



Integration of Technological Tools in Business Intelligence

Fabio Andres Gaince Alvarado, Erminson Garcia Victoria,
Erika Viviana Reatiga Velosa, Claudia Liliana Medina Tovar,
Wilmer Guillermo Piracoca Barragan and
Jorge Enrique Portella Cleves

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

March 25, 2021

Integration of Technological Tools in Business Intelligence

Autores, 07-03-2021

➤ Gaince Alvarado Fabio Andres "fabio.gaince@cun.edu.co", Garcia Victoria Erminson "erminson.garcia@cun.edu.co", Reatiga Velosa Erika Viviana "erika.reatiga@cun.edu.co" Medina Tovar Claudia Liliana "claudia.medinat@cun.edu.co" Wilmer Guillermo Piracoca Barragan "wilmer.piracoca@cun.edu.co" Jorge Enrique Portella Cleves "jorge_portella@cun.edu.co".

Summary – The technological tools in Business Intelligence are used so that companies can enter the data and can also carry out research, reports, and analyze some nonconformity about some value found to present an Alert. Most Business Intelligence tools use them because they need users to enter, study, and report with the most frequent data found in data warehouses, data marts, and operational data warehouses. Most of these platforms are used when you need to analyze, plan, run, monitor, and even measure complex, high-value applications, or verify a customer's need to build applications that are easier to use. BI defines how a company handles its internal and external information. Previously, BI was not as developed by companies because they were very difficult to manage as for reporting as well to access it, because for a person without the necessary experience it would be even more complicated. But, in recent years, the perspective or idea has been changing because more and more tools have been created for the use of Business Intelligence and in addition organizations see the fundamental value that it is representing in giving good results and giving favorable statistics, therefore they begin to use them. For this reason, one of the important tools in most companies or organizations is the database because it allows us to have access to the information, allows us to behave that data with different areas of the company and allow the study of the information in order to make fundamental decisions. The fundamental goals of BI tools is for all data in a company to be transformed into useful information, exposing companies to making the best decisions, and implementing and exposing the best decisions of agreements to business intelligence.

1. INTRODUCTION

This article will allow us to give a general approach to the different uses that can be given to information according to the different problems of a company. It is very useful because it

gives us information analysis to identify trends and shorten useful information. organizations must simplify access to accurate, unified data in real time.

Today, great technological advances have been found in which tools are implemented, applications that contribute to the easy access of the information required by companies in real time which brings with it great benefits both economically and in the development of high availability and efficiency activities such as cloud computing services. Implementing these technologies will show the positive results for companies in decision making. The different concepts, uses and benefits of the various information management tools will then be defined. Likewise these tools allows us to have several benefits such as: Having the ability to separate, refine, secure, simplify and display information data in automated knowledge. Have available to all individuals who request the fundamental information of the data to make the best decision. The data is obtained without the need for other areas of the company, you have the possibility to perform the OLAP process. Equality in the management of both internal and external information, so that it is the same information to calculate everything equally. In a system that handles stable and useful tools.

That is, it is the ability to change the data into information and such information in understanding, so that processes in business can be improved, the following is understood:

1. WHICH IS A DATABASE MANAGER

It is a program that allows the storage, modification and extraction of information.

▪ What database managers are for us for

Los gestores de bases de datos son muy importantes porque

are those that allow us to create databases, which allow us to store, manage and make use of the information.

Companies in Latin American and European countries, whether small, medium or large, at least make use of a type of manager, whether relational or non-relational. To access all that information that is stored basically companies perform the development of some web applications or desktops where they can make use of them through the different operations that the system allows them through the permissions they have in it.

For example: The anonymous company to make use of the exchanged information that they have stored in a database or want to enter new exchange information each month, the solution that the company finds is to carry out the development of a exchange application where the employees of the company will be those users who carry out the upload of the information or extract it according to the schedule that was requested from the team of engineers using the new technologies, to obtain that information, various SQL queries were implemented to perform the various operations that the system will have (Query required information, modify, delete, etc.).

1. What type of database manager to choose

When choosing the type of database manager you mainly have to take into account the type of databases that is best to solve the need, many of the companies choose to use relational databases such as SQL Server, MySQL, etc. To perform the structuring of a relational model that allows them to have the flow of information in order to make large-scale projections since the syntax of these database managers is very easy to manipulate when performing any operation on it. However, non-relational databases like MongoDB allow us to store large volumes of information, with this database the response time is much more agile to process operations.

1. How to do good modeling

To perform good modeling in database managers it is very important to take into account the initial phase of a project, because this is the one that allows us to perform a very thorough analysis about the information collection with which we can identify what are the requirements that the user requests for their information system.

Once these requirements have been identified, the database is modeled in order to solve the problem that the client has, however, this is where many of the companies have to make mistakes when modeling the database since because they do not have an organizational structure within the area of development, what the developer or programmer normally does performs the design.

Therefore, when reviewing modeling we will encounter many common mistakes that make this type of people such as duplication of information, which causes performance to be affected at query execution time.

While a DBA (Data Base Administration) developer at the time of database modeling, they consider the following factors such as:

1. Performance when processing queries.
1. Prevents information redundancy.
2. Improves data access.
3. Ease of maintenance.
4. Reduces resource costs considerably.

The main features of databases are as follows:

1. Logical and physical independence.
2. Reduced duplication of data.
3. Concurrency with multiple users.
4. Data integrity.
5. Optimal query optimization.
6. Information security.
7. Audits.
8. Backups.

Database systems consist mainly of a language of definition, data manipulation, and query types, each database manager has its own data type and sometimes similar or different data.

Databases play a very important role in areas and organizations that use many computers allowing large volumes of company data to be stored, which are used through users, so the data must be well structured to be very easy to manage, plan, analyze and be able to make decisions within an organization.

1. Architecture

Currently some companies train their employees in the use of new software in order to make better use of the tools they have or plan to acquire in order to mitigate risks of loss of information or mis management and increase their productivity.

That is why when implementing a Data Mart (it allows us to centralize the information to take into account the flow of data on a large scale) we need to know the nature of the business to proceed to identify what are those sources of information that will allow us to select a database manager.

When developing an architecture to process large volumes of information, consider the following steps: Designing the Data Warehouse (Default Theme Oriented Data Collections) and Data Mart in order to establish what the flow of information and processes it will have throughout its extraction will be.

Contemplating the processes of extraction, transformation and loading (ETL) of the different sources of information, in order to have a broader view of the data stored in the Data Mart, technologies should be used that allow us to

perform analytical processing in order to obtain strategic results which will provide us with results to facilitate decision-making and contemplate risks. Having large volumes of information or that large repository of data needs to be reviewed how that information can be visualized, there are many ways to consider incorporating BI (Business Intelligence) one of them is to perform report generation, dashboard to review the graphical result of each of the processes to contemplate a greater compression of the information.

1. **What factors should be taken into account when solving a problem**

At a time when some of the processes that are implemented for the development of the system have some failure, it is very important to consider contingency plans to be able to control these errors since they can distort the proper functioning of the tool. For these reasons it is ideal to have mechanisms that allow the user to identify what is the procedure to follow, which can be generated through alerts or mailings.

For example: Information needs to be loaded through an Excel template, the app must read each of the template data, and when a data is not loaded correctly the system has to notify through an alert on which line and column the error was presented so that the user can perform validation of the type of data being loaded.

2. **Importance in the market**

The importance of database managers in the market has a big impact, because large organizations are implementing multiple cloud virtualization solutions, which generate high performance when obtaining information, good infrastructure security, reduced costs on software resources as it would not require many people to support or maintain on-premises servers.

In this great technological age it is inevitable not to think about decision-making for our organizations, so we are turning to business or business intelligence (BI) that allows us to analyze how our company works, to resort to solutions to improve processes that will help us to have better profitability, allowing the allocation of resources and plotting results statistics to calculate or assume what the results will look like in the future of the organization.

Companies to mitigate that the information they contain is not going to make them exposed to ill-intentioned people invest in cloud computing services allowing to solve their problems, where the providers of these services have a number of catalogs available in the cloud supporting the information, implementing security and audit mechanisms to detect or keep control of the operations carried

out by users in order to identify the responsible person who wants to obtain very sensitive information that can generate tax problems to the company, which would lead to the loss of its reputation and sanctions for leaving the information unprotected.

It is recommended in advance that all persons handling sensitive information do not leave it exposed to third parties in order to mitigate such events that are presented in organizations.

A digital country is globally communicated where customers and suppliers meet.

I. BUSINESS INTELLIGENCE

It's a company's data management to help facilitate decision-making by contributing to intelligence.

1. **Importance of business intelligence**

Within a company, business intelligence is very important because it helps people understand the figures or data shown to guide the best decisions in their business and also helps them improve the perspective they can take throughout their company so that they can increase their revenue. In addition, it relies on such data to generate modifications, eliminate inefficiencies and also to be coupled with any market changes.

A main goal for business intelligence is to increase effectiveness and efficiency within the entity, with the great advantage of this technology is that they can turn existing data into new ones. Likewise, BI is critical today as flexible self-service study is being prioritized, data is being digitized into reliable tools, training tool users, and agility to capture the most accurate information. Lately it has been shown how Business has generated an unexpected evolution due to the potential of the tools that companies manage because it helps them in their processes to summarize, gather and examine all the data of the operation in what is better their performance; this is where most organizations consider the following options:

1. Recognize how they can increase their profits.
2. Study the attitudes of your clients.
3. Check data from other companies.
4. Perform constant trace reviews.
5. Improving operations
6. Predicting triumph
7. Determine product ideas.
8. Discover or find the lack of business.

Examples of business intelligence

Tableau Software is a tool that detects incorrect data so that it is corrected by the staff in charge, so it has allowed you to save time, helping companies get a view of business, so you can make decisions for your improvement or be useful to your organization. Also to incorporate more developments without affecting performance, Anónimo (2013) indicates that the

following processes should be taken into account:

1. " Data Mining: Using Databases, Information Statistics, and Machine Learning to Discover Trends in Larger Datasets
2. Reporting: Share data analytics with stakeholders so everyone can draw their own conclusions and make decisions.
3. Performance metrics and benchmarks: Compare current performance data with historical data to track performance against goals. In general, this is done with custom dashboards.
4. Descriptive analysis: Use a preliminary data analysis to find out what happened.
5. Queries: The user questions specific questions related to the data, and the BI extracts the answers from the datasets.
6. Statistical analysis: From descriptive analysis results, data is further explored through statistics, for example, to determine how a trend occurred and why.
7. Data visualization: Transform data analysis into visual representations, such as charts and histograms, to more easily consume data.
8. Visual analysis: Explore data through visual storytelling to share information on the fly and stay in the analytics flow.
9. Data preparation: collect multiple data sources, identify dimensions and measures, and prepare data for analysis."

1. How business intelligence works

All companies have objectives and questions, but to detect it they must make observations, investigations and gather information or data so that such questions can be resolved and the desired goal can be achieved. In the technical aspect, the data is collected, processed and stored, at the time of storing such data, the area responsible may transfer to the information to be analyzed to the questions of the products.

1. How business intelligence work

In Business it should contain the study or research of business and data, although these only form a piece of complete development. Business Intelligence is enabling individuals to draw conclusions starting with data research. People who check the data thoroughly examine the specifics of the data. As they are doing, used statistics and predictive research in order to find present and future models. Data research asks or generates the question, why is this happening

and what might happen next? Business Intelligence verifies those pilots to also separate the effects in some easier way to understand. As indicated by Anonymous (2013) in Gartner's IT glossary, "business analytics includes data mining, predictive analytics, applied analytics, and statistics." Which means that companies generate business studies as an element of fundamental in business intelligence, because it has the ability to answer concrete clarifications, as well as provides a study to make decisions. Although many companies use or can use their study or research processes to continuously improve questions that arise in the course of follow-up.

Research or business study cannot be carried out immediately, for the reason, for the reason that, when answering only one question, it is time to rethink the process. Therefore, it is better to evaluate that the process is a data access cycle, i.e. innovations are made, this is inquiries and even exchange of information; what can be defined as a research cycle, but if it refers to a modern definition it can be said is how organizations use research to respond to the different changes in the questions that are happening in the process.

1. Difference between traditional and modern business intelligence

In previous years, all business intelligence applications handled a traditional standard, which involved a vertical technique, in which business intelligence was estimated by information technology organizations and almost all research questions were answered with graphical reports, i.e. a document could be generated where process analysis could be visualized as a graph; but this was not so favorable, because if at some point an individual had a question regarding the document taken, the requested information could be consulted to the last of the report, but that was not all, it was necessary to start the process again, which generated periods of reporting or documents taken and disappointing, this involved that the members of the companies did not take the time to verify the actual data to be able to make decisions. Despite the drawbacks, traditional business intelligence was used to perform frequent documents so that questions could be answered.

In contrast, modern business intelligence is participatory and affordable, however information technology organizations were critical to data management, but nowadays, even though users did not have sufficient knowledge to handle the tools, this was no longer a problem, because they could customize dashboards and also create documents immediately, then, having and using the right tool.

1. DATA STORE

It is a collection of thematic, embedded, non-volatile and historian data, organized to assist the process of toma de

Decisions. A data warehouse is a data management system model that is designed to organize and support business intelligence activities, especially analytics. Data within a data store is often derived from a wide range of sources, such as application log files and transaction applications.

Its analytical capabilities help organizations gain valuable business insights from their data to improve decision-making. A data store centralizes and strengthens vast amounts of data from different sources. Over time, it creates a historical record that can be of vital importance to data scientists and business analysts.

1. Components of a data store

Datastore databases: Central databases is the foundation of the data warehouse environment. RDBMS (Relational Database Management System) technology is implemented in these databases. As explained, the system that manages storage is another key element in a datastore. Regardless of whether the information collected in the data store can be analyzed using multidimensional visualization, the SGBD can be performed using Multidimensional or Multidimensional RelDatabase technology. Continuing this order of ideas are the comments that are made in the hardware section, for requirements of performance, scalability and technological consolidation.

As with hardware, modern relational database designs, post-relational databases open up a wider variety of choice. The result is that the same application installed in a post-relational database requires less hardware, can serve more users, and use strong mechanisms to access more complex data. Similarly, this technology supports combining the benefits of hierarchical and relational databases at a lower cost.

Here are four types of databases you can use:

1. Analytical databases are developed accurately so that data storage preserves and manages analytics. For example, Teradata and Greenplum.
1. Typical relational databases are the row-centric databases that you might use on a daily basis. For example, Microsoft SQL Server, SAP, Oracle, and IBM DB2.
2. Cloud-based databases can be hosted and rescued in the cloud so you don't have to get any hardware to set up your data store. Examples include Google BigQuery, Amazon Redshift, Microsoft Azure SQL.
1. Data storage applications are not exactly a type of storage database, but different distributors now offer applications that provide software for data and hardware management to store Data. A few examples, IBM Netezza, SAP Hana, Oracle Exadata.

Similarly, this technology supports combining the benefits of hierarchical and relational databases at a lower cost. Generate integrated management of existing data warehouse and Data Marts, integrating extraction, transformation, and loading for corporate data warehouse construction and Data Marts. Planning, logs, interfaces to third-party schedulers, which allows to carry out a management of the planning of all the processes indispensable for the loading of the DW. Support in the exploitation of the data warehouse. Sometimes enough attention is often paid to this stage of data warehouse

management even though it is a big part of the effort in building a data warehouse.

-Metadata: these are data about data that the data store requires. Used to form, maintain, and manage the data store. Metadata allows a person to locate and understand the data, including information needed to determine which datasets exist for a particular geographic location, the information required to determine whether a dataset is appropriate for specific purposes. It can be characterized in two types:

Business metadata, which contains information that allows an easily understandable view of the data stored in the vault.

Technical metadata, which contains information that developers and administrators can use when running warehouse development and management tasks.

Metadata plays an important role for both businesses and technical teams to understand the data present in the warehouse and transform it into information.

-Query tools: Query tools allow users to interact with the data storage system.

Bus datastore: Specify data traffic in your store. Data traffic in a datastore can be organized as inbound traffic, upstream traffic, downstream traffic, outbound traffic, and meta flow.

A data mart is a level of access that is used to transfer data to users. It is used to partition the data that occurs for the particular user group.

-Data stores offer the universal and unique benefit of enabling organizations to analyze large amounts of variant data and extract significant value from it, as well as maintain a historical record.

1. Data warehouse features

Four unique features allow datastores to deliver this universal benefit. According to this definition, datastores are:

1. Topic-oriented: Data is sorted by topic to facilitate end-user access and understanding. It is possible to analyze data on a particular topic or a functional area.
2. Integrated. Datastores create consistency across different types of data from different sources.
3. Non-volatile: Once the data is in a datastore, it is stable and is not erased when new data is entered. The displayed data is read-only and updated periodically. Activities such as deleting, updating, and inserting activities that are performed in an operational application environment are ignored in the datastore environment.
4. Variants over time. In the data warehouse analysis process, it analyzes changes over time. Contains a time element, explicit or implicit. One of those places where the datastore data display time variation is found in the conformation of the registry key. Each primary key contained with the DW must implicitly or explicitly count a time factor.

The following properties are indispensable for the architecture of a datastore:

1. Separation: The analytical and transactional process should be kept separate as much as possible.

1. Extensibility: The architecture should be able to host new applications and technologies without having to examine the entire system.
2. Manageability: Managing the datastore should not be excessively difficult.

1. B data storage efficiencies

Now that we know how a data warehouse works, it's time to learn about the benefits of a data store and how exactly it can help your business approach a deadline. All that needs to be done is to use your data store to get data in seconds.

Data storage optimizes data quality: Refined data quality helps ensure that your company's policies are based on accurate information about your business efforts.

With data collection, you can transform data from multiple sources into a shared array. This way you can ensure the security and quality of your corporate data. This way, you can identify and delete replica data, improperly recorded data, and erroneous information.

Data storage improves business intelligence: You can use a data store to gather, assimilate, and derive data from any source throughout your business commitments. As a result, your BI will optimize in leaps and bounds, due to the ability to effortlessly solidify data from a variety of sources.

examples include cross-references from numerous data banks can be difficult and sometimes impossible to manage. But with a data warehouse in place, everyone on your team can have an integrated understanding of all important information in a timely manner to grow and scale.

Data storage saves time: In the modern and fast world of fierce competition, your ability as a company to make quickly refined decisions is essential to outperform your opponents.

A datastore gives them access to all their required data in minutes.

Business or executive participants won't have to wait until IT produces reports, and IT will focus on its primary goal of keeping the business running.

Data Quality and Consistency: Data Warehouse includes converting data from numerous source systems to a common format. Since each data from the different departments is standardized, each department will generate results in line with all other departments.

Historical Intelligence: Accumulate large amounts of historical data to analyze different time periods and trends in order to make future predictions. Data typically cannot be stored in a transactional database or used to generate reports from a transactional system.

1. Layers of a datastore's architecture

Logs from a web server, with user browsing data. Internal market research data. Third-party data, demographic data, survey data, or as census data. It is feasible that some minimal data, but there is not expected to be a transcendental data transformation.

Having a common area facilitates the process and subsequent unification of the data. This layer is also where the data is

polished. The ETL design phase is often the most delayed phase in a datastore project, and an ETL tool is typically used at this layer. In any system you can find only one of the 3, 2 of the 3, or the three types joined.

These business rules do not affect data transformation rules, but they affect what you can then display in reports. This can be done in the form of a tabular or graphical report through a browser, a mailed report that occurs automatically and is sent daily, an alert that warns users about exceptions, and so on. Usually this layer uses a reporting tool and OLAP tool. A logical data model would be a clear example of something that is in this metadata layer. A metadata tool is commonly used to manage metadata.

1. Data warehouse types

The main goal of the single-tier or single-layer architecture is to reduce the amount of data stored. This goal is to eliminate data repetition. This architecture is not used regularly in practice.

1. Two-tier architecture: The two-layer architecture divides the physically available fonts and the data store.
2. Three-tier architecture: Data is cleaned, transformed, and loaded into this layer using back-end tools. For a user, this application level exposes an abstract view of the database. This layer also proceeds as a mediator between the end user and the database. The top level is the tools and APIs that connect and extract data from the data store.

These could be query tools, analysis tools, reporting tools, managed query tools, and data mining tools.

Designing a data warehouse

Any datastore design must contain the following:

1. Specific data content
2. Relationships within and between data groups
3. The scope of systems that the datastore will support
4. The types of data transformations that are needed
5. Frequency of data refresh

One of the main factors in design is the needs of end users; they are interested in performing analysis and viewing data together, rather than individual transactions. However, end-users often don't really define what they want until emerge una necesidad específica. Por consiguiente, el proceso de planificación debe adjuntar suficiente investigación para anticipar las necesidades. Finalmente, el diseño del almacén de datos debe soportar la expansión y la evolución para mantener el ritmo de las necesidades cambiantes de los usuarios finales.

Enterprise DataWarehouse—This centralized data warehouse integrates all of an organization's information so the entire enterprise can have revenue. It offers a service in which it supports decision making in the company.

Operational Data Store (ODS). Also known as ODS, it is a data store, which when data storage and OLTP systems do not accept the needs of organizations' reports.

Data Mart: un Data Mart is a subset of the data warehouse oriented to a specific area, also known as a departmental database. It is specially designed for a particular line of business, such as sales or finance. In a separate Data Mart, data can be collected directly from sources.

1. DATA WAREHOUSE

1. Components of a data warehouse

Operational data sources: This component is created from the capture of business data it should be noted that the capture of this data depends on the organization of the business since it is not very good to fill the data warehouse with random data, this data should be extracted and organized by sectors in order to have better results in response time, the data must be in a format that is supported by any business intelligence analysis. These fonts can be formed from operational system, erp, cmr, or plain text file etc.the only thing that is important is that these fonts can be extracted.

Extracción y transformación de datos: Es la responsable de que los datos se muevan para poder ser extraídos y transformados por medio de las ETL ya que los data warehouse a través de sus data marts implement a model either star or snowflake because it originally does not have these models and this data is loaded.

Access tools: Para to be able to access and analyze this data we must select which tool best suits the need of the business among them we can find:

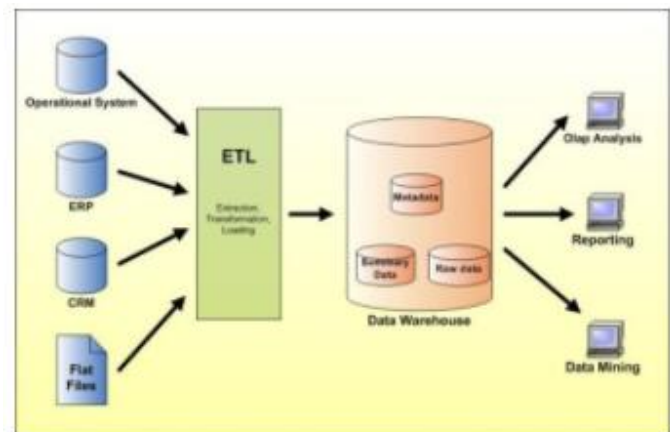
1. consultas SQL (structured query language)
2. herramientas MDA(Multidimensional Analysis)
3. herramientas ROLAP (relational on-line analytical processing)
4. Data Mining tools.

It is necessary to have a good strategy and way to capture important data and in an agile and secure way to be able to analyze the different points of view and that the information obtained is of great benefit to the business and does not become an amalgam of data without utility.

Repositories /metadata:The metadata repository and its design are basically the importance for the success of a good data warehouse, although in development projects they do not give much importance to this section. Its importance is in that stored in this repository, and from here that metadata is responsible for guiding the extraction,

cleaning and loading of data within the data warehouse, in addition to helping query tools and informes generators work properly.

Arquitectura de data warehouse:

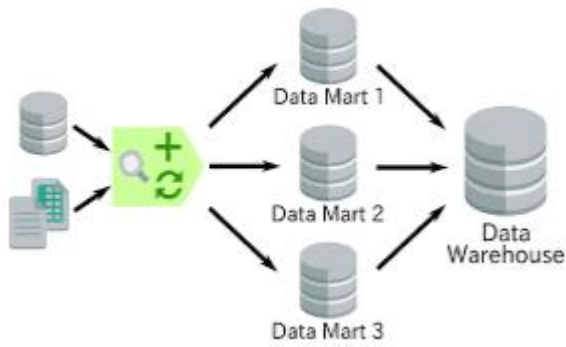


https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.webyempresas.com%2Fque-es-un-data-warehouse%2F&psig=AOvVaw0bks84Bz2A5FdDPYQ12Bcc&ust=1614032441046000&source=images&cd=vfe&ved=0CAIQjRxqFwoTCIify9aB_O4CFQAAAAAdAAAAABAE

▪ Data marts

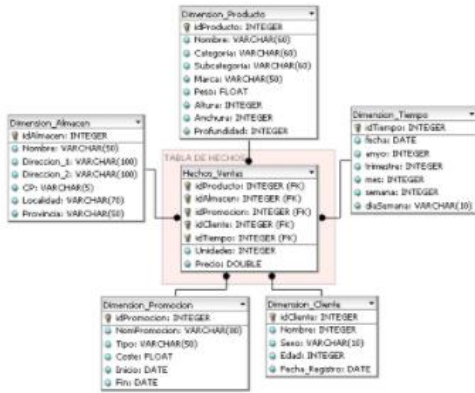
This structure consists of fragments called data warehouse derivatives known as data marts in turn these store models of information such as star and snowflake these models serve to exploit different purposes such as:

1. generation of reporting.
2. analysis of information.
3. dashboards.
4. data mining.



https://www.google.com/imgres?imgurl=https%3A%2F%2Ftroyanx.com%2FHefesto%2Fdata_mart_bottom_up.png&imgrefurl=https%3A%2F%2Ftroyanx.com%2FHefesto%2Fdata_mart.html&tbid=Avj2At0yV1B_mM&vet=12ahUKEwi3yJ6DgvzuAhXRrVkJHSJ6ApYQMygPegUIARDEAQ..i&docid=AEeVDQJJGDmOOM&w=924&h=500&q=data%20marts&safe=active&ved=2ahUKEwi3yJ6DgvzuAhXRrVkJHSJ6ApYQMygPegUIARDEAQ

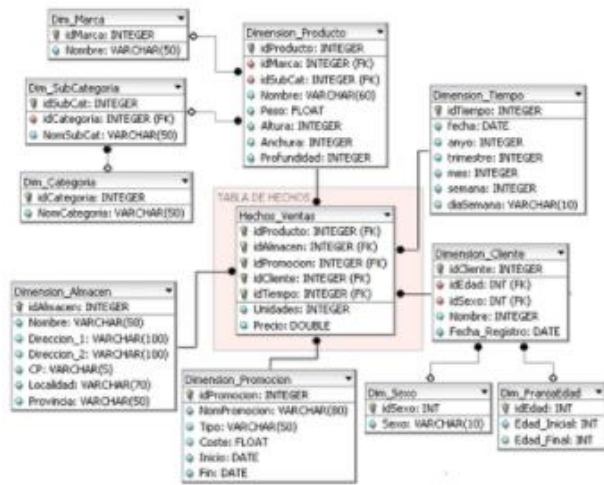
Star model: The star schema separates the data from the business process into: facts and dimensions. Facts contain measurable, quantitative, business transaction-related data, and dimensions are attributes that describe the data indicated in the facts (a kind of metadata, that is, coughing up other



data).

https://upload.wikimedia.org/wikipedia/commons/0/00/Esquema_en_estrella.png

Snowflake or snowflake model: it is a somewhat more complex structure than the star scheme. Occurs when any of the dimensions are implemented with more than one data table. The purpose is to normalize tables to reduce storage space by eliminating data redundancy; but it has the trade-off for generating worse yields by having to create more dimension tables and more relationships between tables (JOINS) which has a direct impact on performance.



https://upload.wikimedia.org/wikipedia/commons/thumb/2/2c/Esquema_en_copo_de_nieve.png/480px-Esquema_en_copo_de_nieve.png

1. Data warehouse factors

Data warehouse success factors: The main factors that lead to success in the data warehouse are:

1. Integrates production data with external data and manages histories.
2. Contains useful data.
3. The data is consistent, up-to-date, and documented (quality).
4. Provides direct access to users.
5. Increases the number of accesses.
6. It provides flexibility that supports the growth of users, tools, as well as the volume of data.

Implementation issues

1. High cost
2. Immediate decision-making assistance (Data Mart).
3. High maintenance cost due to changes in needs, new data sources, change of capacity or technology.
4. Data quality control.
5. Heterogeneity and data integration.

1. Differences between a data warehouse and database

They are polar opposites in terms of design and operations: The database are transactional systems designed to execute high, low or data change transactions, such as a name change or change of balances or withdrawals and the data warehouse is organized based on concepts such as customer product sales types and times there is difference in designs while the database are extremely standardized a data warehouse tends not to standardized data in wineries as in the data marts and star model.

What is a data warehouse:

To. The data warehouse is not much less a branded software or just a database.

B. Data warehouses must adjust to changes as we are going through an era where information or data is very volatile since the most important thing of companies are their data from each area.

C. You also need to be prepared to load massive amounts of data in a short period of time as these information systems are designed for information analysis

D. For this reason it is not desirable for the data warehouse to live together in the same environment as our transactional systems such as ERP or point of sale since it may lower the performance of our servers and you may run the risk that when requested information will not deliver it in an agile way taking into account that the data warehouse can take up to days to resolve only one request.

E. The natural of the data warehouse must be multipurpose.

II. BLOCKCHAIN, BITCOIN, IoT

1. **How many of us know about Blockchain?**

I believe this Mr. "Satoshi Nakamoto" I put him on the line because no one really knows who Tozzi is ok, it is not believed that with aly, it is not known whether a person a group of people the Corporation actually made him in 2008 in a proposal to a few forums and in 2009 the first implementations of Bitcoin appear. Bitcoin is a digital currency that has a value that fluctuates in supply and demand right now the value of the 15 million disks in wanted circulation are about 5200 million euros which is a lot of money and how interesting Bitcoin is not what the money itself represents, The really interesting thing about Bitcoin is what you have just below is the technology with which the Bitcoin. Because this technology is going to allow us to live a new revolution. We live the Evolution of the web because the technology on which bitcoin was built will allow us to live a new revolution that is to bring the data to people, This technology is called blockchain and basically is a unique record of all transactions that are made with Bitcoin, it is something like a ledger of an accounting seat , where all transactions are targeted in such a way that in this digital account book; what is a database where we have stored all the transactions that have been made with bitcoin since 2009, Very quickly for them to see it interests me is to see how blockchain can affect us in the future. If you did not explain this as it may sound a bit Chinese, the blockchain base is the transaction and has several parts one is an address that is unique within the blockchain network; A bitcoin address, which is unique to the person issuing the transaction a unique address from which the transaction will receive the amount of Bitcoin that we want to send to that person and the digital signature and is in the part really Important in all transactions that are made in bitcoin, are signed say the recipient of the

transaction knows who is sending you that transaction in addition to has a dual function, which is that the bitcoin network can validate all transactions. I'm going to set an example. What Bitcoin would look like is one in a bitcoin network the first thing is that the Bitcoin network and is made up of machines by nodes in a bitcoin network, a node is anything with compute and storage capacity i.e. from a Tablet to a server to the PC of our house to the accelerator cards the accelerator graphics cards, That many people buy to be able to work on a bitcoin network the important thing about this type of networks is that they are decentralized there is no node that takes care of managing the information, and this is a rather disruptive part especially because in the information in which we are immersed we are used to centralized models to connect to Facebook and send the photos to Facebook or when we are sending a WhatsApp , We send the message to WhatsApp because everyone here knows that when, I sent a WhatsApp message to someone does not go directly to your phone, right go WhatsApp servers. This is important because the data really of that communication is whatsapp has as in bitcoin is not central server the data is distributed throughout the network. So what we have is a network of nodes in which there is a copy on all blockchain nodes the whole book of transactions and charity and is that since there is no central node, anyone can enter to participate and they are very important because they give us a lot of freedom to sign up or unsubscribe, applications or nodes or people if we wanted to make a transaction between these two people in the blue T-shirt boy and the one of the tie I think would build the transaction and we will connect to a bitcoin node, No matter we can have one of the local we connect that node and send it the transaction, At that moment two things happen the node receives the transaction and stores it then takes the transaction and sends it the nearest nodes in such a way that our transaction goes through all the nodes of the network , at any given time are seconds that transaction is available on the recipient's nearest node. If the recipient asks this node if the transaction is for him he will see that there is a transaction of the blue T-shirt boy and right now he knows the boy has traded him X Bitcoin. But on the other hand. That transaction has been validated by all those nodes, by all the nodes that I pass through allows us that in case of fall of half the network the network continues to work things that with the centralized model on which we are right now in most of the applications that work and Facebook falls Facebook falls Twitter falls Twitter here when a local server falls, all the nodes have to be dropped and yet, even if there were a single node above would continue to work. The problem is not going to be advantages of decentralized networks at all because not everyone would be with the decentralized network. In a centralized model all the data all our Facebook photos are on Facebook servers, In a decentralized model we have to know how to manage that dataset.

Then in bitcoin I solve this problem you must manage the information. Thanks to blockchain and that it is blockchain, I said before that it was a kind of book, where we were pointing all the seats of entrances and exits; is a chain of transaction blocks to go to the interesting part of the future when a transaction arrives a node that node Try to build a block with

all the transactions you are reading and all nodes in the network have to solve a mathematical challenge, the first to solve that challenge, its block is part of the blockchain and the rest is discarded. Once the network has identified the winner of the challenge Bitcoin occurs every ten minutes that block is transmitted over the network in such a way that all copies of the blockchain of all nodes that are distributed across the network have the same copy the same data set of all transactions since 2009, In this way we guarantee the integrity of the information that the blockchain is handled by another interesting part of blockchain difficult to manipulate because if someone tried to modify a transaction on any of those blocks it would be very easy, that we could go to another node and compare it remember that there is another copy there on all nodes of the network or on most nodes , And if you modify the block, There would be another problem that would break the string, you could quickly identify it to be manipulated, Then we have a network that is distributed that resistant to data manipulation and in which the authorship of the transaction is fully identified by the sender and the recipient knows, who has sent it the transaction: This may be interesting for people who have bitcoins, maybe yes, but really most would think I have my money in the banks and I am not interested, basically because well, this is a bunch of geeks who are moving people's money, the really interesting thing about blockchain, is that if we remove the concept of Bitcoin from currency. We incorporate the data the concept of information data we have a tool, tremendously powerful to be able to communicate with machines against people a tool that would allow us to migrate to a decentralized model in which the information is ours, it is not from a company that is giving us a service, and this is a very important thing that will affect us in the future right now we are used to someone's lots of service that is interesting an instant messaging and how it is free and interesting, because I start using it and sending data.

Our information is being used by companies for anything, I'm not saying make money, I say anything, You really want our information to be there good for now at the level where we're on what information we share it are photos and holiday comments, because it can give a little the same, but in the future they may not leave us the same , So you can see that it does not eat geek project there two very important projects is the hyperledger Project which is a consortium of companies of the size accentuated ogame, they are looking to explore solutions on blockchain and the other is R3 is even more interesting because R3 The consortium of the 40 largest banks in the world that are also looking for solutions to certain problems of their financial processes thanks to blockchain , the fun part of the future is coming. There are four areas left in which we could apply blockchain and in which we would solve certain problems that we will find from here to 10 years the first health when I talk about Health I am referring to the information that manages the health system in which we are about us, Right now if you go to an old hospital do a test this test is stored on the servers of the hospital if you go to another hospital and do other tests because servers of that hospital are stored the reality is that our medical record right now is fully distributed throughout the places where you are going, there is no medical record, unless you order all that information and

keep it in the drawer, not for when you go to the doctor. The future of healthcare requires the information handled by health facilities, doctors and all people working in healthcare to ensure that I can rely on privacy and also that it is agile and would allow us to have our medical record stored in a network scattered around the world that that information would be encrypted so that access to information would be limited and I would decide who can see my medical record.

What are you going to do some testing on a TAC And instead of going to collect the results those that the machine sends the test to a blockchain network with the encrypted image where we are the recipient gives you shape that all the medical information of ours would centralize in our virtual file does not tell me which only we have access to. The other area of the Internet of Things to interconnected things is the initiative that can manage a data connects the internet challenge of things that will be thousands and billions of devices with the internet. Connecting billions of devices has a problem, that the centralized model will not support it, We can not have something similar Facebook, of washing machines, I want the washing machine to communicate directly with me with a blockchain network the washing machine would send a message to the network would be stored and that I could see it there is no example of how blockchain will help the internet of things very simple, I think it's in a document I don't remember right now and it was to imagine what that machine you sell that's in the office, not the one you're going to and to buy a coffee imagine that this machine that sells in the future accepts bitcoins, for example you go and pay for coffee. And the machine already has a balance has a balance of that blockchain network on a balance in its favor imagine that the machine, now take your balance and buy the products directly from the suppliers, the supplier would come put the product in and start over, it does not shape that the machine is independent that for machine manufacturers, because it would be an excellent one, because it would save a lot of time, because the typical thing is to throw the money and you do not know that the milk does not come out, Or the coffee is exhausted and the third scope is the management of digital goods. Everything is increasingly digitized, more computerized and things that we used to do how to go to the cinema with an input with a paper stand, because it is all in digital, the problem of digital is that it is very easy to copy. Having a system in which I am guaranteed that I am the person who owns that digital good, because it is necessary and blockchain is a solution to be able to register, that I bought a cinema ticket or bought an airfare, there are a thousand options in which we could apply blockchain, From food control proof of stock authentication in services and Open Data or audit and well also with financial transactions.

1. CONCLUSIONS

By implementing BI technology tools it allows us to optimize the different processes that are carried out in companies. These will support us in decision-making where we implement an organizational structure of the data which facilitates the management of them.

Business intelligence is a tool that makes the right information available to users in the right place. There are multiple advantages it offers companies, including for it to be considered complete it must bring together the following components: data mining, agents and data warehouse. is the generation of a competitive advantage.

APRENDICE

Knowledge is generated about all the intelligence applications and the procedure that a company should govern, to identify the best methodology or processes that a company should use.

REFERENCIAS

- [1] Luna Ramírez, Enrique; García Merayo, Félix. (08 Julio 2000). EL REPOSITORIO DE METADATOS EN UN DATA WAREHOUSE. Chile. Revista Facultad de Ingeniería. <https://www.redalyc.org/pdf/114/11400802.pdf>
- [2] (21 septiembre 2020). Esquema en estrella. WIKIPEDIA La enciclopedia libre. https://es.wikipedia.org/wiki/Esquema_en_estrella
- [3] (7 junio 2020). Esquema en copo de nieve, WIKIPEDIA La enciclopedia libre. https://es.wikipedia.org/wiki/Esquema_en_copo_de_nieve
- [4] De los Ángeles Ibarra, María. (4 octubre 2016). ABC del Data Warehouse. Evaluando Software.com. <https://www.evaluandosoftware.com/abc-del-data-warehouse/>
- [5] (06 junio 2017). ¿Qué son los procesos ETL? PowerData. <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/qu-son-los-procesos-etl>
- [6] Candal Vicente, Isabel. (14 noviembre 2011). Modelo de éxito de una data warehouse. Scielo. <http://www.scielo.org.co/pdf/tecn/v17n35/v17n35a11.pdf>
- [7] Power data, (9 marzo 2019), obtenido de: <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/que-es-un-gestor-de-datos-y-para-que-sirve>
- [8] Marín Rafael. (16 abril 2019), obtenido de: <https://revistadigital.inesem.es/informatica-y-tics/los-gestores-de-bases-de-datos-mas-usados/>
- [9] (2021). ¿SABES LO QUE ES EL DATA MINING? Madrid-España. Universidad de Alcalá. <https://www.master-data-scientist.com/que-es-data-mining/>
- [10] IBM, Obtenido de: <https://www.ibm.com/cos/analytics/database-management>
- [11] Instituto FOC, (19 agosto 2019), obtenido de: <https://www.foc.es/2019/08/19/5104-principales-gestores-de-bases-de-datos-de-la-actualidad.html> Definiciones y respuestas. Disponible: <https://www.astera.com/es/type/blog/data-warehouse-definition>
- [12] uoc.edu, obtenido de: <https://www.uoc.edu/pdf/masters/oficiales/img/913.pdf>
- [13] deyde data Quality, obtenido de: <https://deyde.com/news/que-es-un-gestor-de-datos-y-cuales-son-sus-funciones/>
- [14] Copyright 2007-2021- Sinneexus https://www.sinnexus.com/business_intelligence/datamining.a.Spx
- [15] Llordachs Marqués, Frederic. (08 septiembre 2014). ¿Qué es la data mining? La definición de la minería de datos. Clinic-Cloud. <https://clinic-cloud.com/blog/data-mining-que-es-definicion-mineria-de-datos/>
- [16] Tuatara Tech, Nubia Márquez. MINERÍA DE DATOS (DATA MINING) VS GRANDES DATOS (BIG DATA). corebi. <https://corebi.com.ar/blog/mineria-de-datos-data-mining-vs-grandes-datos-big-data/>
- [17] Omar Sosa Marcelo, Sosa Bruchmann Eugenia Cecilia. (2014). ESTUDIO DE TECNICAS DE DATA MINING APLICADAS AL ANALISIS DE DATOS GENERADOS CON LA METODOLOGIA BLENDED LEARNING. Workshop de Investigadores en Ciencias de la Computación. http://sedici.unlp.edu.ar/bitstream/handle/10915/41977/Documento_completo.pdf?sequence=1&isAllowed=y
- [18] Puente Burbano, Fabio Edmundo. (2005). Repositorio de Tesis de Grado y Posgrado. <http://repositorio.puce.edu.ec/handle/22000/1104?show=full>
- [19] Ana G^a Huerta. 25abr 1997 <https://www.computerworld.es/archive/teoria-y-practica-del-data-mining>
- [20] Ivan Jose Turmero Astros. Marzo - 2020 Como Gestión del conocimiento. <https://www.monografias.com/docs112/data-mining-como-gestion-del-conocimiento/data-mining-como-gestion-del-conocimiento.shtml>
- [21] ORACLE.COM. Qué es un almacén de datos. Disponible: <https://www.oracle.com/co/database/what-is-a-data-warehouse/>
- [22] Tecnologías informacion.com. Datawarehouse: Tipos, Arquitectura y Características.

Disponible: <https://www.tecnologias-informacion.com/datawarehouse.html>

- [23] Nida Fatima. (febrero-2021). Qué es un almacén de datos.
- [24] Nida Fatima. (febrero-2021). Arquitectura de almacén de datos, componentes y conceptos. Disponible: <https://www.astera.com/es/type/blog/data-warehouse-architecture/>
- [25] (2003-2021). ¿Qué es la inteligencia de negocios? Guía sobre la inteligencia de negocios y por qué es importante. TABLEAU SOFTWARE, LLC. <https://www.tableau.com/es-es/learn/articles/business-intelligence#:~:text=Una%20BI%20eficaz%20ayuda%20a,dentro%20de%20su%20contexto%20empresarial.>