

CREATING A BUSINESS AND SUPPORTING DIGITAL TRANSFORMATION

Ayala, Castillo, Miguel Angel, miguel.ayala@cun.edu.co, Jiménez, Bohórquez Luis Rodrigo, luis.jimenezb@cun.edu.co, Martínez, Muñoz, Sergio Andrés, sergio.martinezmu@cun.edu.co, Rojas, Penagos, María Cristina, maria.rojasp@cun.edu.co, Portella, Cleves, Jorge Enrique, Jorge_portella@cun.edu.co

Introduction

Many times, has happened in your company organization that needs to answer questions and make a reminder of what has happened in it since its inception and during the time it has been in operation and has wondered! Why hasn't she emerged or evolved as she pleased? or worse, what were the reasons or why did events that left her lost happen and know what led her to this point and be able to have a vision for the future where you can be expected to pass, that allows you to mitigate risks and not fall into the same mistakes of the past?

Cloud-based applications and infrastructure dominate the IT industry. Most of these applications are designed for use in certain industries. This new expansion of digital applications and services describes new minimum competency requirements for each segment.

The impact is like the beginning of server creation over the past 20 years and cloud computing over the past 10 years. By 2021, vendors will adapt to these changes in cloud and application products by creating solutions that meet the growing demand, scale, and speed of container services.

But what is this container app like? Different types of storage will be required to meet many needs, such as boot partition and logs, transaction data, traditional file application data and new APIs, and files. New container storage methods are being developed to collect Container Storage Interface (CSI) data and store items and backups.

Storage elements will be the first integrated solutions for analytical applications such as Splunk, Vertica, Weka, Elastic, Cloudera, and Spark. This is due to five reasons. The purpose of this application is to use the data directly using a protocol that supports Hadoop as S3A. This allows you to use the well-known AWS S3 APIs.

Unstructured formats include drawings, sounds, PDFs, CAD files, and more. They naturally fit into object storage but are very difficult to manage in the data. Unlimited (non-hierarchical) flat namespaces allow analytics applications to easily manage billions of objects at an almost unlimited level without worrying about directory limitations. The ability to distribute storage tiers at the application level, unlike

traditional systems, is all highly integrated and driving them at the same speed. And finally, features liveblog wear various things and managing forms is an important way to protect stored data from a cybersecurity attack.

Managing hybrid cloud data will be acceptable by providing the best benefits for the public cloud business (on demand, scalable service, speed) and utilities (privacy, security, performance, and control). For data-intensive applications, the benefits are significant by reducing time-to-market, improving access to data, and reducing investment and labor costs.

Data usage ranges from cloud data storage to data exploitation and business permanence (especially disaster data recovery).

Disaster recovery (DR) in two virtual datacenters is no longer required. With an enhanced cloud DR solution that manages synchronized copies of key data at home and in the cloud, IT executives can reach millions without the cost of maintaining and maintaining two remote locations in DR. You could save thousands of dollars without it.

The basic cloud application uses transactions with storage through APIs. The physical model stores objects through interfaces such as the AWS S3 API.

Finally, complex, and distributed applications are deployed in existing cloud environments, centralized data centers, and on-premises endpoints and containers, which carry service connections to integrate and protect workloads. Reliable communication between these services and the ability to manage storage services is increasingly challenging.

Take a "service connection" approach to protect networks and ensure that the ability to access operational services and connectivity increases. Unlike traditional authentication certificates that have been in operation for a long time, we are now turning to short-lived accounts and additional authorizations. This will gradually introduce a "zero knowledge" network (via the Google Beyond Corp platform) that will allow you to design network policies and use/encourage them in a more orderly and consistent way.

In the coming years, digital applications and services that continue to use the country's native cloud systems will be

huge. By 2023, that will exceed 500 million, according to IDC. This corresponds to the sum of all applications developed in the last 40 years. If you are the one you answered, yes!

This article is for you!

Problem

This article will explain the importance of leadership in organizations, the future of data storage systems, data warehouse business intelligence and the properties and application of Data Mining, its stages, mining techniques, their types, as well as their standards, aimed at formalizing good practice for data mining in any area, scenario, or requirement to which it can be applied.

Business Intelligence (BI) leads us to understand everything we need to enable in our company or organization to make it smart in the face of our business needs and for this we need different elements, technologies, and different methodologies.

This allows organizations to collect information or data, which we can debug, visualize, analyze until we can make decisions, execute changes, or manage the execution of those managements and thus make them public in the organization.

When we talk about Business Intelligence (BI) we talk about ERP systems, transactional systems that help us manage the globality of resources in our organization.

This Article deals with the importance of implementing these technologies in companies or organizations, and we will primarily address the issue of Datawarehouse.

This article is divided into 5 parts including this introduction, also found in section2 which is the Methodology, section 3 gift, also section 4 where we find the conclusions on the subject and we will know a little what the Datawarehouse means and finally and section 5 references.

THE IMPORTANCE OF LEADERSHIP IN ORGANIZATIONS

Leadership has been very important in organizations over time and human existence.

At the beginning of the twentieth century the first empirical study related to human leadership has been published, this was because of the interest that had to be able to visualize and identify the characteristics of the human being in the face of this issue to take administrative positions.

After the first world war, industrialization was achieved by leading to new leadership in bureaucratic and business

activities, in most organizations those in management positions come to them for having some link especially with senior managers which leads to the gain of these positions on their own merit and that is where the first studies were directed to the psychological traits of innate and/or effective leaders; characteristics such as intelligence, will, sociability and conditions of authority were marked but over time they did not have the expected relevance.

A second study was given so to speak, the double-factor theories where the behavior of the leader is studied and self-authoritarianism and democracy are considered in which the chief gives a little more freedom to his subalterns to make decisions, that is, to be a little more autonomous. Someone was very interested in these studies and gave rise to different studies on the subject of leadership, this person was Elton Mayo, but all these studies were based on the degree of authoritarianism and democracy "authoritarianism-democracy" and the satisfaction of his subordinates but these studies proved contradictory as time went on as a good relationship could not be defined between the behavior of a boss with his subordinates and the performance obtained by subordinates because of the behavior of the same boss, as performance always varied according to different factors of the entire team.

Between the 1960s and 1970s the scientific revolution, since all workers had to have more defined and strong characteristics in terms of knowledge and knowledge in order to perform and manage new technologies, in addition to leaders demanded to have an orientation towards human resources; motivation, delegate functions, etc., to this style of leader and its characteristics was called situational leadership which took the day-to-day as an element that influenced the effectiveness of their leadership.

Importance

Every organization depends on the leadership of its leaders to strengthen and move forward with the passage of time firmly and robustly, there are four conditions that a leader must have: commitment to mission, vision communication, self-confidence, and personal integrity. Although the leader with these conditions seems to be complete there are also virtues that will lead him to make very good decisions when he needs to make them; prudence, temperance, justice and strength with these virtues the work of the leader is not completed anyway as the communication he has with the rest of the team is very important to make his underlings bring out the best in themselves and there is a fluidity to achieve a good goal allowing integration and thus achieve efficient and coordinated actions.

In essence the leader is forging himself day by day and we can say that it becomes a way of being of the same person that is strengthening with the passion of the mission, in the action and in the fundamental values, the leader with this must know how to delegate functions, know how to grant power to others, have humility and consider himself a person within an organization as an some that can be replaced at any time.

In this order of ideas, we can consider leadership to be a process of integration by directing the team by influencing it with the good characteristics that are available to achieve a common purpose or a goal for the benefit of all by transforming the people involved in this process. We can also say that in an organization if you do not have a good organization, but if a qualified staff at the helm of a leader can get a company out of the bottom.

LEADERSHIP STYLES

Every leader brings business goals to great achievement, enforcing the organization's requirements, leaders have shown different ways they conduct their activities daily; there are three basic styles we can identify within leaders: the authoritarian leader, the participatory leader, and the liberal leader:

The autocratic leader assumes full responsibility for the project, organizes, contributes, and executes together with his underlings, but considers that he is only in that position because the only one who handles the subject and without it will not be able to make the project a reality.

The participatory leader does not delegate anyone, but investigates, investigates, and considers the opinions of others to make on their own a final decision to be executed.

The liberal leader delegates responsibilities in its entirety to its subalterns and waits for each to execute and complete the project on its own.

LEADERSHIP THEORIAS

Leadership takes 2 forms:

- 1) Formal leadership: It is practiced by persons elected or appointed to such formal positions to exercise authority.
- 2) Informal leadership: The one exercised by people who have characteristics or abilities to exercise it.

John Schermerhorn defines leadership like this:

- A) Perspectives of trait theory and behaviorist.
- B) Situational or contingent perspectives.
- C) New leadership prospects.

The theory of traits refers to the fact that there are traits that make the difference between a true leader and the one who is not, studies have been carried out previously and the traits have not been very well determined for this; more recent studies have resulted in some of the qualities that leaders have are: stress tolerance, power motivation, emotional maturity, self-confidence, integrity, among others, behavioral theory is based on the conduct of the leader studies have been conducted at the University of Michigan and Ohio University which consisted of interviewing the administrators of different companies to determine their conduct as a leader; at the University of Michigan they showed leaders interested in business production and staff, they were also seen to have a very productive group, just as in Ohio it was noted in the

results of the study concluding that a good leader must have the two behaviors presented for a satisfactory balance to exist.

Situational perspectives refer to the fact that the leader's traits and behaviors may interact with circumstances or situational contingencies because depending on the situation presented a leader may have a more pronounced behavior and the way they will depend on their reaction and will also depend on when they must react to a contingency. In several circumstances the traits have a lot to do with the results or in being able to distinguish who is a leader and who is not.

The new leadership perspectives developed by Bernard Bass (1990) are those of transformational leadership and transactional leadership.

The transactional has four instances; charisma, inspiration, intellectual stimulation and individualized consideration, the charisma gives vision and sense of mission, inspires pride, as well as respect and trust of the person who follows it. One of the characteristics is that it considers that the worker is autonomous and can develop and perform in different circumstances, as well as bring out the best of himself, causing the interests to flourish with the group and not for himself, this type of characteristic of a leader exchanges prizes for the effort of his subalterns making their motivations for them and raising productivity and fidelity to the goals set achieving them indisputably, but in turn making it a momentary negotiation until the goal is achieved.

ORGANIZATIONAL PARADIGMS

Every organization or company must be in constant learning not only to remain sustainable and solid but also so that in those times of difficulty survive and that same experience remain stronger.

From this arise certain unknowns such as what we should know about: A customer, a product, the business market, technology, social, economic, and cultural impacts, the relationship between the market and the product. If we talk about business learning, we notice the following points:

- New knowledge, skills, values, or behaviors are created or acquired over time.
- What you learned becomes the property of some collective unit (in a topical or explicit way).
- What you learn stays within the group or organization, beyond the physical and individual presence of people.

And again, concerns arise because the organization can do will depend significantly on the way it is seeing itself through the lens s of old or new paradigms; then will the company be learning from its current process? How is it doing? What does it do? and if it does not how to make our company learn in a stable and constant way without losing much.

ORGANIZATIONAL PARADIGMS AND THEIR IMPACT ON ORGANIZATIONAL LEARNING STRATEGIES

Thomas Kuhn lands the idea of scientific paradigms into something already real and palpable, he also treats paradigms as what can be focused on the use of tools focused on success and problem solving, so we could bring to our analysis two great paradigms of recent times: the mechanical paradigm and the holistic paradigm.

Mechanical Paradigm: This paradigm has its foundations in the thoughts of Descartes and Newton, which fragments a whole making it a little more treatable making the cause – effect of these parts predominate and their natural order. This type of fragmentation is specifically treated by a specialist.

Holistic Paradigm: This paradigm has its principles based on physical quantum in the early twentieth century that strengthened its criteria with the study and climate and ecological behavior of the middle of the century, at this time you have a focus on the relationships of the parties and their processes, being the order at least amount to their completion and more considers an interdisciplinary approach instead than the specialization of it.

These two great paradigms defined how we have a perception of the world that if we look in detail at today's organizations or companies, we notice that they have two metaphors: The company or organization mechanically and the company or organization as a living element.

If we take the company or organization as a machine, we focus on its parts whether charges, functions, processes, etc. Thus, corresponding to what these parties maintain are in the cause relationship – effect working with each other to give products services and profits.

We can observe in the organization charts, process diagrams and workflows the characteristics and advantages of this machine. When we touch on the holistic paradigm and look at the company or organization as living, we see within it the relationship with the auto environment – regulated by complex relationships between its components.

If we focus a little more on this living organism, we find synergies, systemic relationships. The answer to all this is social learning with everything that involves a new paradigm, the new approach based on the following premises strengthens social learning:

- Human beings are a social being.
- Knowledge includes the development of competencies related to objectives that are valued.
- Knowledge and knowledge relate in the search for these goals, in actively engaging in the world.
- Learning produces meanings through experience.

THE LEADER'S ROLE IN THE FACE OF ORGANIZATIONAL LEARNING CHALLENGES IN THE NEW PARADIGM

What is the role of the leader?

To understand this role, we need to understand the challenges and how you interact with others, sharing their knowledge since value comes from intangibles such as knowledge, know-how and relationships, leaders have great challenges which is to organize their entire team so that they have organizational knowledge and good learning, just as the leader faces the challenge of managing the confusions presented by networks and communities from matrix roles without leaving aside the responsibility that is available and the importance in decision-making.

Different definitions of leadership agree to mark some characteristics:

- The leader is the one who guides the group, sets the course to follow.
- It is an element reducing uncertainties.
- It is an essential element of individual and group motivation. make things happen.
- Proposes horizons and frameworks for action.
- Its presence is covered by a cultural and situational character.
- It has certain personality attributes that stand out from the rest of the individuals, attributes that are recognized, as sought after or desired.
- Leadership is a two-way relationship, where the leader's behavior motivates and is motivated.
- For the characteristics, motivations, interests, and personality of its followers.

What is the importance of leadership in organizations today?

The tasks to be performed are not determined at the supervisor's discretion, but to the factors that have to do with the benefit and aspirations of the entity, company, or business. In addition to making decisions in front of an organization, company, or business, we can name other leadership competencies, communication, how personal relationships are built, where you want to go with the results, learning capacity and growth of subalterns among other competencies. Leadership positively affects the course of the business, in the motivation of people in their work; it should be noted that it can also negatively influence; for this reason, the development of leadership skills is a requirement in the leader.

DIFFERENCES BETWEEN MANAGER AND LEADER

Manager	Leader
who administers	who invokes
maintains	develops
focuses on the system and structure	focuses on the people
control others	inspires confidence
short term	long term
asks how and where	asks how and why
focuses on the results	focuses on the horizon
does not need influence in the group	influencing the group is their main function

FUTURE OF DATA STORAGE SYSTEMS

DATA STORAGE

Internet has evolved, and as large companies such as Google, Amazon, Facebook, and Microsoft began to focus on their data management platforms, it became clear that it was necessary to improve the way data is distributed and stored.

This context has been trying to experiment with distributed data for years, in the 1990s, people who stored data from a local server were processed by requests to other computers.

In the early days of the Internet, people could create and manage their own files, but construction servers required experience in network programming, web security, and time. Networks such as Napster, Gnutella, and Freenet, which attempted to launch the Internet through the open source and shared P2P files, did not have the right infrastructure because they depended on the primary servers or did not have functional software to collect data correctly across distributed nodes.

Successfully and unsuccessfully, distributed Internet protocols have evolved, including those that store data. However, large data centers continue to dominate the Internet and become part of the Internet industry.

THE FUTURE OF STORAGE

Those who develop applications and platforms for the exchange of web content will find other options beyond the heart of the large companies mentioned above. However, there are other ways to improve the file transfer process, download and store it thanks to the Blockchain network and projects such as Swarm, IPFS, Sia, Story and RIF Storage.

The basic premise of a distributed storage system involves creating a network of distributed computers that can work together to store and retrieve hidden and fragmented data. This protects the various network consumers from predators and attackers.

Services such as Swarm, which benefit from Ethereum infrastructure, and IPFS (Inter Planetary File System) functionality, a P2P protocol that can remove controls and restrictions on current centralized services on the Internet. The effectiveness of distributed storage protocols provides a new way to share data to over the Internet using a peer-to-peer network to host files, based on a project created in 2017, IPFS.

The protocols for this data distribution have been tested on networks such as Swarm, SIA, File coin, Story and RIF Storage. A project that seeks to improve the distribution of storage systems through incentives and methods to avoid data impurity.

A. Swarm

Swarm (BZZ), this protocol implements a distributed shard store with addressed content. Does it look complicated? Shards are blocks of data with 4 KB of fixed size.

For content delivery, this concept is intended to indicate that uploaded content will be collected along with a storage address for identification. This address is derived from the content by a hash function that takes the share as input and returns a 32-byte conversation as output.

Hash functions are unpredictable, unequivocal, and distributed everywhere. In fact, this is what makes Bitcoin have a guaranteed high performance.

The hash field is the address that clients can use to retrieve. This is used to map and return the fragments to the network. In addition, the content mentioned above protects data integrity.

The process by which fragments reach their address is called synchronization. If the node requires a fragment, it only sends a request to Swarm using the content address, and the Swarm redirects the request until it receives the information (or when the request exits).

All nodes that make up the Swarm network provide resources (disk space, memory, bandwidth, and CPU) for storage services.

Swarm has an enablement protocol called Swarm Accounting Protocol (SWAP). It is an "eye for an eye" system that calculates the amount of data you need and provide.

Swarm is meant to be a great repository of tools and deliveries to address the problem of real-time interactive web applications for excessive file usage.

B. SIA

SIA was created in the MIT hack event in early 2013 and was finally released in 2015. It uses a performance test algorithm (Pow) to manage blockchain and integrates it with storage contracts that require merchants to allow multiple tokens to exchange file storage capabilities. When storing files, the publisher must always load the storage test into the SIA block to ensure access.

Deposit providers receive results if they comply with the terms and conditions described in the deposit agreement. If so, they receive a commission from the end user.

When saving files, publishers should always load a storage test into the SIA block to ensure access. The depositor will be rewarded if he/she follows the terms and conditions stipulated in the contract. If so, they receive a commission from the end user.

Each component is encrypted before distribution using the Threefish algorithm system, open, secure, and hidden source code. This provides greater security than existing methods because the host only stores closed parts of the files, not the entire file, unlike traditional cloud providers such as Amazon, which typically do not hide user data. You can verify that each domain is owned by a file by using a method called a storage certificate that uses the Merkel tree function.

Those who use SIA have the same channels as the Bitcoin network. This allows anyone who wants to buy distributed storage to pay with season, and those who provide memory space can turn siacoin against smart cough as collateral.

C. File coin

File coin is a distribution network where participants are rewarded for their storage services, creating an algorithm-based market. The market runs on blockchain using native tokens (commonly known as File coin). This is obtained by miners who make deposits to customers. From a visionary point of view, customers spend their mining income to store and retrieve data.

Data and recovery services are provided by a single storage provider. Here, (1) the customer pays to store and retrieve and (2) mining is rewarded with tokens. Cloud protocol networks also provide security for end-to-end clients, while storage providers do not have access to private keys.

File coin contains two new test storages. (1) Replication test, the storage provider may indicate that the data is repeated only once in each physical library.

Deploying a single physical copy allows the evaluator to verify that the provider is not copying too much data to the repository. (2) Space-time testing, this allows the provider to demonstrate that it has stored data for a certain period.

File coin acts as a support line in IPFS and can provide a storage product for any data. This is very useful for data

distribution, the creation and management of distributed applications, and the deployment of smart contracts.

D. Story

Story was originally built on the Ethereum network, but then developed on a network called Tardigrade. The purpose of this work is to implement a distributed system that will allow participants to use their machine's memory space to create a distributed cloud storage space. The project has a cryptocurrency of this name and is designed to improve the development process.

Like SIA, Tardigrade can load the Reed-Solomon network. In this case, it is encrypted on the client side before being distributed to the network. Each file is encrypted with the AES-256-GCM algorithm.

Story began testing Cockroach DB, a similar data management system with Google Spanner. It is designed to store copies of data in multiple locations, so you can download it immediately if necessary.

Recently, the Story team announced that they have improved the method of transferring data to the Tardigrade network through the IPFS system.

It also increases the assurance that the data stored on the IPFS node will last. This is because the same IPFS network architecture does not guarantee the application developer that the data source will always be there, especially if the number of nodes is suddenly disconnected, there is the importance of rewards.

E. RIF Storage

RIF Storage is an uncensored, uncensored, decentralized storage source created by IOV Labs. The project was launched in March 2019, and the test network was officially launched during the LaBITConf festival in Uruguay. Meanwhile, Swarm is working to improve the growth environment, backed by subscriptions to manage incentives and exchanges in the stock market.

Initially, the network operates with the Swarm protocol, but integration with IPFS is expected in the future as portals, markets, and certain anchor services are updated. The test network operates using RIF tokens as the platform currency to pay for bandwidth support.

RIF's vision is to make a lot of money possible in this network. In the meantime, they propose to allow developers to test the RIF Storage user experience with the RIF token. Another novelty at this point is the actual benefits of SWAP, this allows nodes to calculate and measure the amount of data provided by the official micropayment count.

Encryption, this method allows you to assign a hash to each domain that is shared (using the Merkel tree function) and then combine the same locations where files are searched, and ownership of this data is displayed.

RIF Storage is part of a larger infrastructure called RIF OS (RSK Infrastructure Framework Open Standard), consisting of an identity and reputation system, protocols that allow access to paid networks.

DATAWAREHOUSE

METHODOLOGICAL

Because we can start helping an organization we need to start from the origin and for this we need what we call transactional systems, ERP systems, to manage the globality of the resources of the company or organization, CRM systems to manage relationships with our customers such as supplies, supplies etc.

We talk about any system where we have data, company information or organization from texts and flat files to gigantic systems, all this data we will take them out to make a transformation so that we can take them to a data warehouse by means of a tool that we will call ETL (Extraction, transformation and load) we call this tool Data warehouse which is a large depositor collection of data integrated or structured to a specific topic from which we will start, in order to build what is right for us in this case, which is strategic decision making, this makes it easier for us to process it and the power to try everything and make it much easier, to this we add two more elements, the Data Marts that are so to speak, small cubes or a database centered to a specific area within our organization, this helps us to be even more efficient to search or query the data that we have within this large data warehouse and that we will have at our disposal, information that we consult frequently gaining speed and efficiency.

And on the other hand, OLAP (Online Analytical Processing) systems, very similar to the previous one but these cubes are a more fluid and dynamic system so from there we have a little more the online concept, which allow us to capture or extract the data by selecting it in a simpler way and obtaining large volumes of information.

This increases the probabilities in analytical processing in terms of building intermediate systems between the data warehouse and the systems that eventually exploit this information, so we can say that all these systems that feed off data warehouse DataMart and OLAP systems are all those series of applications with which we will help organizations or companies make business decisions, query systems such as Querying and Reporting, querying and building reports to determine relevant events and pass them on to the different dependencies or functional areas of our company, this helps us learn to know our customers and try to a fluid, natural, empathetic, and pleasurable relationship supported by data and information and business systems.

DISUSED

Business Intelligence Architecture Currently where new and innovative technologies continuously appear in the market are traditional business intelligence, a new generation of architectures are appearing in the market covering new needs that are being required such as advanced analytics functions, based on statistics and predictive analysis, real-time analysis of online data and sophisticated data visualization interfaces, large mass management of Big data data through new concepts such as data refineries, as well as the use of data virtualization to improve and employ the process of integrating data transformation and load extraction and finally new options design and development of platforms based on Cloud Mobile solutions and integrated hardware and software infrastructure these technological changes are driving and pressuring companies to create new and better knowledge and also pressure them to make these new technologies represent a reduction in overall expansion costs of these environments, there is still confusion about where these technologies and functionalities would be integrated within the modern architectures of Business Intelligence, here it is worth asking! does the Enterprise Datawarehouse still have a role in the Business Intelligence infrastructure where Hadoop technologies fit in? How do these new technologies and capabilities meet advanced analytics needs in real time effectively? Is the enterprise data warehouse still critical to the business intelligence infrastructure?

but its role is changing from the classic source data object repository function, for standard reporting comparisons and analysis, the enterprise data warehouse is a better integrated and quality data source for critical and sensitive analytics processes such as financial analysis, compliance requirements analysis and regulations, and for building the dashboard-based control system through critical data such as KPIs or key performance indicators or other metrics business for the various areas of the company.

However, the traditional data warehouse and business intelligence infrastructures do not meet all current needs especially those related to the management of massive and heterogeneous data to computational research-based analytics functions or Hadoop real-time data analysis and other technologies for relational and non-relational data management as well as new analytical solutions require expanding the traditional model of goods and intelligence and adding new components to the architecture, we can see on the screen the three components that should be included in the new business intelligence architecture.

The first component is the computational research and analytics platform which is where new relational software and Hadoop applications shine on this platform is used for data exploration and for the creation of new analytical models based on data mining cause-effect analysis exploration what if pattern analysis and in general unplanned data research these analytical platforms can be used partially or totally by companies as an extension of the Data refinery applications, the data refinery allows you to analyze large volumes of

structured or non-structured data quickly and with excellent results these information outputs can be used by data warehouse platforms in real time through analytics engines or also through applications that are online or online.

The second component of this new infrastructure is the data refinery platform whose purpose is to capture detail data in batches or in real time from new and diverse digital data sources such as social media sensors rf and d&s etc. and upload it to a relational or non-relational data management platform in the same way that a refinery converts crude oil into its derivatives, the data refinery converts unstructured or unstructured dissex data into useful information for sending to advanced analytics platforms or sending it to the enterprise data warehouse, these new platforms require more flexible data management policies in terms of privacy security, storage quality, and data deletion than those traditional data integration platforms in our enterprise data warehouse.

The latest component of this new platform provides real-time analytical functionalities and is within the operational environment its purpose is to develop or implement real-time analytical models such as web event analysis, risk analysis and data flow optimization since the rules functions and analytical models of this platform in real time have probably been developed in the enterprise data warehouse, the computational research analysis platform and real-time analytics components must be integrated and share information in a fluid manner the new components of analytical business intelligence and data management work together fulfilling their purposes in the future it is unlikely that the architecture will change much as the requirements of real-time and targeted analytical processes and computational research will remain constant however technologies will change when new developments and requirements of these processes are implemented with new technological approaches business intelligence platforms currently focused on providing autonomy and analytical capacity to the business, have reached their turning point and are now looking more insistently to provide better tools to the strategic business user in 2018 the business intelligence platform offers the business user of analysts autonomous tools for advanced data preparation and analysis, in the same year autonomous platforms are alone for data preparation, they will have expanded, integrated with analytical applications always in 2018 intelligent platforms based on Hadoop data exploration and discovery converge on a single advanced generation platform for data discovery processes that will include autonomous data preparation and natural language the evolution and sophistication of the autonomous data discovery and data preparation platforms available on the market have changed the focus of buyers in the market of analytical business intelligence platforms, directing friendly and interactive solutions that provide a wide range of advanced analytics functions and do not require significant IT participation for defining the data model to be used in the process.

According to Gartner the business intelligence and analytics platforms can be described and evaluated based on five functional areas these are agility and centralization of the

platform that allows management of information flows from strategic users from centralized content with the help of autonomous data management functions.

Second, autonomous decentralized analytical processes on the part of the user.

Third, data discover platform that includes autonomous analytical processes.

Fourth: Business intelligence features built into applications.

Fifth: and latest development of external networks that include business intelligence analytical functions for external users. Likewise, suppliers are evaluated through 14 critical functions classified into four categories these categories are:

First, infrastructure this category chair contains four functionalities.

First: Platform management feature that enables platform scalability by improving performance and ensuring disaster recovery availability.

Second, cloud services platform for and analytics applications to design deploy and manage analytical processes in the cloud with data in the cloud or on-premises systems.

Third, user management and security functionality.

Fourth, data source connectivity functionality that enables the integration of structured or unstructured data from various cloud data platforms.

Second category data manager: This category has three functionalities.

First: Governance and metadata management are tools that allow sharing thematic data models and metadata this would allow administrators to look to capture, store, reuse and publish goal objects to this as dimensions hierarchies measured layers report design objects parameters etc. in a simple way. Second: functionality. Removing autonomous TV and text content capabilities to integrate transform and load data into stand-alone storage layers.

Third: Functionality autonomous data preparation services drop work functions for handling and combining data from various sources and creating analytical models such as larger data areas groups and hierarchies, these advanced functions include discovering semantic joined layers and intelligent profiles generating hierarchies from various data sources that include multi-structured data.

Third category: this category contains four features these are:

First advanced analytics built in.

Second, dashboards represent interactive dashboard creation features with visual exploration and advanced geospatial analysis features for user consumption.

Third level: interactive visual exploration through color graphics illuminations zooms figures and visual movements representing areas of data being analyzed.

Fourth and last functionality: mobile browsing features that allow you to deliver content to mobile devices interactively through publishing on our portals and that takes advantage of

the functions of mobile computers such as touch screen camera location of devices and queries in natural language. Fourth and last category: compare and share results: this one has three functionalities.

First, incorporation of analytical content are those functions to modify analytical content to incorporate them into business applications and portals based on libraries for developers' apps and support for open standards these functions of integration of the store is analytical intelligence with the application architecture will allow the user to choose where in the business process the analytics should be. Incorporated.

Second functionality: publishing analytical content. Third functionality: Collaboration and social business intelligence that allows the user to share and discuss information and analytical content through forums of discuss on chats and other annotations.

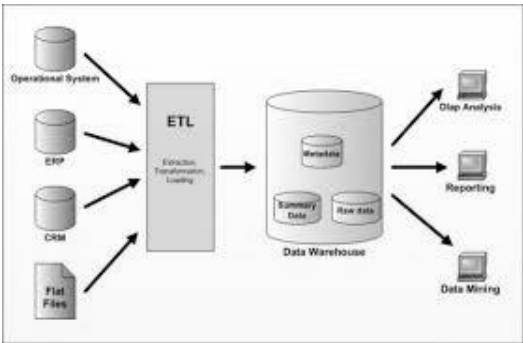
F. Early Data warehouse

-It originated in the mid-1980s and this technology provides us with an architecture for data output in operational systems in decision making.

-What does Data warehouse mean?

The most important element during the implementation of a business intelligence project is the data warehouse is where we will store all the information obtained from the different sources of our company and why it is one of the most important elements or elements,

Table I



https://www.google.com/search?q=data+warehouse&client=firefox-b-d&source=lnms&tbn=isch&sa=X&ved=2ahUKEwiWpd_KzfnuAhXk01kKHcRXBq8Q_AUoAXoECAIAw&biw=911&bih=436#imgrc=4bW5GUI6yL_TKM

here we're going to concentrate all the information with the structure and design where we can exploit this information, these structures are composed of fragments derived from the data warehouse known as data marts and these Santa Marta in turn have designs where we store the information also known as star models or snowflake models that we will be talking about later and these designs will not allow to exploit information for different purposes for example for reporting or analyzing information through onlap cubes data mining

dashboards among other types of solutions in the following image we can see the role that the data warehouse has within a deployment for a business intelligence solution.

We should observe how on the left side we have operational data sources these can be formed from ERP systems systems CRM flat text files or some legacy systems the sources may be diverse the only important requirement is that these can be extracted and the tool that will allow me to extract the data from these sources is known as dt it is that by its acronym extraction transformation allowed and is a tool for extracting data from data sources transformation because it requires good transformation because we had mentioned that data warehouses through their data marts implement a model that is known as a star 1 snowflake and good need to be transformed because at source they do not have these models and well because the load or boom of this data towards these special models within the ties them in the data warehouse and well finally we can see how we already have the representation that is often visual or good that is what is intended but many other times not as for example in the case of data mining that we will also be talking about later and well since the visual manifestation can be for example for solar analysis through the cubes the reports or what we had mentioned a moment ago that are also known as boards of control something very important to mention about the data warehouse is that it is not a software and much less a brand or a single database in general we can abstract ourselves from the logical and physical models of the databases that make up our environment and we refer to its whole as a data warehouse for this reason regardless of what business intelligence solution you are implementing if you are not sure what the purpose of the data warehouse is likely to arrive to have many deficiencies when exploiting data and analyzing information one of the purposes of the data warehouse is that they should focus on the whole company which means that it must provide information for the area of sales, finance, human resources, production, operations, et cetera and of course also to the general management another of the repositors is that its design must conform to the changes as possible we live in an era where business and information is very volatile and good because the data warehouse must to be prepared to receive these changes prepared for bulk data loading should be designed to load massive amounts of data preferably in a small period of time let us remember that these systems are primarily designed for the analysis of information i.e. queries and therefore it is not advisable to put it to live together in the same productive environment of our transactional systems as our ERP or point of sale as they can slow down our server's performance by querying large volumes of data or also running the risk of not delivering information in an agile manner and the nature of the data warehouse must be multipurpose

your data must be in a format that supports any and all possible forms of business intelligence analysis in any and all your technologies differences with a conventional database the traditional databases used by transactional systems in relation to a data warehouse are polar opposites as to your operating and design requirements on the one hand transactional systems are designed to execute transactions of the type high low

changes and data queries such as a charge a credit memo an inventory return the registration of a new customer etcetera on the other hand a data warehouse is organized based on concepts such as customers products time sales etc. there are also differences in design while transactional databases are extremely standardized a data warehousing to not be standardized by organizing the data in conceptual warehouses known as data marts and its star model as a personal opinion I consider that the development of a data warehouse in the process of implementing a business intelligence solution is equal to or even more Important that the technology selected for exploitation since without a good model we can face problems such as prolonged response times inconsistent information problems to display information among other things so even if you have the most expensive technology to the latest trends if you do not achieve the purpose of the data warehouse and its correct design is almost a guarantee that your business intelligence project will not succeed.

DATA MINING, A FOCUSED USE FOR TODAY

IN THE CURRENT ERA

And they exist different methods for analysis focused on obtaining information from Big Data, each type of analysis being as a different result. This allows several different analyses to provide different results.

Data Mining is characterized by the discovery of uncovered patterns in previous scenarios, unusual records linked to dependencies, detecting anomalies forcing the company to develop methods to better work big data.

Of course you have to be clear about the concept of data mining and how it can be applied to almost all areas of computing, starting from a simple database, to on-premises or cloud servers, whose information is the raw material to work with; and is that the great problem of growing companies is the lack of experience in being able to manage all the load of information and responsibility that contracts their databases, affecting in principle their infrastructure, since they do not have the tools (such as Software, servers, qualified personnel, structured cabling network or strategies oriented to database management), their ability to manage in the face of production of 100% in their workload, the constant search for better methods to manage Big Data and new goals of the company or business.

The consequences of missing information, capable of leading any project or organization into complete chaos, may sound worrying on paper; but this is also a stimulus in any area that leads to the constant improvement and growth of better systems, infrastructure and staff trained for the management of the Bigdata: Data Mining is the solution par excellence that determines how to order, plan, sensitize or optimize the

management of any company, project or business that requires management (not only information, but personnel, customers and active resources of a company) appropriate and accurate to suit any company regardless of its size.

We are in the information age, and it itself becomes the most asset in the world, since the one who controls it and knows how to handle it, controls the world, however, it becomes complex to have to deal with Gigas and Terabytes of information, including areas and topics that we do not understand in our competition, which makes it complex not knowing how to control the database that is being generated and disorganized.

This is where data mining comes in; and is called mining is by the concept linked to extracting bases important, significant, and vital information for the company, project, or company, which focuses on using these bases to design models and statistics oriented to the construction of strategies that allow the improvement of the company's work or activities, bringing with it benefits such as cost reduction and productivity improvement. All this is organized into a pyramid in which this information flows at different levels of abstraction that can be handled; at the first level you can see the storage of the data where the databases that accumulate from the operation of a company or organization are stored, on a second level, this information will become typified, organized and consolidated, leaving clean data in a consistent format, becoming a Data Warehouse, at the next level we find the standard reports, focus, for example on the presentation of information to users or customers, they can also be called information cubes, this tool being a good method to obtain the information more quickly and accurately for their analysis; at the next level, information becomes part of various scenarios, devices, and environments (the most common being the internet). Now at this point it can be stated that all the information extracted or obtained is sufficiently processed to be considered a Mined Data.

These last 3 levels focus on the Business Intelligence area geared towards intelligent business management, Characterized by its large intelligent analysis of business information at the enterprise level, all this automatically or semi-automatically, no matter how large the databases, this exploration and analysis will cover it in its entirety, once the information is processed, a much deeper analysis can be done giving opening to the research applied towards patterns, relationships, and summaries of the data understandable to those who own the databases.

The goal after having already the processed information is to generate in the company or company a memory to start working what is called intelligence, which is the consolidation of the information processed to start working on the improvement of the operations of the company or company in all its aspects, as we had spoken in the beginning not only in gross information of the company, but any area that has as active the company, and this is very advantageous with respect to time and cost management.

Current technology already allows this type of analysis (unlike 20 years ago), and they do not have to be such complete or state-of-the-art equipment or NASA or servers, only a single office team with the right software and person to undertake such exercises is enough; this brings a price and is that companies are beginning to understand Data Mining as one of its pillars, starting a strong competition with the aim of creating an advantage over the other companies; About 15 years ago programs to implement Data Mining processes in their entirety were payments but today the software that allows us to make Data Mining is not only paid, but also migrated to the Open Source area, which makes even smaller businesses able to access it optimizing monetary costs. And leaving the door open so that anyone, group, or project has more facilities to grow business ally, so there is no longer the excuse of not having the necessary resource to mine data, it only takes will and good practice to perfect the mining methods.

As mentioned above, our raw material is the database, and more specific, a modern database which has certain characteristics such as its large size, but there are 30 observations as it was 20 years ago, this today is microscopic, databases today are transactional having billions of observations; this generates the need to move this large amount of information, the process cannot be intuitive as it is humanly and individually impossible to work this type of data without models or Software.

We will mention the discovery of things that are not obvious because we have not explored the database, things that would never have occurred, without the help of modern methods which becomes a huge advance, not only technological but transitional, since it was unthinkable to generate some strategies that for today are so viable and manageable that it would adapt to any type of business, activity, or project.

THE BENEFITS

Big-Data leads us to manage a lot of information, which as we mentioned is impossible to manage without systems or automation programs, this is where our Data Mining model looks like allowing that great management of resources on one or more platforms, leaving us out of experiments, tests, to discover needs, expose variability and improve the performance of activities. Let's talk about more specific instances, in this case population segmentation, leaving instances like being able to classify them into different groups in order to treat each group in an independent way, in a way that allows the customer to be treated in a better way and in the process, all this to generate better production and profits for a company, having access to this data can really generate an automated system that can even replace intervention or decisions mitigating mistakes, saving time and money.

For a growing company, the knowledge of new market strategies and business models is so beneficial and beneficial that good mining practice makes the process much more comfortable; It should be noted that not only Data Mining

would focus on business, but also for other areas also important for the development of society such as the area of medicine, public administration, the global location of personal data, profitability management and taxes of a state, manufacturing and industrial production.

All these industries could be beneficial both logistically and monetarily, reaching profits of \$300 billion per year according to a study conducted by McKinsey Global Institute Analysis, being a very beneficial part and practically money from the ground to be collected, leaving several examples such as the European economy that generates 250 billion euros per year where the application of data mining allows this type of profit in this region.

At this point we can conclude that the correct management of Big Data is the key to today's industrial and business competition for individual companies, becoming the most important asset, we can no longer rely solely on the rudimentary methods of analysis that could serve in the past, now all this range of strategies have evoked to contribute to the entrepreneurship sector what really exercises the good development of the economy in the industry.

THEIR APPS

Data mining allows us to take it to many areas:

- Marketing.
- Analysis for customer retention.
- Study of consumer behavior.
- Production processes .
- Cost study.
- Use of capital and labor.
- Quality control
- Inventory Management.
- Fraud Prediction.

For the marketing sector, we can specify strategies that can be a better management of customer service (CRM) based on knowing what the customer wants, benefiting both parties and generating solutions faster for each type of customer or need, resulting in better customer satisfaction who will be the most benefited by raising the response rate.

Another great application for data mining is to know the profile of customers to know them, determine their needs, this leads to benefits such as creating more effective solutions that the customer can appreciate and buy more and more. It is also trying to find groups of different customer behaviors, a single company that works various sectors such as infrastructure, manufacturing and sales will have different groups of customers, they cannot be treated in the same way, this causes each sector or activity of the company to segment the group of customers, which behave in different ways or in some cases in similar ways thus calling itself a classic segmentation in marketing.

Customer retention creates convenience by better understanding customer turnover and desertion (cell phone companies, credit cards, software licenses, cable subscription, among others) where they are areas that customers commonly want to defect and jump from company to company generating additional costs for companies that have invested large amounts of money for the best management of their products and services, but that despite the effort is not able to meet the needs of the customer.

Another improvement towards customer management is to typify their behavior, i.e. how the user currently sees the product or service they are receiving from the company, or business, the user's opinion should be taken into account both in the technical margin to discover their needs and start undertaking a better product or service, also know that product or service is not being consumed to reduce their production (supply and demand) and is also the framework of creating new needs for the user according to their own needs, such as the creation of a new product, service, subscription plan, improvements or additional wildcards etc.

We can say that the cost study focuses on several areas, from seeing the spending-benefit ratio, to minimizing costs in production or projects, being an indispensable asset in any company should not be taken as an unlimited resource or also scarce, since the business tends to grow when more profits are seen.

The advantage of having the costs calculated in the operations of any company reflects good management, since some companies believe they have their finances under control, but some companies later reflect the problems that have been ignoring for years for not predicting or obtaining the right information at critical times of their projects and operations.

The management of resources and personnel are closely linked to the previous process, since you also see the financial part but now as internal management of the company, since each worker in the company requires a cost to be able to continue in its work, and, depending on your position in the area you perform, this will be your salary, social security, pension, endowment and tool, whether a fixed or temporary worker, always an employee will generate costs in the company. Now an employee study, exams, training, filters, and annual tests would also be applied here, all to qualify staff and know that people have the necessary attitudes and skills for each area or project that is present in the company.

Even toilet staff need to be evaluated, let us not forget that information is what it takes to be able to process it and quantify what quality of labor the company has, being able to predict more and more accurately that people are contributing to the growth of the company and that they can be considered unnecessary burdens and that they are not required at least in the projects or operations of the company.

Quality, both in services and in the product is the first thing that customers or beneficiaries of them see, quality is measured both in standards that the same company defines, and in the customer experience when testing or receiving a

product, an example is the IBM which in the 70s did not have as a viable concept that a person had a computer in their homes, so absurd and wrong was their vision in the face of this idea of a home computer that would later be astonished to see that what they themselves said as unthinkable was reality and prosperous, making innovation their role in the quality of services and products, no longer as impossible to achieve but as an opportunity for the quality of what is being given to the customer to be much better.

We have talked that the most important thing that a company can have at the asset level is its information and of course knowing how much I have, how much I have, how much it comes in and out is no exception, here we see that, if we do not have control over, for example, our manufactured goods or products will be impossible to know whether our business is growing or not, resulting in an unknown so great that it almost always ends up preventing a company from growing. So important is to know how much I have that if, for example, in a company that handles computer parts do not know where the processors are located and a customer requires several of them, it is money lost, besides that in some scenarios the company does not know that it is missing these resources and it is necessary that the company knows it to generate the corresponding correctives, as can be the monitoring of the goods, putting plaque on their equipment, counting lots and relating them to a summary that will give the actual values that the company actually has, also applies for simple tools such as office elements (Papers, Pens, Staples, etc.), it should be clarified that in the technical area of some companies it should be very clear that if the worker does not have the necessary tools or in optimal conditions for his work, the expectations that the company will achieve its objectives in the provision of its services cannot be achieved or does not become the objective projected by the company, the worker also requires to give information about the conditions in which it works, and control and inventory methods are needed to find patterns that degrade the tools and equipment of the company.

Another very important application in Data Mining and statistics is risk detection and fraud detection, which has been successfully used by companies (especially financial institutions) to assess the risk of some clients who can do something that they mention as Default, very crucial to avoid losses and potential years in your company. Identify the associated characteristics such as good-paying or poorly paid behavior, measuring their compliance performance with indicators and scoring systems that allows to identify the company how good a customer is certain people, assigning each client a profile with valuable information for the decision making and future of that client and dealing in a differentiated way and if the client is very good, punctual with their payments, you can start to give profits, quite favorable options looking, of course pointing to the benefits of both the client and the company.

Another benefit in controlling customers is to ensure that no fraud is committed, such as at the transaction level, avoiding movements or actions considered suspicious, comparing, for example, a transaction with a customer's purchase pattern, if

pattern moves far away from what is normal for a customer it can be marked as a possible fraudulent transaction.

THE TECHNICS

The following techniques are used as a methodology and best practices for Data Mining as fundamental principles in data collection and management exercises:

Visualization: This is the initial step, it is the method in principle in the database to get the first data for information analysis, there is a mistake of people believing that Data Mining is a difficult method, but the truth is that it begins as an exploratory process, and there is no better way than to explore the data than in a visual way; humans are very bad at processing ones and zeros, but we are very good at processing patterns if presented visually. Graphs are an important factor, as it allows us to capture those patterns that we detect becoming easier to identify the information we require.

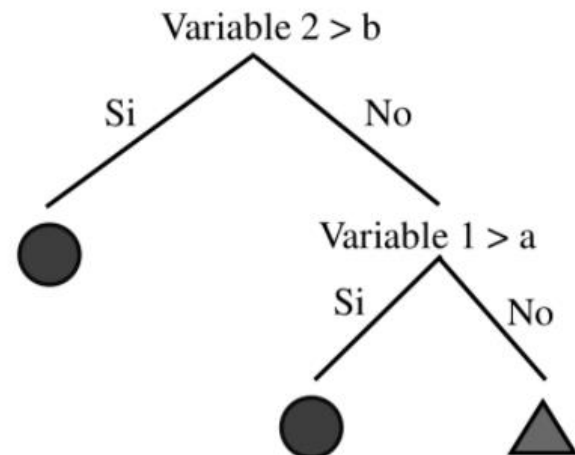
Prediction: Your goal is simple, forecast the value that the variable of interest takes as a function of an individual's characteristics and attributes, this variable can take a value in a certain range; you might think that you are building a customer profile or prediction template. A good example is how much a customer buys me in the 1-year period, here the variable is the amount of money you enter measured in dollars (\$). This helps segment customers according to their value, that is, the different customers of a company are not the same as mentioned in previous points, it is a misconception that appropriates the concepts of a company, this concept is false, some customers are more valuable than others, then you can predict how valuable a customer will be, you can generate a more preferential treatment for each type.

Regression: Predicts the relationship of attributes, a known technique is linear regression, summarizes the relationship between variables to be predicted and the variables that are used to forecast next to a linear equation. The idea is to consolidate all these variables to formulate a simple forecast, an example may be how much a person spends on beer according to their demographic characteristics, here the idea of Data Mining is that it gives us a database with an N number of observations based on an N number of attributes, here we can work with the variables necessary to be able to consolidate the prediction to be solved, such as age, sex, region, income and purchase periods. All this data must be translated into numerical values or equations, allowing the results of that prediction to be accurate enough to become trusted.

This method can be heard simple and practical and is, even with an immense number of variables and characteristics. What predictions can the previous exercise yield? One can be what sex consumes the most beer?, beer can be connected with education?, where you can see which range of ages are the most likely to consume beer, what kind of people are the ones who consume beer the most, this can transform the equations into a profile of who is an optimal customer for the business,

store or company that sells this product that is beer and can design action plans for the different groups of Customers.

Other techniques are regression trees, entering in the graph data the attributes mentioned above and making connected relationships with each one, reflecting the connections between each:



For our example, a tree is created for each type of person, charting their attributes, analyzing characteristics that the person especially has, this type of tree allows to compare conditionals in the style of an algorithm, such as if the person has an X amount of income, it can be predicted that the person will have a Y amount of expenses and if it is below generate new variables and also generating new profiles that have different characteristics and that they spend different in the company, allowing segmentation by customer value.

Classification: This technique is a comparison at the group level to be able to show if a person belongs to any group or several, giving us room to ask questions that will help us classify the client:

Example:

- Can we forecast if a customer brings us problems?
- Will a customer leave the company to go to the competition?
- Is the customer going to pay me or not?
- Is a transaction fraudulent or not?

It sounds like prediction or regression, but its area is simpler, since here we must assign each group of clients to the individuals to evaluate, and then through progressive regression, define the value of the client and its potential future, according to the questions that have been asked for the valuation of that individual; we have at our disposal information on some demographic characteristics of the individual: sex and age. We want to convert these 2 variables into a kind of score that is related to answering one of the questions we have asked ourselves.

Example:

Score $-26.46 + 0.78 \text{ age} - 0.55 \text{ female}$.

The idea of dealing with such formulas is to condense all the information received into a single line, giving room for a quick classification of each individual and starting to see common patterns or characteristics that generate a filter to know whether a certain group of people are customers or not. Here comes the concept of speculating on a probability of, for example, buying or in the case of banks, probabilities of payment.

Prob	Score
0,9	2,19722
0,8	1,38629
0,7	0,4873
0,6	0,40547
0,5	0
0,4	-0,40547
0,3	-0,8473
0,2	-1,38629
0,1	-2,19722

The Score can be considered as a qualifying method for evaluating a customer's purchase values, for this example if the Score is higher, the odds can increase, so that the person is a good customer.

Other classification techniques:

- Classification trees.
- Random Forest
- Neural Networks

Grouping: it is about defining groups of individuals, we first started in our analysis with predefined groups and the work was to develop a model to know how the new observations should be classified.

For grouping there are no pre-classified groups, but already organized structures, what is sought is the formation of groups according to their observations and if they resemble each other, being logically affirmative that observations belonging to different groups are different, in this case groups of similar clients may have a similar behavior as well.

Age	Salary
20	40
25	50
24	45
23	50
40	80
45	85
42	87
35	82
70	30

This example works with 2 variables: age and salary, here you can consider the exercise of how many groups we have in this table, generating multiple observations.

Although the table can be simple you can get a lot of information that will allow us to evaluate multiple similar characteristics for this group, where in each observation it will allow the common relationship and creation of the groups. This helps a lot to create segments and anomalies.

Association rules: it is based on determining "what goes with that", a classic example is a supermarket where you want to find combinations of objects that frequently occur in a transaction database, Here you can relate baskets of customers who frequently go together, this being a common pattern, you can also observe the frequency or intensity of this type of purchases, all this in order to impact more positively on the merchandise that you want to offer, optimizing and being more precise in the supply of products and of course leaving the customers more satisfied since you know what the customer is really looking for with higher priority, with this would also improve the revenue of the supermarket, since special offers can be made on the products, inciting the customer not to have to look for each product, but to offer this common basket at a single cost.

Example:

If the customer buys me milk, then he buys me coffee.

This allows customer interaction with the product in a more friendly and effective way, you can generate strategies so that the customer does not have to try too hard to look for the products it requires. For the example above if my customer frequents the purchase of milk with coffee, I can place on the shelves both products, I can generate in a single product coffee and milk, I can generate advertising offering these 2 products more frequently so that the customer knows what brand or product to count on, I can on certain days of the week lower the prices of these products in order to increase sales of them, and also letting the customer know that they can trust the warehouse, implementing personalized marketing.

Also, for the industry that manufactures the products can benefit in this example, since they can increase the quality in production, for this example, of the coffee or milk they

process, adding new presentations, different quantities, higher production, and manufacturing priority, because the demand for their products will increase.

All these techniques are applicable in any business sector, regardless of its size or objective that it has in mind, each method, pattern, demand, or activity that the company carries out for its growth is considered a business model, and each company has the immense opportunity to grow if it applies each of the techniques mentioned above, it should be noted that the technology that is applied in the current times (XXI century) , allows the management of businesses much more effectively than in times of yesteryear; the Software will be the main tool for the proper practice of Data Mining.

Conclusions

As all theories have been shown to have strengths, threats, weaknesses and opportunities, companies are adapting to the theory that is most accessible to their leaders as some of them have been proposed many years ago and others until now are being incorporated.

I believe that the kind of leadership that can best be adapted is the situational one as the leader can let the team act according to the circumstances and needs of a project, leaving more flexible leadership to see rather than being rigid and static. The motivation, capacity and knowledge of the team should eventually be evaluated as this depends on how the capabilities of the entire team can be used in each project and thus give the best possible result, so the leader himself must also adapt to the circumstances and thus be part of a group growth.

Efficient leaders are those who know your team to respond appropriately and efficiently with certain projects while recognizing that you have a human team that at any time for any reason can change a person on the team and thus not let you see the gap that this person can leave on the team.

All styles of leaders have been successful in due course the evolution of the same leader has been reflected over the years as needs arise in the face of projects and these in turn a select team must be prepared for it headed or hand in hand with characteristics that have been shown above.

A good leader will always be characterized by the strength and charisma that impedes his work environment and within him his team so that they copy these behaviors when the strengthening of behaviors is needed. You can also be recognized for being focused and demanding as this way things can be done in a better way and thus get the best results; it is to be known and understood that new leaders are forming in a different way for the 21st century.

Decentralization depends to a large extent on how rewards can become a means of adding value in terms of collaboration and adoption. This is because once users use these services to avoid controls from large companies or governments, or to generate money (storing 1 TB on a network such as SIA could

cost between \$1 and \$2 per month or \$10 in Story, compared to the \$23 it would cost on Amazon S3) it is necessary to ensure that there will always be an interconnected node to help make everyone's files safe and available.

SIA, indicates that they have a capacity of 2 bytes and that currently 730 tera bytes are used, distributed in more than 320 storage providers. But these storage capabilities, similar on networks like Story, are insufficient when you consider how much information is added to the Internet daily.

As technology will continue to advance over the next few years, there are underlying performance issues that make it difficult for users and businesses to start migrating to a decentralized network.

For decentralized cloud storage to increase in popularity, they must be able to offer a service like or greater than that offered by the centralized market. On the other hand, as the technology is on the move, we are several years away from that happening; therefore, it will remain a niche.

Each of these services aims to move the way forward so that the Internet is a less controlled means by individual institutions or governments. For this reason, decentralized storage systems are a piece of a broader set of tools for the creation of the Internet of Valor.

We can conclude that Business Intelligence and Data warehouse with tools that companies, and organizations must improve, extract, debug, analyze, visualize the decision-making in our organization.

Business Intelligence (BI) we talk about ERP systems, transactional systems that help us manage the globality of resources in our organization.

Referencies

- [1] Mendoza de Graterol, Eva and Mendoza de Lorbes, María Antonia. (2008). The paradigm shift in organizations as the foundation of the ethical leadership style. <https://www.aacademica.org/000-096/678.pdf>
- [2] Franco, Carlos. (2015). LEADERSHIP IN THE NEW PARADIGM OF ORGANIZATIONS. <https://cesoftco.net/el-rol-del-lider-en-el-nuevo-paradigma-de-lasorganizaciones-y-las-comunidades-de-practica/>
- [3] Introduction to Administration: Paradigms in Organizations. <https://clea.edu.mx/biblioteca/introducc%20a%20la%20administracion.pdf>
- [4] Martinez, Ignacio. (2017). The importance of leadership in organizations, 7 factors that support it. <http://ignaciomartineza.com/liderazgo/importancia-del-liderazgo-en-las-organizaciones/>
- [5] Martinez, Esther. (2013) The importance of leadership in organizations. https://ddd.uab.cat/pub/tfg/2013/110463/TFG_elopezmartinez.pdf
- [6] Ballina Ríos, Francisco 2000 Management Theory. An alternative approach, FCA McGraw Hill, Mexico, pp. 95
- [7] Fielder, Fred E. 1995, Leadership and Effective Management. Threshing, 2nd. Mexico, pp. 11-12.
- [8] Koontz, Harold., Weihrich Heinz 2006 Administration. A global perspective. McGraw-Hill, 12th, edition, Mexico, pp. 532
- [9] Lussier, Robert N., Achua Christopher F. 2005 Leadership. Theory, Application and Skills Development, 2nd. Edition, Thomson Learning, Mexico, pp. 359-360
- [10] Robbins, Stephen P. 1999 Organizational Behavior, 8th . ed. Prentice Hall, Mexico,
- [11] Schermerhorn, John R., James Hunt 2004 Organizational Behavior, Limusa Wiley, Mexico, pp. 293, 296, 302
- [12] (2009, May). *Centralized database*. Obtained: <https://aishiteru-onna.blogspot.com/2009/05/bases-de-datos-centralizadas.html>.
- [13] (2012, August). *Centralized Architecture*. Obtained: <https://normalizacion-bd.blogspot.com/2012/11/5-arquitectura-centralizada.html>.
- [14] F. Rosa, (2018, April). *Blockchain and decentralized databases: are they the same thing?* Obtained: <https://www.criptonoticias.com/tecnologia/blockchain-y-bases-de-datos-descentralizadas-son-la-misma-cosa/>
- [15] Vegagestion, (2017, September). *Cloud storage: features, pros and cons*. Obtained: <https://vegagestion.es/almacenamiento-la-nube-caracteristicas-ventajas-desventajas/>
- [16] Incubicon. (2019, August). *Types of distributed storage systems for blockchain networks*. Obtained: <https://blog.incubicon.com/tipos-de-sistemas-de-almacenamiento-distribuido-para-redes-blockchain>
- [17] J. F. Bolaños, (2018, October). *Easy explanation of IPFS protocol, distributed web*. Obtained: <https://medium.com/@jfbolanos/explicaci%C3%B3n-f%C3%A1cil-del-protocol-ipfs-la-web-distributed-8e3e725240ff>
- [18] L- Scrap, (2018, September). *5 decentralized storage solutions based on cryptoactive technology*. Obtained: <https://www.criptonoticias.com/negocios/5-soluciones-almacenamiento-descentralizado-basadas-tecnologia-cryptoactivos/>
- [19] https://www.google.com/search?q=data+warehouse&client=firefox-b-d&source=lnms&tbn=isch&sa=X&ved=2ahUKEwiWpd_KzfnuAhXk01kKHcRXBq8Q_AUoAXoECAIQAw&biw=911&bih=436#imgsrc=4bW5GUI6yL_TKM
- [20] <https://www.youtube.com/watch?v=SHrfXdRDOH4>
- [21] <https://virtual.cun.edu.co/contenidos/migracion2020/sistemas/s9/business-intelligence/u3/recurso3.pdf>
- [22] Data Mining for Market Intelligence, INCAE Business School, 13 Oct 2011, <https://www.youtube.com/watch?v=YiR9k157iA0>
- [23] Which is the Data Minería, Eduardo Lecaros 10 Oct 2011 <https://www.youtube.com/watch?v=a2jBG-74tC8>

Biographies

Miguel Angel Ayala Castillo, system engineering student,
national unified education corporation superior, Bogotá -
Colombia 2021.

Luis Rodrigo Jiménez Bohórquez, system engineering student,
national unified education corporation superior, Bogota -
Colombia 2021.

Sergio Andrés Martínez Muñoz, system engineering student,
National Unified Higher Education Corporation – CUN,
virtual modality, Bogota – Colombia 2021.

María Cristina Rojas Penagos system engineering student,
national unified education corporation superior, Bogotá -
Colombia 2021.