposes making reviews of the areas that may be a bit problematic with the employees, but the only thing it seeks is to ensure that the tasks that were initially assigned are fulfilled through controls or follow-ups.

Classification

They are a series of systems that have a specific function for the development and sustainability of the company:

1. DSS (Decision Support System) it is clear what BI is looking for, through a user-friendly interface without prior technical knowledge, which will provide information according to their role in the company.

2. GDSS (Group Decision Support System)

Its purpose is to improve the communication of a certain group of people, who as a consequence have a more assertive decision-making.

3. EIS (Executive Information Systems)

It is responsible for displaying relevant indicators for everyone and easily accessible, which allow almost immediate decision-making.

Database types.

To reduce the risk of improper manipulation of information, whether financial or personal data, they must be stored in a database in BI are:

OLTP (Online Transaction Processing)

It is a technological solution that manages online transactions faster than other types of database, due to its easy structure composed of only recent information, completely discarding the old one, the storage capacity is only 100Mb to 10 Gb that brings as a consequence of being simple consultations and therefore more vulnerable to attacks, it is commonly used for banking procedures, air reservations and of course wholesale sales.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

OLAP (Online Analytical Process)

Unlike OLTP, it seeks to review the data analytically, but not the massive information record that may become unnecessary, its architecture is more robust and seeks to risk a little in terms of speed for security and quality of service since it has a storage of 1TB to 100Pb, does not discard old information without first being reviewed, because it can be useful today, it tries to have queries with multiple filters and access keys with specific functions and thus achieves with all its characteristics to be better than OLAP but requires a more trained staff but it will surely make a difference.

Schematics

In order to have a correct relationship of information, it is necessary to implement some schemes for different situations, these are:

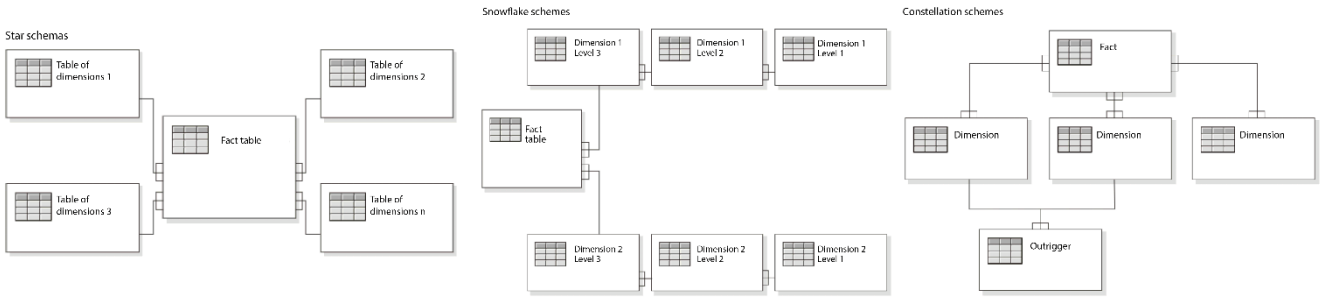

Figure 2: Database Scheme

- **Star**

As a center, it must be made up of a table of facts from which a series of dimension tables are derived, thus forming a star.

- **Snowflake**

It has two dimensions of 3 levels each and that reach the fact table.

- **Constellation**

It is composed of a fact table that connects to two dimensions that in turn reach an Outrigger and a last dimension with a fact-only link.

IV. DATABASE MANAGEMENT SYSTEM

To understand the subject well, we will first define what a Database Management System is.

It is the one that allows any type of company to have their financial, personal, blueprint, credentials, etc., in a friendly interface, in addition to providing security with access control through roles, complex programming languages, and recognized platforms. worldwide such as Oracle, because not all personnel can have access to it and even less modify it without prior authorizations, the administrator of said platform must be a person of full confidence and even so with all possible certifications, since any type vulnerability will put the company at risk.

Components

Depending on the type of information that we have, we must choose a database engine, which is basically a software that works as an intermediary between the database and the applications that will have the information record and access control, some are:

1. Oracle

Founded on June 16, 1977 in the United States, it is an American corporation that offers both local and cloud development solutions whose president is Lawrence J Ellison, its database engine is so reliable

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

that it is used in multiple government entities in everyone, managing finances for example in SAP that allows the planning of business resources, therefore it is quite expensive to implement this alternative and the technology staff must be strictly trained but this does not mean that it provides a user-friendly interface, as well as it has easily accessible support channels in all languages.

2. Microsoft SQL Server

Launched on April 24, 1989, despite not having been in the market for as long as Oracle has the advantage of offering more applications within the same sales package, in addition, Microsoft is known not only in this sector, as a pioneer in technological solutions of all kinds and affordable.

3. MySQL

It started in 1995 in Sweden, whose preference of use is for content managers that use the PHP language, it has as a derivative another engine called MariaDB, but they were bought in 2008 by Sun Microsystems, which in turn acquired the following year Oracle for the sum of 7.4 billion dollars. After seeing the characteristics, you have to choose an engine from which, if or if, you must execute several parameters such as independence, consistency, integrity, backup and recovery of data, minimum response times, control and of course security.

In the database there are three subsystems that classify the data as follows with specific functions:

Data definition subsystem

It seeks to describe the data, its usefulness with other sectors, architecture, and analysis inside and outside the entity.

Handling subsystem

They are those data that can be modified between users, easy to search and add because it is a collaborative work.

Administration subsystem

It is where information is protected through roles, the structure of the database changes, backup copies are made and the access of people outside the entity is restricted.

LANGUAGE

For the operation of any database, a programming language is required or called DDL (Data Definition Language), LMD (Data Manipulation Language), 4GL (Fourth Generation Languages), which are according to each type of user such as administrators, designers, programmers, end users, we start with the modeling that is divided into 5 types:

1. Hierarchy

It is a one-to-many relationship, that is, if we have a database, it will continue with a server and finally an end user, but the last two can only have one database.

2. Network

Unlike the Hierarchical, it allows the relationship of several sectors with others, being a little more insecure, but increasing the company's production.

3. Relational

They are data sets or tuples, for the correct dynamic relationship between tables.

4. Multidimensional

It is the most used in the OLAP model for online transactions, because for each dimension it allows defining a fact, using the snowflake or star schemas.

5. Purpose

It is the model that covers the previous ones, that having one or more dimensions, facts, is given the name of object. And on the other hand we have the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

query language which is the final state of the database, which are also categorized as follows:

Database query language: Used to query the database and information systems, some examples such as SQL, which is object-oriented, D, which is used to manage relational databases, TRDBMS, LDAP, which allows consulting and modifying services. Active Directory using the TCP / IP protocol, DMX as the model for data mining and Information Retrieval Query Language.

ARCHITECTURE

It is the structure that makes a database safe, which must be made up of the following levels:

Internal

It is the level where the configuration files are created, through a physical scheme that describes in detail the type of storage for the database.

Conceptual

In charge of distributing entities, relationships, attributes, restrictions and user operations.

External

Here is the description of interest for a group of users to the rest using views.

STRUCTURE

Now we will explain a little more in depth the structure made up of:

External interfaces

This has as its function the final communication between the database and the user thanks to the forms.

Interpreter

It does the opposite operation, that is, every action that the user does will be translated into a programming language understandable for the database (SQL).

Query optimizer

Since each query that a user wants to make has been online for some time that affects the performance of the database, to which the views are created as a solution that allow the information to be filtered dynamically and thus increase the quality of the service.

Database Engine

In charge of doing all the operations on the aforementioned database.

Storage mechanism

It is any system that physically houses all the data, for this reason it is necessary to implement information backup, access control, which also means a strict control and optimization protocols for the transfer of files.

Transaction engine

It works as a direct intermediary for the database engine or storage mechanism, it must generate logs for each transaction, no matter how small, to have traceability of the queries in addition to having the responsibility of guaranteeing the arrival of the data to the sender as well as the receiver.

Management and operation of DBMS

This section works as a system guide for each phase of the database, all the queries that the user makes with the database, ensures its operation, performs all quality operations such as storage management, monitoring of features and of course the storage mapping that seeks to facilitate the transfer of information and integration between multiple tables of the same database in order to find possible failures.

V. DATA WAREHOUSE

The primary virtue of this class of databases lies in the constructions in which the information is stored (due to its star table models, snowflake tables, relational cubes, among others). This kind of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

information persistence is homogeneous and reliable, and facilitates the consultation and hierarchical regime of the same, always in a different environment from the operational systems.

The primary functionality of this kind of architecture is to contain data that is necessary to later transform them into useful information for the user, for which it will have to give the precise information to the indicated individuals at the optimal moment with the correct format. It is a highly advisable satisfaction when you want to ensure that individuals inexperienced in the knowledge of the system manage to compromise their information because, in this way, users have the possibility of easily carrying out inquiries about their data stores without touching or impair the operation of the systems involved.

It is a type of data hosting architecture that makes it easy for business executives to sort, encompass, and use their data to make strategic choices in their organization.

What is a data warehouse?

A Data warehouse is a corporate information bank that stands out for integrating and debugging information from one or more different sources, and then processing it giving permission with enormous response speeds. The construction of a data warehouse represents in most of the opportunities the first thing that must be done, from the technical criteria, to establish a complete and reliable satisfaction of Business Intelligence.

Data warehouses contain huge amounts of data that are sometimes subdivided into smaller logical entities in relation to the subsystem of the entity from which they originate or for which they are primary.

In this trade a data collection is run oriented to a precise subject, usually an organization, to assist in making choices. All the data that is collected from the different information systems with which a company works are unified and that will later be used to analyze, do and report.

We would have the possibility to say that the Data Warehouse is a key and essential ingredient in Business Intelligence or Big Data that its main function is to stimulate the analysis and dissemination of data in a more efficient way.

These types of commerce have the possibility of being of the physical or logical type and they capture data from different sources of information, the purposes of these data that are captured are mainly for analytical purposes.

Data warehouse location

Like any program or type of architecture, a Data Warehouse can be installed in different types of locations. It is usually installed on a corporate server, although this kind of implementation has become increasingly obsolete, giving way to a great number of possibilities and greater flexibility in the implementation of these kinds of technologies and commerce architectures in the cloud. .

Installing in the cloud is a broad-based implementation that makes it easier for organizations to actively confront the increasing demand for data. Today's circumstance is that the business demands to be able to enter more and more data, it is so much that simultaneously business researchers are exploring the data and their main priority is to achieve high system performance.

The installation in the cloud shows some virtues that are described below:

• Greater speed and speed: allows greater efficiency in the capture and processing of information, allowing the end user better response times for the delivery of results.

• Greater control: this type of cloud service provides us with greater flexibility, more quality in relation to the security and custody of data throughout its life span.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

• Co-location: this provides us with an efficient data load. It allows us to make a joint copy in the cloud for the Data Warehouse with the systems that provide us with the information and its main function is to achieve a very efficient data load, thereby providing the administrator with continuous access to their data.

At this moment we can already understand that a Data Warehouse is a business that centralizes its main functionalities, main uses and areas of performance, we ask ourselves and what relationship could exist between everything mentioned with my business or company? And the answer is that the data Warehouse allows us to rethink and streamline reporting processes or reports eliminates the need to resort to different data sources to verify if they are updated.

Sustain its renewal on an annual basis or check if information has been lost due to the fact that all the data arising from different information sources are in the same central store. This also does not guarantee excellent data quality apart from improving the primary time for generating reports, as well as reducing waiting times.

Using a central data store helps to group data and make quality information available to all members of organizations or companies more efficiently. It rules out inefficient processes that need waiting times that are too "excessive" for the user when its main function should be the concern to communicate some information, managing a single version that allows a suppression of repeated records, errors and inconsistent information, leaving much more reliable and information accurate, complete and updated to carry out that will allow us to carry out our reporting tasks and data exposure and precision.

Business Intelligence solves this data chaos is the Data Warehouse which is a group of processes and actions, it is a collection of data oriented to a topic, integrated and non-volatile in support of the development of decision-making by management considered a comprehensive and timely solution to develop business that is based on Data warehouse as a solution and technique that helps us collect and manage data from many sources in order to provide an accurate response on business issues and make more accurate and accurate decisions based on a large information backup, in a way that was not feasible. This system constitutes an intermediate capable that its correct operation will depend on an excellent coupling and communication between them and which we will deal with below.

Data source

Before exposing the data and information, commonly very diverse, a transformation must be made to a homogeneous format. A data trade draws on both internal company sources and fundamentally relevant external sources:

Internal data from idealization systems for company elements (Enterprise Resource Planning, ERP), from systems for managing relationships with users (Customer Relationship Management, CRM), from transactional databases, from content managers (CMS), from files (Excel, CSV, text), emails, etc.

External data from apps and systems of others, internet pages, communities, cloud services, etc.

Systems that work at this level establish communication with the operational systems of a company through interfaces that are used in the first stage of hosting data warehousing. The central functionalities of these systems are the collection and incorporation of data.

Data extraction

The data is collected using the following extraction procedures:

• Trigger: if a company's operational systems support database triggers, they can be used to automatically extract data. Triggers accept to determine operations that are executed as fast as

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

possible, the certain events that are used to process in changes of the source system's information bank. This leads to the extraction of the edited data into data trading.

• Protocol files: if an operational system is not compatible with trigger technology, the data collection layer may contain programs capable of investigating the protocol files or logs of the source system and obtaining the operations that they have recorded.

• Surveillance programs: if there is no trigger or protocol files, it is customary to use supervision programs that extract the changes in the data of an operational system from algorithms with the ability to periodically create copies of the data and which will subsequently be compared with data already processed.

As it is not feasible to enter the databases and the systems do not support any of these three procedures, they must themselves define the probable changes and send any important modification to the data warehouse.

Testing area

Every day, data is imported into data commerce from the different entity element planning (ERP) systems and other business-related program systems for further transformation. It is common practice to normalize data before mixing it in data trading using extract, transform, and load (ETL) utilities. These utilities read the primary data from local databases, make the development of transformation to the data trade (filtering, coherence, format changes, etc.) and write in order to be processed. In this sense, the book "The Data Warehouse Toolkit" (Ralph Kimball) is a fundamental reference since it establishes what it should be in a company, there are many forms of interface and creation of DW.

Every business wisdom program developer seeks to impose a methodology with their products.

ETL

Common usage scenarios for ELT are within the macro-data domain. Among other things, you can start by pulling all your source data to flat files on scalable hosting such as the Hadoop Distributed File System (HDFS) or the Azure Data Lake Store. You have the ability to use technologies such as Spark, Hive, or PolyBase to find out the source data.

The key point of ELT is that the data trade that is used to do the transformation is the same data trade in which the data is consumed. This data commerce reads directly from the scalable hosting, rather than uploading the data to its owner hosting. This approach skips the data copy step that is provided in ETL.

This is a time consuming operation in the situation of huge data sets. In custom the target data trade is a data host using a Hadoop cluster (with Hive or Spark) or an Azure Synapse Analytics instance. Generally, a schema is superimposed on the data in the file at the time of the query and is stored as a table, which makes it easier to query the data as in some other table in the data trade. These are known as foreign tables because the data does not live in the data merchant managed hosting itself, but rather in some external scalable storages.

Data trading only manages the schema of the data and applies the schema on read. Among other things, a Hadoop cluster using Hive would describe a Hive table where the data source is, aptly, an address to a group of files in HDFS, in Azure Synapse the same result can be achieved - with PolyBase creating a table with the data stored externally to our information bank. Once the source data has been loaded, the data present in the foreign tables can be processed by means of the data trading functions. At macro-data levels, this means that data trading must have the ability to do massively parallel processing (MPP) that splits the data into smaller extracts and distributes the processing of the extracts across numerous

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

machines in parallel to their In the process, the final stage of the ELT pipeline usually edits the source data to a final format that is more efficient for the types of queries to be supported.

Among other things, you have the ability to partition your data. Additionally, ELT could use improved hosting formats like Parquet, which stores row-oriented data with the appearance of columns and provides improved indexing.

DATA STORAGE

A data warehouse data trade is a collection of data oriented to a precise topic for a non-volatile and time-varying company or company, which assists in making a choice in the entity where it is used. It speaks, more than anything, of a finished file of an organization, beyond the transactional and operational information, stored in a data source of information created to benefit the examination and the effective dissemination of data the OLAP online analytical processing.

The hosting of the data should not be used with current usage data, data warehouses commonly contain huge proportions of information that are sometimes subdivided into smaller logical entities in relation to the subsystem or entity from which they come and for which are essential, in a data trade what you want is to contain data that is necessary or usable for an organization, which is used as a data repository to later transform them into useful information for the user, a data trade must give the accurate information to the crowd indicated at the optimal time and in a correct format. Data trading provides answers to the claims of professional individuals, using Election Support Systems (DSS), Executive Information Systems (EIS) or utilities to carry out inquiries or reports.

In relation to the incorporation of data, most data warehouses have online analytical processing (OLAP) functionalities that accept moving the data to multidimensional constructions. The so-called Online Analytical Processing is an analytical procedure supported by ETL development with which the important administration data for the company is compressed.

The final individuals have the possibility of easily carrying out inquiries about their data warehouses without touching or damaging the operation of the system, the performance of a data trade must be paramount and the following must be taken into account:

The integration of the data that comes from databases distributed by the different entities of the organization and that will continually have different constructions should make a global specification and a comprehensive examination of the entire organization in data trading simpler.

Separation of data used in day-to-day operations versus data used in data trading for the purposes of disclosure, assistance in making choices, for organization and for control operations, the two types of similar they have to fit into the same information bank, since they serve very different objectives and could get in the way of each other.

DATA LOGIC

Among the accessible process and evaluation utilities in data warehouses, OLAP apps have been consolidated as a user-level work platform. The utilities that are used in the framework of online analytical processing show functionalities with which individuals have the possibility of formulating particular queries to the data commerce that agree to examine its multidimensional content.

Its representation by the OLAP procedure makes it easy to model the refined data in relation to as many predefined dimensions as are essential. In this development, researchers have the possibility of resorting to different operations with which an OLAP cube can be modified

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

METADATA

The metadata is the data about the data that is why it is often considered as the element that defines the properties of a note. Its importance in Data Warehouse projects is given for a customary reason. In a BI startup, 80% of market profits are shown with a metadata cover that makes it possible to translate the content of the data that comes from a field in a table into a business criterion. On the planet of DWH and onboarding processes, the creator or architect of DWH has equivalent claims that force him to find a way to manage this metadata, as is generally the case in onboarding processes.

The use of the term from a computerized approach takes this generic conception as its starting point, both individually and when a group of metadata details a group of data or elements.

Metadata are characterized by being highly structured data that describe properties of the information such as content, quality, relevance and other situations or attributes, present differentiations that will depend, in a more recent instance, on the rules within the apps to know the internal composition of the data schemas, they can be classified in functionality of different criteria, such as their content, variability or functionality.

At present, Big Data, Internet of things and cloud computing, metadata have a unique importance compared to sources proportions of information that grow exponentially, the administration of metadata can be of great impetus to gain operational efficiency and / or make superior choices to obtain competitive virtues. Its multifunctionality, in fact, is among the primary properties of this vigorous utility for data management as an asset, specifically these are some of the features that most superiority of a passable metadata management at the time of doing better data administration and governance processes.

SYSTEM OPERATIONS

• Extraction: in the data extraction stage, important information is selected from different sources following a push or pull plan, if they are extracted by the pull strategy, it is the data trade that drives the extraction.

• Transformation: at this stage the data that has been extracted is cleaned and translated into the format of the information bank where it has to be stored.

• Loading: at this stage the data already transformed are saved in its corresponding information bank in the data trade. This first level of collection may contain a so-called staging area, a temporary sector in which the data that is going to be loaded undergo pre-processing, which can be essential in the situation of very complex ETL processes.

As there are different sources of origin, the incorporation of these data in commerce uses different transformation and cleaning utilities that have the possibility of being classified according to these three categories.

• Data migration: these programs accept to determine simple transformation rules that convert heterogeneous output data into a target suit format.

• Cleaning: in this situation, these programs rely on the so-called fuzzy logic and neural networks to improve the quality of the data while solving errors, incomplete data and repetitions. When performing this data cleaning, it uses predefined rules, algorithms or look-up tables, this operation is also often called quality management.

• Data audit: these programs are used to find rules and relationships between data, as well as to detect those that contravene the established rules and are predictably erroneous. After incorporation, the extracted data enters the central information bank

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

(Core Data Warehouse) through programs that meet the following rules.

- They test the integrity conditions
- Classify the data
- Calculate the aggregations
- They calculate the access constructions
- They fragment the data so that the entry is more efficient

VI. FORMULATING THE RESEARCH PROBLEM

Thanks to the Pandemic that we are currently experiencing, the business sector also had to implement business alternatives so that they did not fall, therefore, Business intelligence arrives, which seeks to help by collecting information for decision-making, they will be shown in In advance, which will be the best investment with the lowest risk to the entities. Now, all those who want to implement BI, generates many doubts as it is a fairly broad subject, that is, it makes them insecure because they do not know with certainty how to carry out this project, it should be noted that there are many advantages if applied in the right way, that's the goal of this article. Therefore, the research problem is what is the structure of business intelligence?
With the previous one, we seek to give a brief advice to guide all those interested in business intelligence.

General purpose

Guide about Business Intelligence, taking into account the digital channels and information sources of companies focused on the subject.

Specific objectives

- Perform the acquisition of viable and legitimate information to know how to make the best decisions.
- Organize the information collected in Databases that guarantee access through roles, backup copies and monitoring of the application, in addition to being friendly with the end user.
- Distribute the files with the existing information security protocols, to the servers that will house this project, which will guarantee the development of the entity.

VII. CONCLUSIONS

In Colombia, most companies have knowledge of business intelligence tools and consider the need to implement it, however, they do not know the proper way.

Starting with the advantage of Information Systems that seeks above all to have control and acquire an ability to make decisions in a company, in addition to the fact that for information security issues it is better to use the OLAP database option that is It is less vulnerable, but this does not mean that it should be fully trusted as it also has its disadvantages, being slow and perhaps generating a feeling of poor quality for some users, but even more information can be stored as long as the entity has the right financial.

Being such a broad subject, it turns out to be somewhat confusing, after following the conceptual map acquired by the teacher Jorge Portella, it is possible to organize in an optimal way what a database management system is, understanding its structure and clearly noting that it is not It is something that is implemented immediately but requires years of practice, certifications, advice, because it is a service that has been professionally worked for many years that allows ease of online transactions, information tracking for investigations, among other processes that were carried out. only in person, to which its contingency plans play an important role, as is Oracle, which is responsible for providing feasible development solutions from the software level, provides trained personnel, offers resources at an academic level for those who want to start in this alternative of business, to robust servers that will be in charge of hosting the data and making the most of it.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 24 March 2021 doi:10.20944/preprints202103.0579.v1

VIII. REFERENCES

Rosa, S. L. (2010). Business intelligence: Audit and control. Leganés: Carlos III University of Madrid.

Lagunes., X. A. (June 15, 2016). Inteligencia de negocios. Una herramienta para el futuro https://www.gestiopolis.com/inteligencia-negocios-una-herramienta-futuro/

Martín, B. (October 15, 2015) *https://www.beeva.com*. https://www.beeva.com/beevaview/estrategia-negocio/un-viaje-en-el-tiempo-por-la-historia-del-businessintelligence/

Devens, R. M. (1865). Cyclopedia of Commercial and Business Anecdotes. Appleton https://books.google.com.co/books?id=vqBDAAAAIAAJ&printsec=frontcover&hl=es&source=gbs_ge_summary_r&cad=0#v=onepage&q=Business%20intelligence&f=false

Pérez, A. (June 17, 2019) Strategic decisions in the company: how to make them. https://www.obsbusiness.school/blog/decisiones-estrategicas-en-la-empresa-como-tomarlas#:~:text=Las%20decisiones%20estrat%C3%A9gicas%20son%20las,afectan%20m%C3%A1s%20al%20corto%20plazo.

EmpresaActual (February 15th, 2019) Predictive analytics in companies: applications and examples. https://www.empresaactual.com/analitica-predictiva-en-las-empresas-aplicaciones- y-ejemplos/.

Grupo Logicalis (December 13, 2015) Business Intelligence: Budgeting and Decision Making https://blog.es.logicalis.com/analytics/presupuestacion-y-toma-de-decisiones

Iñigo, S.J. (2018) Decision making under risk: Sequential decisions, application examples in business management https://addi.ehu.es/bitstream/handle/10810/33174/Toma%20de%20decisiones%20bajo%20riesgo.%20Decisiones%20secuenciales.%20Ejemplos%20de%20aplicaci%C3%B3n%20en%20la%20gesti%C3%B3n%20de%20empresas.pdf?sequence=1&isAll owed=y

Sinergia (2007) Decision Support Systems (DSS) https://www.sinnexus.com/business_intelligence/sistemas_soporte_decisiones.aspx

TechTarget Contributor (July 2019) OLTP (online transaction processing) https://searchdatacenter.techtarget.com/definition/OLTP

Evaluando Software (August 24, 2016) OLTP Systems: Transaction Processing, Management and Maintenance https://www.evaluandosoftware.com/siste mas-oltp-procesamiento-administracion- mantenimiento-transacciones/

IBM (April 18, 2013) Dimensional drawings https://www.ibm.com/support/knowledgecenter/es/SS9UM9_9.1.2/com.ibm.datatools.dimensional.ui.doc/topics/c_dm_star_schemas.html

Infase (September 6, 2017) Database Engines https://infase.net/motores-de-bases-de-datos/

Wikipedia (Feb 14, 2021) Oracle Corporation https://es.wikipedia.org/wiki/Oracle_Corporation

Diarioti (July 6, 2010) Creator of MySQL Appeals Against Oracle's Purchase of Sunhttps://diarioti.com/el-creador-de-mysql-apela-contra-la-compra-de-sun-por-parte-de-oracle/26998#:~:text=En%20febrero%20de%202008%20www,Europea%20en%20enero%20de%202010

Ctisoluciones (2015) what are and what are database management systems for?

 doi:10.20944/preprints202103.0579.v1

https://www.ctisoluciones.com/blog/son-para-sirven-los-sistemas-gestion- bases-datos

IBM(April 18, 2013) The Dimensional Data Model https://www.ibm.com/support/knowledgecenter/es/SS9UM9_9.1.2/com.ibm.datatools.dimensional.ui.doc/topics/c_dm_star_schemas.html

Figure 1, Josehn (December 26, 2020-Business Intelligence Tools 2019) Software and hardware, the best Business intelligence (BI) Tools in 2019-2021 https://softwareyhardware.com/software/herramientas-de-inteligencia-de-negocio/

Figure 2, IBM (April 18, 2013) Dimensional Drawings https://www.ibm.com/support/knowledgecenter/es/SS9UM9_9.1.2/com.ibm.datatools.dimensional.ui.doc/topics/c_dm_dimschemas.html