BUSINESS INTELLIGENCE AND ITS BIG EVOLUTION (March 2021)

authors

Andrés Velosa Andres.velosa@cun.edu.co,
Anderson Quijano
Anderson.quijano@cun.edu.co,
Camila Martínez Nydia.martinez@cun.edu.co,
Gustavo Pabón Gustavo.pabon@cun.edu.co,
Jorge Portella Jorge_Portella@cun.edu.co

Keywords: Business Intelligence, Data warehouse, Data Marts, Architecture, Data, Information, cloud, Data Mining, evolution, technologic, companies, tools, software.

Abstract - Information has been and will be a vital element for a person or department groups in an organization. That is why there are technologies that help us to give them the proper management of data; Business Intelligence is responsible for bringing technological solutions that correctly and effectively manage the entire volume of necessary and important information for companies.

Among the solutions offered by Business Intelligence are Data Warehouses, Data Mining, among other business technologies that working together achieve the objectives proposed by an organization. It is important to highlight that these business technologies have been present since the 50's and have been evolving through time, improving processes, infrastructure, methodologies and implementing new technologies, which have helped to correct past mistakes based on information management for companies.

There are questions about Business Intelligence. Could it be that in the not-too-distant future it will be used as an essential standard or norm in any organization for data management, since it provides many benefits and avoids failures at the time of classifying information. On the other hand, Cloud storage has been the best alternative to safeguard information and not depend on physical storage media, which are not 100% secure and are exposed to partial or total loss of information, by presenting hardware failures or security failures due to mishandling that can be given to such information.

I. INTRODUCTION

Over the years until today, since the origin of the companies, they have sought to facilitate their methods of obtaining and organizing their information, so each company at that time had its own ideas to use this information as effectively as possible.

Where the first terms of BI technology (Business Intelligence or business intelligence) really began to be seen was in the 50's, where the IBM company created hard disks which were assembled in the computers where the information storage was implemented.

This was the key point, since later this term was already mentioned in a publication that was made in an encyclopedia in the 60's where they began to implement methods of this technology, and with the passage of time there were important dates where important advances which began to be universal where different companies used similar methods to create, obtain, straighten, organize and provide information in a pleasant way for any type of end user.

II. OBJECTIVES

- To know the different techniques and strategies used by business intelligence for a successful result in the analysis of information.

- To understand the type of process in terms of large-scale data migration as data is very important and therefore today requires certain features that require safeguarding information in a broader sense than what has been seen with the evolution of data.
- To learn from BI technologies for our future as systems engineering professionals.

III. BUSINESS INTELLIGENCE

Business Intelligence (BI) is a group of methods or applicable technologies oriented especially in companies, so that they can more easily, to create, organize their data and have a better understanding of these, achieving through these activities the functions are in a timely manner to meet the proposed goals.

These strategies make the information collected is represented in knowledge and thus take better alternatives for success for a company in the growth and change of the market.

Talking a little about its origin, it is indicated that it was created or developed in 1965 according to the publication of an encyclopedia called Cyclopaedia of Commercial and Business Anecdotes, but it was not until the twentieth century with the rise of technology and electronic information that began its design as it is known today. Some important dates to highlight its advances are the following:

o 1956: Hard disk drives were created by the company IBM (International Business Machines) which were assembled in the first computers of the time to store information.

o 1958: Peter Luhn, considered the father of the term business intelligence, being a senior researcher at IBM, developed tools to understand how information was examined to make the best business decisions using the term Business

Intelligence.

o 1970: In this decade, other companies also became involved in business intelligence, such as SAP, which developed software to facilitate data entry into databases.

o 1980: In this decade, databases are constantly developing and make it possible to gather information from different sources and store them in a single database.

o 1990: In this decade is where a progressive advance is seen, since several companies realize how important and effective this technology is, so they begin to create and launch several business intelligence tools, but they had a main drawback focused on the fact that they were perceptive and difficult to use by the end user, since they had to have the knowledge to use this technology or to access the information they had to resort to the IT department.

o 2000: It was in this decade when the potential of the use of Business Intelligence software truly began to be realized, forcing more companies to require its use with significant improvements which allowed any user without knowledge of systems to use it easily performing the different actions that this technology provided with the information.

o 2007: It is evident that only this year has reached a higher percentage of information storage through these tools than in previous years, making its accelerated growth in companies.

o 2017: According to the figures provided, it is expected to increase by 7% with revenues of US$18,000 and by 2020 this value is expected to grow gradually reaching or exceeding US$23,000.

On the other hand, these technologies create an action which collects information in a more effective way, thus being an innovative and proactive tool to answer the questions that a company has: What happened, what will happen, why did it happen, and what will happen? Capturing, processing, and analyzing the information provided.
BI is used in many work areas, but the main ones are the following:

· Sales: It is used to study orders, buyers, trade, items, business opportunities and future reach.

· Marketing: Used for sectioning buyers and innovative items.
· Finance: It is used to study the costs and other financial areas of a company.

· Logistics: It is used to study the methods necessary to carry out the organization of a service.
· Manufacturing: It is used to study yield, waste, quality and inventory.

Business intelligence is characterized in several ways which optimize corporate management, such as the following:

· Data access, as this is the main reason to ensure that consumers have access to the necessary and requested information.

· It helps in the decision-making process, where we want to go beyond showing the data, but also to allow users to interact in an analytical way with the information they need to use.

· End-user oriented, with a user-friendly interface so that the user can use the tools necessary for their work.

IV. BUSINESS INTELLIGENCE ARCHITECTURE

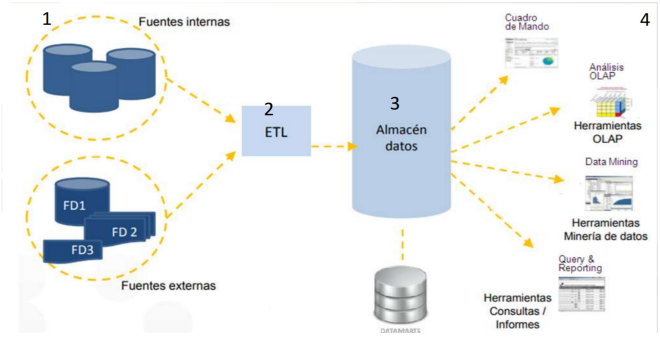

Figure 1. BI Architecture

Below, we will learn more about each of its components:

a. Information sources

Information sources are provided by the company, in which data can be obtained from internal or external sources that are handled. In the internal, is the stored information of the organization (databases); in the external, are the data for example from social networks or data provided by a third party.

The data are the most important in every BI analysis, since from them hundreds of results can be extracted according to what each company requires. These sources of information must be with all the necessary columns to find the solution,

As any start in BI requires the data with which the analysis is going to be performed. In the information source, the data from the electronic warehouse will be provided. This data can be in .xls format, text files ETC.

The provided data must ensure:

- Safety
- Reliability
- Ease of recovery
- Ease of administration

In order to obtain a successful solution, it is necessary that each process is carried out properly, starting with the information sources, since they are the driving force of all the analyses.

b. ***ETL Process***

As its acronym indicates (Extract, Transform and Load), this process is in charge of extracting, transforming and loading the quality information, that is to say clean and purified before being loaded to a DTW (DataWarehouse), this information must be purified, because all the data that the organization handles does not serve for the same thing, in this case the decision making.

ETL phases:

- Extraction
- Transformation
- Load

Data warehouses

Figure 2. Data warehouses

Nowadays, Data Warehouses play a very important role in companies or organizations, because they are able to store large amounts of information in a secure, reliable, easily retrievable and easily manageable way. Since the end of the 80's the term Data Warehouse was defined as the collection of data destined to a specific subject, which is integrated and changes through time; all these characteristics must support the respective changes to obtain a good data management.
Innovations in technologies are advancing very quickly according to the needs required by the business market. The handling of large masses of information is an issue that worries the owner or customer, who wants their data to be safe and to be able to access them without causing damage or loss. Technologies such as Big Data, Predictive Analysis, Real Time Analysis, Cloud solutions, Mobile solutions, among others, have been decisive to form a Hardware and Software infrastructure for business, using elements of Business Intelligence.

It is said that an DW is oriented because the information is classified according to the interests of the organization; It is integrated when taking information from different sources. This data must be verified before being added to the DW; this is known as Extraction, Transformation and Load (ETL). The Data Warehouse was created to be read, but unadulterated. This information is kept intact.
The architecture of an DW is made up of the following components:

· OLTP (On Line Transaction Processing), it refers to all the exchange information that a company generates on a daily basis and where these external sources come from.
· LOAD MANAGER, In this process the ETLs are in charge of extracting the data from the OLPs to manage them and subsequently load the results obtained in the DW.
· DW MANAGER, in charge of transforming and integrating source data, it is also in charge of intermediate storage for decision making.
· QUERY MANAGER, is in charge of supporting management processes and executing relational queries of a database.
· TOOLS AND DATA CONSULTATION, is the medium in which the user can interact with the data to perform queries.
· USERS, are those who use the system and make the planning of activities..

TYPES OF DATA WAREHOUSES

Departmental Data Warehouse: Obtains partial business information. They only need the data that are of interest to the department or to a group of people.

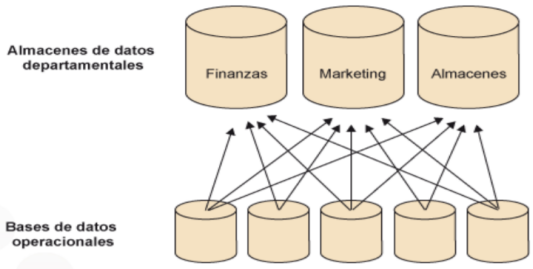

Figure 3. Departmental type

Corporate Data Warehouse: Stores all the organization data with the maximum detail.

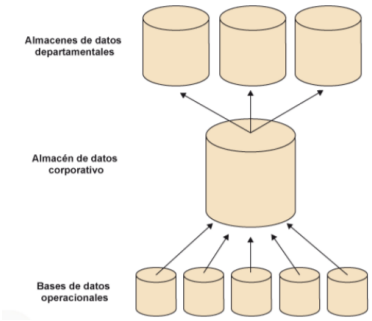

Figure 4. Corporate type

Operational Data Warehouse: It has a structure between operational and decision making.

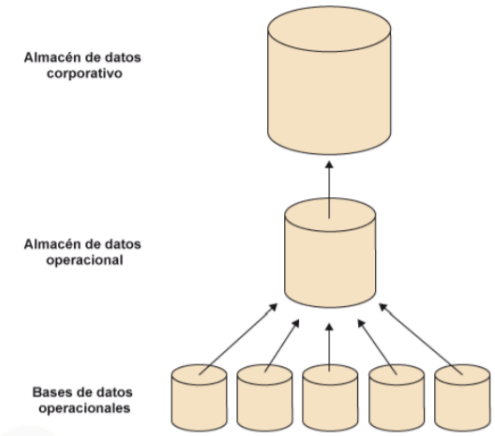

Figure 5. Operational type

Having gone into the subject of Data Warehouses (DW), we have already identified what they are used for, their importance in the BI world, but we do not know what they were used for before, and what is the future of this storage technology? In the past, DWs were created using repetitive and structured data that was filtered before being passed to the Data Warehouse. Today, the Data Warehouse has evolved so much due to the high amount of information that is being handled and distributed globally, contextual information can now be attached to unstructured data, and this in turn can be stored in an agile and simple way.

Data Lakes have been emerging due to the needs of Data Management, however, it is not the replacement of Data Warehouse: they come to complement existing resources and support new experiences.

Some differences between Data Lake and Datawarehouse are the following:

· Data: a Data Warehouse only stores data that is modeled or structured, while a Data Lake stores all types of data, whether structured, semi-structured or unstructured.

· Processing: when loading data into a Data Warehouse, they must first be given shape and structure, in other words, they must be modeled, which is called "schema on-write". In a data lake, the data is loaded unprocessed, as it is, and when it is ready to be used, it is given shape and structure; this process is called "schema-on-read".

· Storage: Big Data technology, such as Hadoop, has a very low data storage cost compared to the Data warehouse. There are 2 reasons for this: 1. Hadoop is an open-source software, acquiring this type of licenses and support is free. 2. Hadoop is designed to be installed on low-cost hardware.

· Agility: a data warehouse is highly structured, it can take a long time to change the structure, due to the processes that are linked to it. A data lake is not structured like a data warehouse, which allows a developer or scientist to easily configure their models, queries and applications.

· Security: Data warehousing has been used for a long time, while Big Data (the basis of the Data lake) is a bit newer. This means that data warehouse security is more mature than that of a data lake. However, security improvements are being made for the Big Data industry.

¿What is Data Warehouse?

We talk about the Data Warehouse as business intelligence on the amount of structured information where most of the data is stored, which allows the great exploration of them to obtain the best performance in accessibility when being managed by the end user, since it is a corporate database and is responsible for integrating and debugging information from different or one source in order to process it and have an analysis with high response speeds because today's business users need to generate decisions to ensure the availability, performance, backups and physical recovery for the client which requires the collection of information created to support applications in terms of different decision making.

As large companies need to debug information, they ask themselves the question about the data warehouse: and why is it one of the most important elements? Previously and since 1970 and today is where all the data is concentrated with a special design to exploit information by fragments known as data marts that are certain types of models such as star and snowflake that allow report generation, information analysis, dashboards, data mining, among others, which allows the difference in terms of design as transactional databases are quite standardized because if you do not have a good model in terms of design, it does not meet the purpose of the data warehouse would not get a good Business Intelligence project.

Since its beginnings in the 70s and 80s, when it began to see the great need for organizations that required an easy way to store their data and access information, is where the scientist Bill Inmon, which is considered the father of data warehouses, defines the concept with the type of standards and schemes by which the information is governed, as well as the permissions of the user that gives a competitive advantage that provides quality and reliable data, which leads to better decisions by a company as it unifies all databases in one place which with the type of information processing, solid platforms of consolidated historical data are obtained for analysis in which there is a business approach of an information base that accesses the manager or directive which takes the type of decisions as to know: sales, customers, events, invoicing, products or inventories which has in a historical way the data to support the decision making processes for it is required to analyze the type of sources either by indicators that validate the repository that can be physical or logical.

The Data Warehouse known as an electronic storage architecture where large organizations and companies seek to store securely, easily, reliably and in such a way easy to manage, and in some cases easy to retrieve consolidated data from various sources or systems of the company dealing with structured data which has a main objective and high quality which is to have all the information together in order to make an analysis and divide to determine sectors or strategies in a historical and current mode of data, which to generate

analytical reports for workers and to have the knowledge of the entire company.

What are Data Warehouse structures?

We rely on the architecture as it can be divided into three structures as basic, basic with a staging area and with data marts area.

Basic Structure: Operating systems, Blueprints and files that provide raw data which are stored together with metadata so that end users can access them for analysis and reporting.

With Test Area: is added or can be placed between the data sources and the warehouse where data can be cleaned before going into storage.

With Data Marts Area: add systems designed for a particular line of business, which are separated by inventory windows, purchasing and users could access data from one or all data marts departments.

Figure 6. Data warehouse

Properties of the architecture are essential security properties with which to monitor the types of access, separation of which consists in the analytical process must be kept separate from the transactional mode, scalability requires a good hardware and software architecture with easy way to update data with which it can meet the type of purpose to the user, extensibility which must support the architecture to host new applications and technologies and administrability that requires the management is not difficult to use.

¿What elements make up the Warehouse's data?

Data access services, as opposed to traditional information storage.

Data Preparation (data checks and repairs).

Modeling, which is unique if not adaptable to business processes and needs.

We can say that, as well as stores large amounts of data that are collected on a time basis and integrated by multiple sources and among it is fundamental the architecture of a Data Warehouse consists of the following layers:

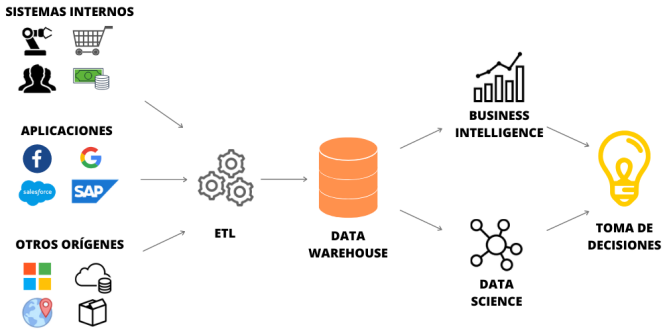

Figure 7. Data warehouse architecture

Data source layer

Which is fed by different data that can be in any type of format such as: plain text, Excel files, other types of databases of which can vary between:

-Operations data such as accounting data inventories and marketing data.

-Web server logs

-Third party data

Data extraction layer

These are extracted to the Data warehouse system where some data for the type of cleaning can be applied.

Test area

The transformation of cleansed data by transforming data mart and data warehouse data is applied to facilitate further data integration.

ETL layer

One of the most interesting because it applies the type of logic and intelligence to transform the data from a transactional nature to an analytical nature and where the design phase is implemented, in which requires a good elaboration, and a good cleaning of the data is obtained.

Data warehouse layer

Data transformed from the three types of entities of the data mart, data warehouse and operational data warehouse.

Logical data layer

Business rules are applied and stored which do not affect the data transformation rules but would not allow viewing reports.

Data presentation layer

Information in graphical or tabular forms through browsers such as a report that is generated on a daily basis as a horoscope or alert, etc.

Metadata layer

Consists of the logical data model and consists of reporting on the data stored in the data warehouse..

System operation box

It gives in real time information on how the system is performing and gives system performance and user access history.

DATA WAREHOUSE IN THE CLOUD

Figure 8. Data Warehouse in the Cloud.

Currently, the world is undergoing two major transformations that differ in terms of their potential to have levels that merit business innovation.

One of the first transformation tasks is the area which and as everything requires a boost to increase agility in general with Information Technology departments are experiencing the rapid increase in demand for data so managers want to get more and more historical data as well as the implementation of new data flows in the warehouse and have more and more increase.

The second area of transformation is based on looking at how to improve. Because of today's data-driven environment, there is a need to improve cost control as it reflects the growing need and today's pandemic to ensure that all sensitive and strategic data is fully secured in a cost-effective manner. Benefits and solutions for a Data Warehouse in the cloud.

Why migrate to the Cloud today?

The update will always have great impact in terms of business transformation with cloud, data and analytics, and internet of things of which these three technologies are essential and key part of today's ecosystem which has great potential to always generate significant high impact on business.

The importance of the cloud for the data warehouse is linked to 3 key factors:

Agility improvement

Companies increasingly want to go further and seek to support new areas. To do so, they need a new infrastructure that is fast, efficient, and secure. These objectives can be achieved using cloud services.

Increased cost control

IT departments are needing to consolidate existing data marts, each running on dedicated or proprietary hardware, into an integrated environment. Cloud services offer to keep data protected throughout its lifecycle and are more cost-effective in terms of security.

Co-location

It is much better if the data warehouse sits alongside source systems that are already running in the cloud. This offers faster data loading, which means that users get access to their data in a timely manner.

Advantages of moving the data warehouse to the cloud:

- Easier consolidation y rationalization
- Faster monetization of data in the cloud
- Cloud offers protection.

Data warehouse: Migration as opportunity?

Recently on June 13 of 2020, the type of strategy that can be given use to the data so that it occupies a central place for the domain of companies wishing to seek and give scales that allow it to be strengthened and have greater assertiveness at the time of making decisions that make data-driven organizations which is to make more efficient provision in terms of information and today there is talk that the data is no longer a matter of IT, the democratization of data is encouraged.

Today it is essential to evaluate all variables and make excellent decisions as data grows exponentially and the pace of collection is progressively increasing as business requirements also increase, which requires challenges for any migration in the data warehouse environment.

DATA WAREHOUSE VS BIG DATA VS BI

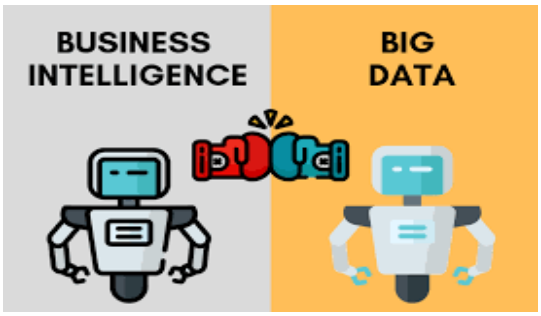

Figure 9. Business Intelligence VS Big Data

There are 3 different concepts of the way in which data will be used, however, in all 3 falls the responsibility of having a large volume of information in different formats, which can be structured or unstructured. The objective is to find benefits in the different technologies and choose the best one according to the needs of the organization.

BIG DATA: This is a large volume of data or information that has variety, complexity, speed of increase and has characteristics of being unstructured. The objective of this technology is to analyze data in real time, which can come from different sources, such as social networks, images, emails, among others.

DATA WAREHOUSE: stores consolidated and structured data from different sources. Its main objective is to be very accurate and of high quality to provide good support when making decisions. The idea of this technology is to gather all the information and divide it to make a better analysis of the information.

BUSINESS INTELLIGENCE: This is a process that filters the data obtained by the data warehouse, selects it and transforms it into useful information for corporate strategies.

Returning to the architecture of AD, it operates as follows:

Data is extracted from applications, database, among others. This information is housed in different types of systems, with different formats, the data is integrated, transformed and cleaned to be loaded into the AD. The information is structured in multidimensional cubes to be used in dynamic queries.

Users enter the cubes, using different query tools, reports, analysis etc.

A multidimensional cube converts flat data into a matrix of N dimensions. The most important objects found in a multidimensional cube are attributes, hierarchies and indicators, in this way the elements will be related and the intersection will represent the final value being evaluated.

It is evident that the most common problems in the DP are found in the collection of requirements, analysis and design. This is due to the fact that there is no standard methodology for its execution.

There are different methodologies that allow them to work. Some are new and others have been modified for different needs. Some of these are:

- Kimball life cycle
- DWEP
- HEFESTO
- RWM
- KM-IRIS

Kimball life cycle: Also known as Dimensional Modeling or DM, this methodology proposes different phases: Project planning, project definition, dimensional modeling, physical design, ETL (data presentation design and development), deployment, maintenance and growth.

DWEP: The Data Warehouse Engineering Process is based on the modification of the standard Unified Process (PU) and consists of four phases (Initiation, Elaboration, Construction and Transition) and seven workflows (Requirements, Analysis, Design, Implementation, Test, Maintenance and Post-Development Review). The difference of this methodology with the PU methodology is that the DWEP consists of more flows compared to the PU, and that in the DWEP these phases end and start again.

HEFESTO: This methodology consists of four phases: Requirements analysis (identify questions, identify indicators and analysis perspectives, conceptual model), OLTP analysis (Determination of indicators, establish correspondences, level of granularity, extended conceptual model), logical model of the DW (type of logical model of the DW, dimension tables, fact tables, joins) and ETL processes.

This methodology has the following characteristics:

- It is easier to understand and identify what is wanted.
- It is flexible to the needs of the user or company.
- Allows the end user to participate at each stage.
- By using conceptual models it is easier to analyze.
- Phases are not repeated

RWM: Rapid Warehousing Methodology is based on DW growth development and consists of five phases: Goal Definition, Requirements Definition, Design and Modeling, Implementation and Review. This methodology is iterative like DWEP.

KM-IRIS: This methodology was developed by the Integration and Re-Engineering of Systems (IRIS) group of the Jaime I University (Universitat Jaume I in Valencian) in Spain and consists of five phases: identify, extract, process,

store and share. This methodology is not well-known and is not repetitive.

It can be concluded from these technologies that they all have their importance in data management and that they must be interconnected with each other to provide the best solution for the organization.

Data mining

Figure 10. Data mining structure

The technology era is advancing more and more, today we find great engineers with incredible skills that has helped the technological science to emerge in a way that no other market grows easily, years ago there were not many electronic means to store data in a systematized way and yet with the passage of time were available different tools to organize information and have a safe and reliable access. With the evolution of the storage capacity, this allowed this era to carry out new research to get great results from all the data, although at the beginning there were not many with time the data increases every day, since with every action we perform on the Internet everything is aimed at storing data, from entering the mail, to an enrollment in any virtual course. With all the increase of data it has been required to manage large databases and with the treatment of all these data arises the Data mining or data mining.

It is possible to be evidenced that in the beginnings of the data mining it was very difficult to acquire a software that will realize these functions, since this software had disproportionate charges that not any company could buy them, in that time a software of these could get to be worth 40.000 dollars these costs could only be solved by big companies these are some of the advanced software in the beginnings with a high cost.

- SPSS Clementine
- Insightful miner
- IMBs DB Miner

However, we can not rule out the idea that there are some low-cost software, even free for those companies that do not have ample finances but willing to apply these methods in order to emerge in the midst of high competition, for the greater benefit of many companies technology advanced so that any company could use data mining and have multiple options for growth and prevention of any problem. One of the software developed open source are:

- Rapid miner
- Weka
- Orange

Thanks to the creation of advanced software to deal with exaggerated amounts of information, data mining became one of the most vital tools for any company, considering every day more information is generated, due to the fact that each and every person provides important data, either to their trusted operator or a simple internet purchase that is capturing the most sensitive data of a citizen, a previously unknown topic. Now, you may be wondering, what is data mining or what does it consist of? Data mining is a deep analysis of large amounts of data, which allows the creation of new strategies that benefit the company, strategies that are difficult for the competition to detect. Therefore, all this is focused on the growth of companies and the sustainability of large companies.

To be able to reach an information analysis, it is essential to consider the whole process that must be done to the data, starting from storage, having a database engine that is constantly updated with proper backup to prevent the information from being lost at some point. As a second is the ETL, where the information is cleaned and transformed for the customer to see it in a friendly way. Continuing with the creation of reports whose purpose is to perform queries depending on the requirements of each of these for once created can be published in different environments and that there is an easy search for those who require it. Data mining is used in many environments, such as large providers of telephony, internet, and TV services, which contain amounts of users nationally and internationally, where they must conduct a market study, shortcomings to improve the quality of their service and offer better options to customers and thus gain more users every day. For data mining is essential to have different techniques to find the treasure that you want to find, it is essential that the analysis of the information is not obvious, but to go beyond what is required, as today there may be many who run this type of activity, the issue is to find the key point for the company to emerge or level any kind of fall that may have had.

DATA MINING TECHNIQUES

The main objective of the descriptive method is to identify patterns in which data can be identified and thus described. One of the best known is Clustering which consists of discovering sequential patterns, we can give as an example the way in which a customer or a group of customers relate the different transactions carried out by the same in the course of time, this type of methods has been used to identify which products should be purchased when shopping at the supermarket.

- **Clustering technique**

This technique is also used to classify, but in this case, it is a grouping technique, groups that currently are not defined, but it is necessary to create to relate them, since today there are items that for example go hand in hand, or customers always buy in company because dependence is required.

Almost 100% of the people when they buy a product carry it together with another product because it is indispensable in its preparation, such as rice and salt. This type of technique works to know the type of products that are most consumed and that go hand in hand to prepare offers to customers, offering low prices and quality rice, which is what housewives fundamentally like for its texture and easy cooking.

This is called Cross selling which is a cross-selling, usually the other product is needed to prepare it and upselling works as growth, which in this case would be the famous offers that all customers are looking for.

Another of the methods that compose it are the predictive methods, they have also covered large fields of medicine in which they help us to identify whether a tumor is benign or malignant, this technological advance helps us to advance hand in hand with science in which we no longer have to undergo arduous, long and painful clinical procedures that in the end when we knew the result were not very reliable and could lend themselves to diagnostic errors.

We must keep in mind that the predictive method is not only applicable to the medical field, in recent years we have been identifying how this method has also made inroads in the sports world. For example, in Soccer we can determine data such as the result that a soccer team will have at the end of the season, what is the possibility that the team will be relegated or achieve its objectives, what types of players are the ones that are lacking, and in which positions we should reinforce according to the information provided by this method. The betting market lately has had a significant growth, we can use the predictive method for betting markets where we can predict the result that will have, or which would be the most accurate approximate.

- **Regression technique**

This technique joins variables in order to form an equation that helps me answer a key question. Something like: What kind of customers are we going to have? Is it viable to buy the product? Each variable will be in a box, which will allow us to analyze each type of customer and be able to segment them in order to know their profiles and thus find the most accurate answers and create the growth strategy.
An important issue in data mining is that the system is partially unknown, why? Because where the data is already known, there are no more variables to be discovered, it is necessary to unknown the variables. This is functional when a fault is unknown, when a product is going to be manufactured to know if the product is viable or not, for the study of financial issues, to detect frauds, etc.

The CRISP-DM standard, are the steps that allow me to apply data mining, currently has 6 steps that are as follows:

1. We must know the problem and get more involved with the area that has the problem.
2. We must have a database where the information is being collected and know what type of information it has and what it is related to understand the relationship between the data.
3. Generate a data cleansing, eliminating data that is not required and, in this way, guarantee the reliability of the information that is stored.
4. Set in motion the creation of the model to be used, validating if it is descriptive or predictive and carry out the previous tests of the model.
5. Perform an analysis of the results of each of the tests and identify the possible answers to the situation.
6. Execution of the data mining process once it has been reviewed.

With all the information related to data mining we realize that the more information is obtained, it is much easier to start the whole process, because with the amount of data you can take advantage of the problem, to raise quality indexes and / or improve a service that is being provided, in addition to making a favorable offer for both the company and its customers, making a big difference to the competition.

Nowadays it can be seen that large companies have many drawbacks, many of them want to cover customers, but do not measure the magnitude of the quality of what is offered. Returning to the issue of telecommunications providers that are the most visible, they have serious shortcomings when providing service, there are sectors in which their antennas work properly, but others that have serious flaws, even so, their strategies in these sectors do not favor their customers, because of this many users come to the last circumstances before proceeding with a withdrawal. This is not only because of their quality of service, also in call centers the attention is not very good, many times the strategy of an advisor is simply to cut the call, to which the user is not at all satisfied and for this in the country have the highest percentages of nonconformities, but not only telecommunications entities have problems, as a company is constantly growing they have to have shortcomings or to grow must promote their product or service with good offers, avoiding damage to the finances of the company.

As we see not only in this type of field applies data mining, but many companies also today use them, this is the search for new ways every day more potentializing markets, service and / or products that are provided to users. Always as a company with a high user rating is essential to be innovative, creative when creating a strategy, strategies that help maintain that status and that the company does not lose its prestige. This became a competition, because if a company takes out a promotion or a big discount, other competing companies want to participate in the same thing by raising their promotions, perhaps many times they are visual deceptions, offers that have no discount, so it is seen many times the clothing stores that in the famous "Cyber day" aims to capture the attention of

users, Many people in the midst of their ignorance fall into this famous game without thinking that many times the value is doubled and then make a supposed "discount" that sometimes ends up being more than what you normally have. However, this is a great move of many companies when making their offers, but this helps to raise their sales rates, they favor themselves financially and users are happy with their "discounts". Although not many people fall for this deceptive advertising, there are still many more who are willing to buy unnecessary things, at unnecessary times, just to get a great discount, something never seen before.

Many of these strategies are carried out with different methods and techniques that have the Data mining, having favorable points many times for both parties, other times not so favorable for users but before the user is convincing.

Thanks to data mining, many investigations have had an incredible evolution, because it has managed to contribute to large scientific, financial, technological studies, etc. This is something that has helped that everything that has to do with data management is enhanced, due to the benefits that has been achieved as a result of the analysis and studies based on the step by step to create an excellent method and achieve effective strategies that many companies have had the facility to adapt them to the decisions of the company for free and without limitations of companies, simply anyone who wants to implement such processes can do so with the proper knowledge and good handling of data.

There are several companies which designed business intelligence tools which meet different needs and costs, some focused-on SMEs and others on multinationals, in the most popular ones we find developments made by large companies such as SAP, Microsoft and Oracle, and others by companies growing in these technologies.

- Microsoft Dynamics: This management software from Microsoft is focused on customer relationship management through an integrated tool called CRM (Customer Relationship Management) which provides options to have a detailed view of customer information from the first link to the purchase of products.
- Factorial HR: This software from Factorial is focused on providing solutions and services to the problems presented in the human resources area of small and medium-sized companies.
- IBM Cognos Analytics: This management software from IBM powered by AI (Artificial Intelligence) is focused on visualizing data and insights, supporting the analysis cycle from its origins to commercialization, sharing information with members of the organization with multi-cloud environments.
- SAP business intelligence: This SAP (Systems, Applications, Products in Data Processing) company management software is organized in models corresponding to each business area of the company with a client-server structure where the database is stored on a server and users can access the information from the server as long as the client computers have this software installed.
- Oracle Business Intelligence: This management software from Oracle company is focused on a BI server implemented adaptable to enhance the use of several users as possible being simultaneous with parallelism giving added value to applications based on web services (J2EE and .NET).
- Tableau: This management software from the company Tableau Software created in 2003, is focused on facilitating the exploration and management of information allowing to find the necessary information and to share it in a simple and effective way for companies.
- Sisense: This management software from the company Sisense powered by artificial intelligence (AI) focused mainly on the cloud with a high-performance elastic engine with a friendly interface to any type of user with or without the technical knowledge can use the tools provided API (application programming interface) realizing these custom functions for any workflow.
- Clear Analytics: This management software from the Clear Analytics company is focused on the end user based on Excel, but fed with accurate information as it is audited, it is very useful, since practically all users use Excel in their daily life making changes such as creating, typing, examining and visualizing company information with the features of BI tools.

V. CONCLUSION

The main objective of this work is to address the general topic of business intelligence (BI - business intelligence) from its origins, its implementation, its use, its development and worldwide use, as each time its use has been booming and in this century with the constant transformation of technology is widely used, as it covers the necessary tools for information management coupled for any type of company from SMEs to multinationals to reach the fulfillment of their projections.

Thanks to the researchers of different companies, which with their contributions have made these methods easier and simpler for any type of user without having knowledge or computer studies to use the information consulted and to be able to capture and share it in the most agile way with the required results.

Currently we have observed the great growth of business intelligence technology which is estimated that in the period from 2020 to 2025 will have annual growth of 7.6% driving business growth which is estimated at 33,300 million dollars according to a report by Markets and Markets, moving about 23. 100 million dollars, although by issues of the pandemic occurred worldwide on March 11, 2020 has reduced its growth by issues that several companies have ceased its activities, as some market sectors are customer presence and have had to gradually transform to the digital age, where his forte was in the areas, insurance, banking and financial services, with the gradual revival of activities business intelligence is bringing new ideas which will optimize procedures with greater control of their services for the necessary deployments analyzed for the confrontation of the crisis.

Figure References

[1] Figure1. BI Architecture
https://virtual.cun.edu.co/contenidos/migracion2020/sistemas/s9/business-intelligence/u1/recurso5.pdf

[2] Figure 2. Data warehouses
https://sites.google.com/site/lainformaticajf/base-de-datos/almacen-de-datos

[3] Figure 3. Departmental type
https://miriadax.net/documents/54270034/55985398/M2-Intro-DW-FIC.pdf

[4] Figure 4. Corporate type
https://miriadax.net/documents/54270034/55985398/M2-Intro-DW-FIC.pdf

[5] Figure 5. Operational type
https://miriadax.net/documents/54270034/55985398/M2-Intro-DW-FIC.pdf

[6] Figure 6. Data warehouse
https://www.rberny.com/estrategia-de-ti/data-warehouse/

[7] Figure 7. Data warehouse architecture

https://www.mistralbs.com/definicion-arquitectura-data-warehouse/

[8] Figure 8. Data Warehouse in the Cloud.
https://carlospesquera.com/wp-content/uploads/2021/02/logo-dwh-nube-150x150.png

[9] Figure 9. Business Intelligence VS Big Data
https://www.iebschool.com/blog/big-data-business-intelligence-diferencias-big-data/

[10] Figure 10. Data mining structure
https://virtual.cun.edu.co/contenidos/migracion2020/sistemas/s9/business-intelligence/u1/recurso5.pdf

Bibliographic References

[1] Data Warehousing : Migration as an opportunity by Redacción PowerData on November 13, 2020 Available:
https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/data-warehousing-la-migracion-como-oportunidad

[2] Benefits and solutions for a Data Warehouse in the cloud by Redacción PowerData on November 13, 2020 Available:
https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/beneficios-y-soluciones-para-un-data-warehouse-en-la-nube

[3] Components of Bi Isis Lay 12003204 Carlos escobar 12003202
https://es.slideshare.net/elemento3000/componentes-de-business-intelligence

[4] Layers of any Data Warehouse architecture Cun 2021 Available:
https://virtual.cun.edu.co/contenidos/migracion2020/sistemas/s9/business-intelligence/u5/recurso5.pdf

[5] Management News.analysis and information 2021 Diponible:
http://www.emb.cl/gerencia/articulo.mvc?xid=1769#

[6] What is a Data WareHouse 15 Jun 2017 Available:
https://www.youtube.com/watch?v=jFsRdTcljeU&ab_channel=AuriboxTraining

[7] data-warehouse - PowerData

https://www.powerdata.es/data-warehouse

[8] Introduction to Data Warehouse the Corporate Information Factory by Angels Rius Gavidia - miriadax.net

https://miriadax.net/documents/54270034/55985398/M2-Intro-DW-FIC.pdf

[9] Design of a Data Warehouse based on Data Warehouse Engineering Process (DWEP) and HEFESTO - Castelán García Leopoldo, Ucharán Hernandez Jorge Octavio

https://virtual.cun.edu.co/contenidos/migracion2020/sistemas/s9/business-intelligence/u3/recurso6.pdf

[10] Methodologies for developing Data Warehouse - Architecture and Engineering Journal, Vol 7, No 3, December-2013, pp. 1-12 by Leonard Brizuela, Eric Ismael, Castro Blanco Yudi - Researchgate.net

https://www.researchgate.net/publication/316634417_Metodologias_para_desarrollar_Almacen_de_Datos

[11] Data Warehouse - EcuRed by Berta Delfos - June 10, 2010

https://www.ecured.cu/Almac%C3%A9n_de_Datos

[12] Data Warehouse Features - vmware Docs - 07/02/2020
https://docs.vmware.com/es/VMwarevSphere/7.0/com.vmware.vsphere.storage.doc/GUID-057D6054-0A51-4023-B90A-D737DB0426F4.html

[13] Data Warehouse: Types, Architecture and Characteristics - Information Technologies

https://www.tecnologias-informacion.com/datawarehouse.html

[14] The future of data warehousing is in the cloud - PowerData, August 11 2020

https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/el-futuro-del-data-warehouse-esta-en-la-nube

[15] Data Warehouse Architecture: Corporate Data Warehouse data areas - Evaluando Software.com, 1 November 2011

https://www.evaluandosoftware.com/arquitectura-del-data-warehouse-areas-de-datos-del-almacen-de-datos-corporativo/

[16] Data warehouse - profundizaU3, CUN 2021 Available

https://virtual.cun.edu.co/contenidos/migracion2020/sistemas/s9/business-intelligence/u3/recurso5.pdf

[17] Data mining.SAS

https://www.sas.com/es_co/insights/analytics/data-mining.html

[18] Data Mining- Find out how data mining will predict our behavior 2021.Iberdrola

https://www.iberdrola.com/innovacion/data-mining-definicion-ejemplos-y-aplicaciones

[19] Data Mining Vs Big Data

https://www.campusbigdata.com/big-data-blog/item/82-data-mining-vs-big-data

[20]Data Mining- Classification Rules 2013.Jul 01, Universidad Miguel Hernández de elche

https://www.youtube.com/watch?v=Pf9E3nI_8WA

[21] Data mining process- 2021 CEUPE

https://www.ceupe.com/blog/proceso-del-data-mining.html

[22] Data mining. How is it applied? 2007 IMF

https://blogs.imf-formacion.com/blog/tecnologia/data-mining-como-aplica-2020 2007/

[23] Data mining techniques- 2016 July 13

https://www.evaluandosoftware.com/tecnicas-data-mining/

[24] What is data mining for? SYLOPER

https://www.syloper.com/blog/recursos/para-que-sirve-la-mineria-de-datos/#:~:text=y%20su%20%5B%E2%80%A6%5D-,Significado%20y%20caracter%C3%ADsticas,datos%20en%20un%20contexto%20espec%C3%ADfico.

[25] Importance of data mining in the business world- METRIC BANK

http://metricbank.com/datamining-importance/

[26] Data mining software: make data analysis of data more effective 30-Jan-2018 IONOS

https://www.ionos.es/digitalguide/online-marketing/analisis-web/software-de-data-mining-las-mejores-herramientas/

[27] Free tools for data mining 02-Nov 2018 Dia Garcia

https://www.inteldig.com/2018/11/herramientas-gratuitas-data-mining/

[28] ETL- SAS

https://www.sas.com/es_ar/insights/data-management/what-is-etl.html

[29] What is Data science? - 12 Mar 2019- NEOLAND

https://www.neoland.es/blog/que-es-data-science

[30] What is an ERP? - ANER

https://www.aner.com/que-es-un-erp.html

[31] What is OLAP? 29 DEC 2011-Business intelligence made easy

https://www.businessintelligence.info/definiciones/que-es-olap.html

[32] Design and implementation of a scalable data warehouse architecture

https://www.dataprix.com/es/articulo/business-intelligence/arquitectura-del-data-warehouse-areas-datos-nuestro-almacen

[33] Tools to implement a Data Warehouse- 2016 Agos 30-softpedia

http://softpei.blogspot.com/2016/08/herramientas-para-implementar-un-data.html

[34] Data Warehouse, technical aspects, features,uses,benefits,components, OLAP tools

http://fccea.unicauca.edu.co/old/datawarehouse.htm

[35] Guide to build a Data Warehouse, case study-10 Feb 2017-Paula Rochina

https://revistadigital.inesem.es/informatica-y-tics/guia-construir-datawarehouse/

[36]What is metadata and what is its use? - 2016 Mar 19- Power Data

https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/que-son-los-metadatos-y-cual-es-su-utilidad#:~:text=Es%20com%C3%BAn%20describir%20el%20t%C3%A9rmino,la%20informaci%C3%B3n%20de%20los%20mismos.

[37] What is reporting and business intelligence? -2019 Sep - Gestalia

https://www.gestalia.es/que-es-reporting-y-business-intelligence/

[38] Database vs data warehouse: key differences

https://www.guru99.com/database-vs-data-warehouse.html#:~:text=Database%20to%20analyze%20data.%20is%20a%20collection%20of%20related%20data%20that%20represents%20some,is%20designed

[39] Top 15 Best Business Intelligence Tools

https://mopinion.com/es/las-15-mejores-herramientas-de-inteligencia-de-negocios-una-vista-general/

[40] What is Business Intelligence (BI) and what tools exist? 5/10/2017

https://blog.signaturit.com/es/que-es-business-intelligence-bi-y-que-herramientas-existen

[41] What is Business Intelligence? - Sinnexus

https://www.sinnexus.com/business_intelligence/

[42] What is business intelligence and what are its key aspects? 03/21/2018

https://www.universidadviu.com/int/actualidad/nuestros-expertos/que-es-business-intelligence-y-cuales-son-sus-aspectos-clave

[43] Business intelligence

https://es.wikipedia.org/wiki/Inteligencia_empresarial

[44] What is Business Intelligence?

https://es.workmeter.com/blog/bid/177356/que-es-el-business-intelligence

[45] What is Business Intelligence and what is it for 25 March 2019

https://www.obsbusiness.school/blog/que-es-business-intelligence-y-para-que-sirve

[46] 10 concepts to understand Business Intelligence

https://www.il3.ub.edu/blog/10-conceptos-basicos-para-entender-el-business-intelligence-del-presente/

[47] Business Intelligence. Theory and concepts

https://www.gestiopolis.com/business-intelligence-teoria-y-conceptos/

[48] Fundamental BI Concepts

http://www.solidq.com/ebs/BI_y_Analytics_Volumen_I.pdf?utm_campaign=Descarga+e

[49] Business intelligence: definition, characteristics and solutions 17/08/2017

https://vegagestion.es/business-intelligence-definicion-caracteristicas-soluciones/

[50] Business intelligence: concept, features, software and examples 30/11/2020

https://blog.lemontech.com/business-intelligence-concepto-caracteristicas-software-y-ejemplos/

[51] What is BI? Business Intelligence Definition and Solutions

https://www.ciospain.es/big-data/que-es-bi-definicion-y-soluciones-de-business-intelligence