

Data Mining in the Financial Sector of Colombia and its Value in B2C Methodology

Luis Eduardo Muñoz Guerrero

Universidad Tecnológica de Pereira, Colombia, ORCID: <https://orcid.org/0000-0002-9414-6187>

Summary:

In the origin and evolution of industries, systems engineering has been fundamental, offering effectiveness in process optimization. The financial industry, in particular, has benefited greatly. Next, we seek to characterize and identify the impact of data mining on the modernization of the Colombian financial sector, from the 1990s to the present. The methodology that is executed has a qualitative and correlational approach compared to official surveys from the Bank of the Republic, whose purpose is to highlight the magnitude of digitalization in Colombian society through Business to Consumer -B2C- methodologies, with a post-pandemic increase of 71% in the use of electronic commerce gateways by the adult population between 2020 and 2022.

It is considered that data mining has developed new tools in the financial sector, integrating business intelligence modules. These tools range from data extraction, transformation and loading, simplifying the analysis and representation of markets with a large flow of information. This allows the analysis of portfolios and transactions, strengthening virtual relationships with clients through the automation and decentralization of processes. Finally, the digitalization of the banking market is analyzed, concluding in the development of electronic payment platforms with new applications and their contribution to financial security, exploring measures that strengthen control over money laundering (SARLAFT), tax avoidance and evasion.

Abstract:

In the origin and evolution of industries, systems engineering has been fundamental, offering efficiency in process optimization. The financial industry, in particular, has benefited enormously. In the following, we seek to characterize and identify the impact of data mining in the modernization of the Colombian financial sector, from the 1990s to the present. The methodology that is executed has a qualitative and correlational approach against official surveys of the Bank of the Republic, whose purpose is to highlight the magnitude of digitization in Colombian society by Business to Consumer -B2C- methodologies, with a post-pandemic increase of 71% in the use of e-commerce gateways by the adult population between 2020 and 2022.

Data mining is considered to have developed new tools in the financial sector, integrating business intelligence modules. These tools range from data extraction, transformation and loading, simplifying the analysis and representation of markets with a large flow of information. This allows the analysis of portfolios and transactions, strengthening virtual relationships with clients through the automation and decentralization of processes. Finally, the digitalization of the banking market is analyzed, concluding on the development of electronic payment platforms with new applications and their contribution to financial security, exploring measures that strengthen control over money laundering (SARLAFT), tax evasion and avoidance.

Keywords: Data mining, e-commerce gateways, tax collection, SARLAFT, programming, software, databases and B2C.

General objective:

Evaluate the impact of data mining on the modernization of the Colombian financial sector from 1990 to the present, with emphasis on the value generated for individuals in the B2C methodology.

Specific objectives:

1. Analyze the digitalization of the banking market in Colombia, focused on end consumers as natural persons from data mining.
2. Explore the increase in security of the Colombian financial market
3. Identify optimization in the management of tax collections and reduction of SARLAFT risks

1. BACKGROUND:

Access to digital channels has become an essential component of everyday life. Communications, entertainment, the way of relating, working and interacting with others has undergone a drastic change compared to how it was carried out three decades ago within Colombian territory. Access to an interconnected network of information and the provision of data available to anyone has brought with it rights and obligations that, as natural individuals, generate a process of adaptation.

With this, the understanding, use, development and promotion of digital platforms, especially financial ones, have been crucial for income management and the administration of personal, business or home finances. In Colombia, according to the Digital Divide Index report (2022, s. f.) executed by the Ministry of Information and Communications Technologies -MINTIC- published in December 2023, it concludes in “a decrease in the period 2018 to 2022 of 9%, allowing large communities to access information and financial services more efficiently.” The report highlights the significant decrease in the digital divide during this period, as a consequence of the global pandemic. Various factors led to a change in face-to-face care for users, based on the need to reduce the spread of COVID-19 through physical contact between people. Sectors such as health, education, security, tax and especially banking had to strengthen and rethink the scope of non-physical contact with their clients.

For its part, Bancolombia, as a financial entity, was a pioneer in aspects of digitalization, successfully entering into non-personalized attention to clients and consumers. In this way, it opened the first Virtual People Branch in electronic commerce starting in 1996, offering a new alternative for Colombian consumers; In 1999, the Virtual Business Branch was inaugurated, which promoted the first Internet payment system in Colombia with 9 million transactions and was linked to 893 companies during this year (Portafolio, 2015), opening the field to electronic commerce and being an ally of many for the timely management of their obligations as employers.

Due to these obligations, electronic commerce begins to play a fundamental role, which is presented in different ways depending on the actors in the transactions; The main modalities according to Lara Navarra and Martínez Usero (2002, p. 408-409) can be classified as Business to Business (B2B), Business to Consumer (B2C), Business to Government (B2G), Consumer to Consumer (C2C) and Mobile Commerce (M-Commerce). On the other hand, we have authors such as Laudon and Traver (2021b), who on July 1, 2021 gave way under the Pearson Higher Ed publishing house. (Table 1) the following consolidated e-commerce modalities in post-pandemic environments:

Table 1 Post-pandemic e-commerce models

Modality	Participants	Description	Examples
B2B	Company → Company	Transactions between companies for wholesale sales of products or services.	SAP, Alibaba (wholesale), manufacturers and distributors
B2C	Company → Consumer	Companies sell products or services directly to the end consumer.	Amazon, Walmart, Netflix
C2C	Consumer → Consumer	Transactions between consumers for the sale or exchange of products or services.	eBay, Mercado Libre (C2C), Wallapop
C2B	Consumer → Company	Consumers sell products or services to companies, often as freelancers or content creators.	Upwork, Fiverr, content creators for brands
B2G	Company → Government	Companies offer products or services to the government through contracts or tenders.	Software or infrastructure companies for governments
G2C	Government → Consumer	The government provides direct services to citizens through electronic platforms.	Payment of taxes, issuance of passports online
G2B	Government → Company	Government interacts with companies for purchases or procedures, such as electronic bidding.	Public tender platform

Note: Adapted from E-Commerce 2021-2022: Business, Technology and Society, Global Edition, by Laudon, K. C., & Traver, C. G. (2021a). Pearson Higher Ed.

Given the above modalities of commercial transactions, it is crucial to decide to focus on one of them for a better approach. In this research we will take B2C (Business to Consumer) electronic commerce, defined by Medrano García, Díez de Castro and San Millán Fernández (2010, p. 81) as “the strategy that companies develop to directly reach the end customer.” Although both companies and consumers are involved in this type of electronic commerce (Arango et al., 2020), the object of study of this research is to increase the channels that companies generated to analyze the benefits to a large niche in the Colombian population.

2. Data mining as an extraction, prediction and conclusion factor for financial and banking markets in Colombia

It is a reality that the consumer user is the main focus of all business activity. Therefore, knowing the client's needs better than the client himself would give an unprecedented competitive advantage; How to design or develop a solution to that need, and offer innovative products that are irresistible?

Information units or data collectors are strategic pieces and raw material in the vision in each organization. Through this, there is the input to carry out automation projects that generate greater added value in the services offered in each entity. To the surprise of the entities themselves, their current processes are that unit of information. Measuring, quantifying, calculating and projecting the resources required in the execution of interactions with clients can be the gold mine for extracting and predicting internal and external movements.

In the case of Bancolombia, the opening of the Virtual Business Branch was a complete success, the flow of transactions reached 9 million being generated by the 893 companies, only the first year (Portafolio, 2015). Analyzing the architecture or the level of development that the application required for this period goes beyond the proposed study methodology; However, the robustness of the system for storage and reach in a little explored terrain is highlighted. In mention of what has been little explored, a relationship is made with the digitalization gap that was conceived at the end of the century. The vision of this banking entity encompassed an information unit, creating access and facilitating manual processes that they had been executing throughout the territory, given that its information repository was vast and generated organically.

Due to the abundance of information generated daily, converting excess data into quantitative information became essential. The information units generated with each transaction, such as the registration of dates, users, type of transaction, scope in documents, file extensions, weights, transferred amounts, loading capacity and requirements per minute of the application, formed clear and solid data. These data allowed data mining processes to be carried out using tools such as SQL, Hadoop and Python, according to true information, which implied a minimum adaptation of the data to carry out the relevant analysis. With the above, the new requirements of the entities were met, a clear example was the creation of new modules or operation profiles, for which an adequate analysis was required in the adaptation of timely solutions to the needs of the market.

As a solution to this great flow of information, the term data mining arises. As mentioned by Morales and Carrión (2020), in the data mining process, "statistical techniques and techniques for pattern recognition and identification of trends are used in the stored information. In this way, bibliomining provides a way to get to know our users", offering a great contribution when studying their movements and thus offering them quality services that in turn generate value.

Regarding the data mining process, it begins with the registration of operations by consumers. Therefore, banking entities have complete information originating from the history of transactions, to predict movements and customer requirements in personal or business banking. In effect, optimal designs and strategies are generated for each segment in the value delivery chain. The above is done through the exploration and analysis of large volumes of data stored in repositories of the information unit. However, reaching a decisive conclusion is finding the hidden correlation in the large magnitude of data, dispersed in the vast number of transactions. Well, in line with de la Puente (2010), there are some phases in the data mining process:

1. Determination of the thematic fields of interest.
2. Identification of internal and external information sources.
3. Collection, purification and process of hiding user identity in the system data warehouse
4. Selection of analysis tools.
5. Discovery of patterns, trends and preparation of reports.
6. Analysis and implementation of the results. (p. 4)

Starting from the object studied, Business to Consumer -B2C-, we can say that the first three phases have been defined. Therefore, the field of interest is the identification of the interactions that users have with financial platforms; Likewise, the source of information is the record of transactions that reach the Bank of the Republic, as a regulatory entity in the organic law of the Colombian financial system. These banking movements or transactions generate reference records of payment, sending or request of monetary mass, and in turn formal requests to the banks, whether for documents, credit studies, disbursement or transfer of money and those transactions that are aimed at the security of consumers, such as reports, consultations or accesses.

Because of these movements, we have a solid, first-hand information unit. However, there are cases where platforms offer a large flow of information and a variety of transactions, which requires purging, identity concealment, cleaning of said sources of information and correlation of variables for better analysis. Which brings us to the fourth point -Selection of analysis tools-, considering the connection of statistical techniques that allow us to extract behavioral patterns from large volumes of data; According to Morales and Carrión (2020), the most used data mining techniques are descriptive factor analysis, market basket analysis, clustering techniques, Bayesian networks, neural networks, decision trees and genetic algorithms.

For the practical purposes of the financial sector in Colombia we can associate the benefits provided by neural networks (sequential learning and prediction), clustering techniques (proximity between individuals), Bayesian networks (correlations and predictions), decision trees (consumer preferences) and genetic algorithms (behavior prediction). Thus, this set of statistical and prediction tools is useful for timely recognition of the organizational strategy, hence the extraction of user behavior patterns in the use of library services.

In February 2011, an innovative money management service was launched in Colombian territory; it was not until 2018 that Daviplata was more widely received in the Colombian monetary market. Among the pillars of its creation, prediction of customer movements and the increase in the use of new technologies such as social networks were argued, a key point for the marketing of its product. (History, Corporate Information and Strategy | Davivienda, 2024). The application, which at the end of 2019 had 6.1 million clients, by 2021 had more than 12.8 million users, opening accounts in Colombia and Costa Rica.

Its strategy was to identify the need and the access channels that the population had, so that anyone who had a SIM card was able to open a virtual wallet without carrying out the procedures required to open savings or checking accounts. Providing a competitive advantage in adapting to populations with disabilities, new functionalities to the ATM network, free virtual card with immediate opening and contactless availability, which earned Davivienda the international Model Bank award in 2018.

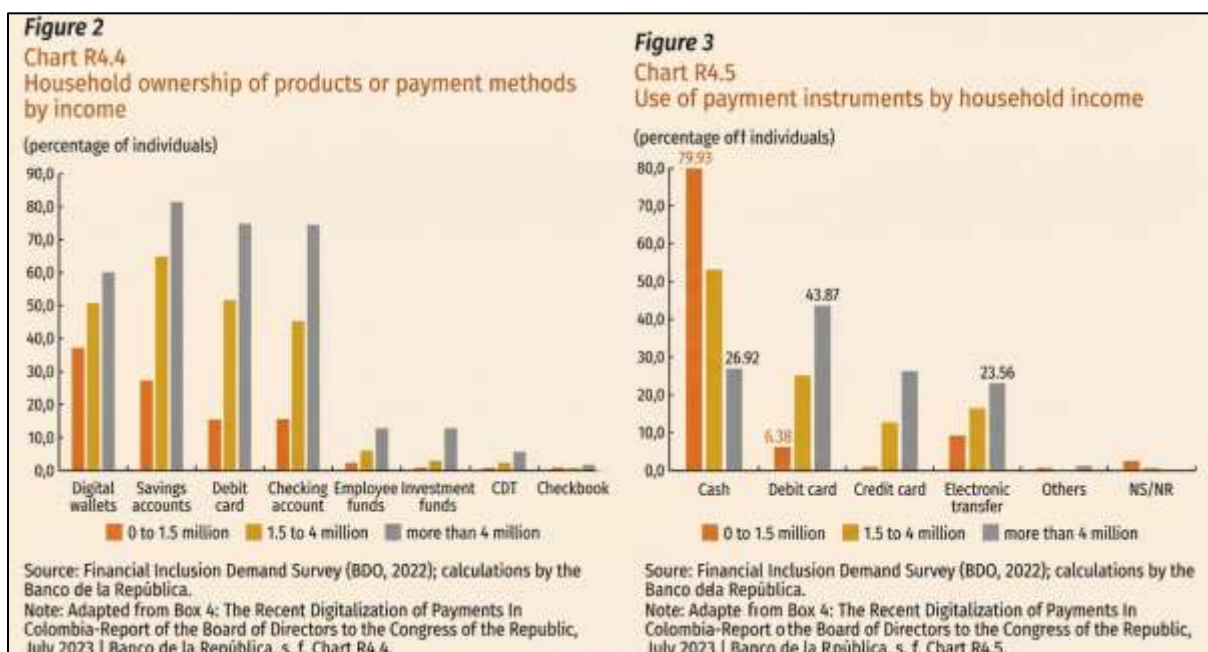
This platform has implemented various accessible technologies, such as artificial intelligence and voice narrators, to facilitate access to financial services for people with disabilities. Additionally, Davivienda has adapted its offices and ATMs to be more inclusive, offering services such as video conferencing with interpreters, braille signage, and hearing technology for ATMs so that they communicate or interact with users' electronic equipment. In this way, there is the possibility of credit requirements, payment services as beneficiaries of the State for vulnerable communities and document management from electronic portals. These initiatives seek to ensure that all people, regardless of their abilities, can access and benefit from financial services equitably.

Davivienda is a clear example of business success, and on the brink of a pandemic, it managed to add to the Bank of the Republic's figures for 2022 in its report, “Recent Digitization of Payments in Colombia.”

The study document “Recent Digitization of Payments in Colombia” is the Report of the Board of Directors of the Banco de la República to the Congress of the Republic published in July 2023, which is based on the survey carried out between April and May 2022; for which the central bank of Colombia or Banco de la República (Banrep) in collaboration with the Financial Superintendency of Colombia (SFC) carried out the Banca de las Oportunidades (BdO) financial inclusion demand survey.

The report was executed to identify the means of payment that Colombian adults, men and women over 18 years of age, have, especially after the health crisis caused by SARS-CoV-2 (acronym for Severe Acute Respiratory Syndrome Coronavirus 2) connoted by the World Health Organization (WHO) as Covid-19. In this way, growing technological developments for payment methods were identified, showing a 71% increase in the use of electronic commerce gateways by the adult population compared to two years after the most significant health crisis of the century began. Crucial figure that identified the top 10 payment methods (Figure 1 - Graph R4.1, 2023), used by Colombian nationals. As a result, a 64% increase in the frequency of use of mobile applications by financial institutions was reported and 56% reported an increase in the use of digital wallets in the same period (Figure 1 - Graph R4.1, 2023).

It is worth clarifying that the solution provided by the banking union -Asobancaria- and the commitment to solve the informality of transactions and reduce the use of cash, provided the input for the developments in software engineering, using applications on smartphones and access through web channels in various programming languages, where ease, security and strengthening of payment channels were provided. Due to the above, the figures that accompany the way in which individuals receive their main income increase (Table R4.4). According to the survey, about 53% of the adult population in Colombia, people over 18 years of age, men and women residing in all municipalities of the country, receive their income through electronic transfers (35% through transfers to bank accounts and 18% through transfers to mobile wallets and digital purses).



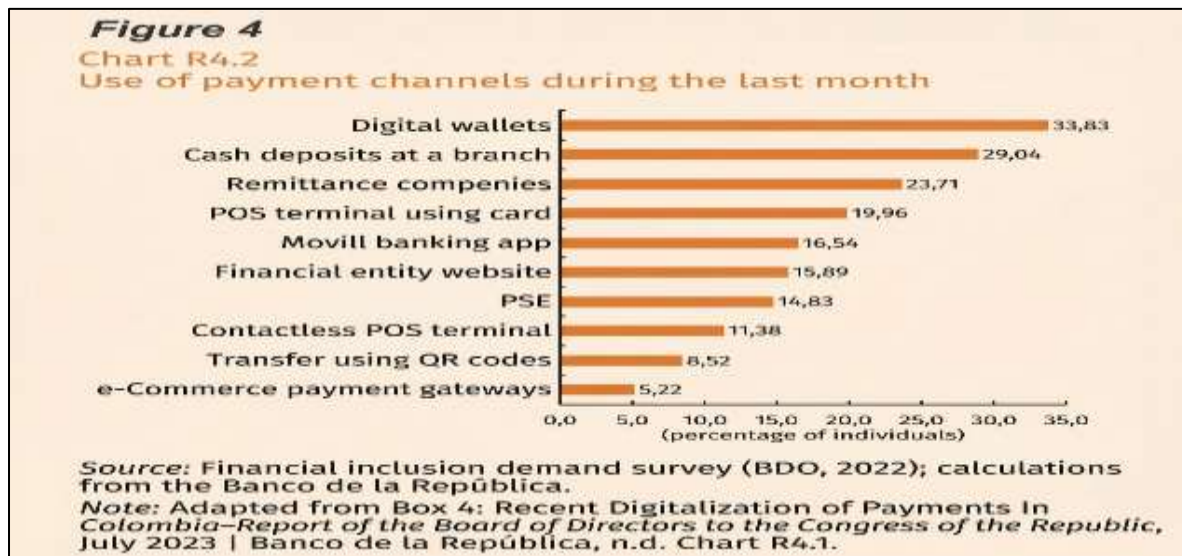
Within the study method, families were segmented according to their income: low income (0 to 1.5 million), medium income (1.5 to 4 million), and high income (more than 4 million Colombian pesos). As seen in Figure 2 - graph R4.4 generated by the survey “Recent Digitization of Payments in Colombia”, as income increases, greater ownership of financial products is generated, which relates the traceability of money and ease of movement of monetary mass: allowing basic transactions such as the purchase of inputs, payment of bills and in a broader model of financial education provision for savings, in products such as employee fund, investment fund and Term Certificates of Deposit (CDT).

In the following Figure 3 - graph R4.5 and in the same sense of increase in family income, it is observed how payment methods vary, decreasing cash handling. From there the study document concludes that the education perceived by high-income households provides security when carrying out digital transactions, which can justify the figures to some extent, protecting personal data and financial information provided in payment or transfer portals, such as key data from cards or savings accounts, dynamic keys, biometric data for authentication.

In the Financial Inclusion Demand Survey of Banca de las Oportunidades (BdO), carried out in collaboration with the Financial Superintendence of Colombia (SFC) and the Bank of the Republic (Banrep) in 2022, the angular input was generated to analyze the digitalization of the banking market in Colombia B2C, focused on end clients such as natural persons. For their part, financial entities that had platforms previously created in 2020 reinforced the security of the applications and generated a connection network for the benefit of the population. These platforms were key as intermediaries of an interconnected financial system during 2020, for the distribution of income, subsidies and aid by the State; It even opened up financing opportunities, accepting a reality with little physical contact.

The Colombian State considers, within the aforementioned document, that the increase in the use of electronic payments and its relevance in the development of the payments ecosystem represent a crucial advance to achieve the efficiency objectives of the public payments policy. However, it is essential to continue efforts to reduce significant gaps in the access and use of electronic payments, as well as in formal transactional and savings products, especially for people with lower incomes, with lower educational levels and residents in rural areas.

Transforming a traditional, conservative union, where the security perceived by society has been built little by little, given the economic and historical circumstances of the country, was quite a challenge. However, the results of a health crisis granted all social classes the opening of accounts or digital wallets. Highlighting the impact on the transformation of the financial sector as an ally at the service of companies and end consumers, in the simplification of processes



The evolution of the applications was evident, at the close of the survey since April and May 2022, digital wallets, such as Nequi, Transfiya, Daviplata, Movii, MercadoPago, PayPal, Tpage, to mention some regulated by the Financial Superintendence of Colombia, continue to be in constant movement and open to users. Clarifying that for the month of July in 2022 these applications were the preferred ones with 33.83% of the movement of money in the country (Figure 4-Graph B4.2). Generating intuitive platforms with a wide range of functionalities, being the key to opening generational doors, in all municipalities of the country, the increase in the use of digital channels in the last two years was greater than that observed in rural areas (54% versus 45.6%). The characterization of income based on households with high incomes (73%) versus those with low incomes (42%), “and in those with higher educational levels (70.7% of adults with a postgraduate degree and 64.9% with university studies)”.

Compared to the age reference, greater dynamism was shown in the use of digital payments in the groups between 46 and 55 years of age of 62%, expanding technological knowledge and adaptability to new technologies. Those between 18 and 35 years old saw 55% usage of pay channels, while a smaller proportion of adults between 36 and 45 years old increased their usage (46%). As we mentioned previously in the final phases of the data mining process, patterns and increases in payment trends are observed that, by virtue of a technological era, seeks the strengthening of systems and at the same time the simplicity of processes.

In this order of ideas, one of the characteristics that the systems have shown is the border interconnection and the strengthening of technological systems, it is the simplification in the transfer of remittances and migration flows, maintaining regulations from the generation of access to portals between nationals and foreigners, as well as the tracking and collection of information. According to data from the Inter-American Development Bank (IDB), in Latin America there was an increase of 9.5% in the arrival of remittances by August 2023 (Maldonado & Harris, 2023), which consolidated fifteen consecutive years of growth, representing a record in the sending of remittances to countries in Latin America and the Caribbean. Within the report published for November 2023, destination, origin, amounts, labor sectors and purchasing powers are determined according to currency fluctuations with the predominant countries for the flow of migrants such as the United States and Spain.

Specifically in the United States, the report highlights the increase in “the average weekly salary in the first quarter of 2023, where it reached \$860, the highest value observed in the

last eighteen years, and which meant a growth rate of 3.9% compared to the value observed at the end of the previous year” (Maldonado & Harris, 2023).

The interconnection capacity of databases has been able to provide governments and companies -G2B- with the notification of accurate information, as well as detailed information to banks for the registration of accounts and identification of consumers, as is the case of remittances. In other words, security and traceability mechanisms are generated in the flow of currency and people, as well as identifying the origin and destination of the resources.

This interconnection and alignment of databases with state entities not only improves the security and traceability of information, but also allows for a more efficient and effective integration of technologies in the financial sector. As these databases become interconnected, robust and flexible technology architectures, such as Service Oriented Architecture (SOA) and APIs, are required to ensure that systems can communicate efficiently and securely, ensuring connectivity over secure ports at any time. These technologies make it possible to develop banking and remittance platforms that offer integrated and accessible services for both nationals and foreigners, facilitating the flow of information and the accurate recording of transactions.

Service Oriented Architecture (SOA) is a software design approach that allows you to create services that communicate with each other. SOA is very useful for developing applications that need connection with others. It offers benefits such as service reuse, flexibility and scalability. In turn, APIs, or Application Programming Interfaces, allow different programs to communicate with each other. APIs define how requests and responses should be made. They are essential for the integration of systems and services, facilitating interconnection.

Software development approaches maintain their relevance since they represent the skeleton or support system within the financial scheme. Official portals within the Bank of the Republic such as Single Deposit Accounts (CUD) or Central Securities Depository (DCV), National Tax and Customs Directorate (DIAN) or National Administrative Department of Statistics (DANE), provide access through IP ports to interact with direct, accurate and updated information from web portals in real time.

The relationship between SOA, API and data mining is fundamental for the efficient management of information. SOA makes it easy to create modular services that communicate with each other using APIs, streamlining the integration of data from multiple sources. APIs act as bridges, allowing services to access and share information in a secure and structured way. This is crucial for data mining, as it allows you to gather data from different systems, analyze it and extract valuable information. The combination of SOA and API improves organizations' ability to process large volumes of data, generating accurate and timely insights that are vital for strategic decision making.

It should be added that this same alignment of information and consultation to unified databases can be used in various ways, allowing the banked society to offer services to business clients (B2B) for the consumption of individual clients (B2C). A clear example is the possibility of having access to the payment of public services, education, housing and credit modalities, even in different banked entities and with payment flows organized in real time, without delays in the interbank registration of transactional operations. The user interface of each application, whether from a web portal or an application for iOS or Android, generates familiarity and highlights attributes that encourage interaction and, in some cases, consumption. The potential for product development is then what marks the vision strategy of the companies, increasing interaction and possibilities within the same platforms.

Well, the value that companies give to their consumers can be reflected in the number of interactions they have from the portal of each banking entity or the connection they generate with other commercial platforms in the payment chain.

3. Explore the increase in security of the Colombian financial market

The Colombian financial market has improved its security significantly in recent years. The implementation of advanced technologies has been key in this process. Among these technologies, biometric authentication and the use of blockchain stand out. These tools have reduced the risk of fraud and increased user confidence. Additionally, financial institutions have adopted stricter regulations to protect customer information. These measures include data encryption and double authentication in transactions. Thanks to these actions, the financial market is now safer and more reliable.

According to Camejo et al. (2015b) is understood as an authentication method, “a step within the Identification-Authentication process in which it is determined if a user or entity can have access to a system or resource.” Following the guidelines of Asobancaria and applicable international regulations such as Basel II and III, the authentication methods for identifying the owners of assets in the financial system are the obligation and responsibility of each banking or credit institution, so that the security of the assets of each of the guarantors who have resources in said financial institutions is provided.

To provide a personalized, safe and tailored service, models have been created that allow system savers to safeguard their resources and dispose of them in a timely manner. These models include double authentication, secure keys, alternate keys, blockchain, and biometric authentications. These technologies allow financial institutions to recognize users and reduce the risk of theft. A clear example is the development of personal verification methodologies in application logging, identification-authentication.

The development in personal verification methodologies resulted in the decentralization of other processes, such as the opening, closing, transfer and status and account queries from any device; as well as the sending, receiving and disposition of money from remote accounts or pockets within the same accounts. To understand a little more, it is necessary to have a look at the interconnection and employability of services at the level of system components.

Well, as mentioned, user authentication and recognition is a step to determine access levels within the applications. However, to reach these accesses, filters are initiated at the architectural level, in systems where many databases are touched or consulted, not only in the company, but also in the entities that safeguard the information. For this, computer security layers with greater coverage are required, before, during and at the end of the messaging flow.

In Colombia, state databases with sensitive information can be consulted, whether notarial data, birth records, obtaining real estate, credit capacity or credit history, as long as users have detailed information for each of the properties or records they wish to investigate. As mentioned above, software development technologies that have SOA or API architectures can generate access to this type of requirements.

One of the new methodologies applied is the use of biometrics as an identification and authentication method, taking advantage of the resources that consumers possess and constantly update, such as cell phones or smartphones that have a camera, microphone and location in order to open the door to biometric identifications such as face and voice. As Mendoza (2020) mentions, “in the execution of daily operations, it has proven to be a

safer mechanism that reduces fraud and reports results with more certainty and confidence, although at the same time it entails privacy and security risks.”

It is in the branch of biometrics where new areas of study and application development have been generated, highlighting the need to clearly define terms of security and privacy. Progress in biometric identification systems, with the increase in data collection, as we see below, has increased with technological and computational development, ultimately seeking to identify individuals on the move in any context. “Biometric systems are methods based on pattern recognition that distinguish physical, psychological and behavioral characteristics, to achieve personal identification” (Mendoza, 2020b).

The international regulatory framework -ISO/IEC 19794- based on the international standard developed by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC), Defines the format of exchange and storage of biometric data, such as fingerprints, faces, irises, voice, and others, used in biometric identification and authentication systems. In Colombia, it was adapted to consumer protection through Law 1581 of 2012, which promotes a healthy use of biometric data within Colombian territory.

For its part, the FICO company, founded in 1956 and based in Silicon Valley, has been identified as a pioneer in the use of predictive analytics and data science to improve operational decisions. According to the “FICO 202 consumer survey: identity verification and digital banking” carried out by the entity and published in Mexico City on March 23, 2021, 14 countries were taken into account, including: US, Canada, United Kingdom, South Africa, Indonesia, Vietnam, Philippines, Malaysia, Thailand, Australia, New Zealand, Brazil, Colombia and Mexico. This survey highlighted that by that date, 63% of people in these territories had their biometric data registered with financial entities, which represents a significant increase of 19% compared to 2019. (FICO Consumer Survey 2021: Identity Verification and Digital Banking, s. f.)

A paradigm appears in access to the supply of information. Although it is possible to better understand customers through location tracking, their interaction with applications, social networks, content and purchases, an ethical dilemma arises. The line between accessibility and privacy of information, at this point, may not be clearly defined for the consumer, which raises important ethical questions, which we will not address on this occasion.

It is worth clarifying, financial education has played a crucial role in this increase in security. Banking institutions have invested in campaigns to educate their customers about safe practices. These campaigns include phishing prevention and the importance of using strong passwords. In addition, the use of secure digital platforms to carry out transactions has been promoted. This education has empowered users to better protect their assets. As a result, financial market security has improved, benefiting all participants.

4. Identify optimization in the management of tax collections and reduction of SARLAFT risks

The collaboration between banking entities, insurance companies, credit companies, central banks and tax collectors nourishes the information centralized in the DIAN. This effort seeks to reduce the loss or manipulation of monetary resources. The implementation of unified databases and the promotion of electronic invoices have reduced money laundering and tax evasion. Effective cooperation between banks and tax collection entities has guaranteed adequate management of tax collections, resulting in more effective and timely supervision of payments.

Alignment with international agreements such as Basel II and III, which establish standards for banking risk management and global financial stability, is considered the basic input of

the organic law of the financial and insurance system in Colombia - Decree 2555 of 2010. In this legal framework, SARLAFT risk prevention (Risk Management System for Money Laundering and Financing of Terrorism) through prediction algorithms in data mining has proven to be an effective strategy in the financial sector. These algorithms can analyze large volumes of transactional and behavioral data, identifying patterns that could indicate suspicious activity. For example, by analyzing frequent, high-value transactions in high-risk regions, algorithms can alert financial institutions to potential cases of money laundering. This proactive approach allows entities to take preventive measures before risks materialize, increasing the security and reliability of the financial system.

Furthermore, prediction algorithms in data mining allow for continuous, real-time monitoring of financial activities. This capacity for constant surveillance is crucial to detect anomalous activities that could go unnoticed through traditional methods, within these activities you can see some of the following: changes in spending behaviors, transfers to high geographic risk destinations such as tax havens or poorly regulated territories economically, transfers with unusual amounts or outside of normal hours, division of transfers to avoid withdrawal limits, as well as changes in personal information for the account holder, to name a few. By using advanced machine learning techniques, these algorithms can learn and adapt to new fraudulent behavior, continually improving their accuracy. In this way, financial institutions can respond quickly to any indication of illicit activity, mitigating the risks associated with money laundering and terrorist financing.

Finally, the integration of prediction algorithms in the SARLAFT risk management systems seen previously such as neural networks, clustering techniques, Bayesian networks, decision trees and genetic algorithms provides a significant advantage in terms of regulatory compliance, since they are aimed at predicting behavior. International and national regulations on the prevention of money laundering and the financing of terrorism are constantly evolving and financial institutions must adapt quickly to these changes. Within the international legal framework we find the Financial Action Task Force (FATF/FATF), the United Nations Convention against Transnational Organized Crime (Palermo Convention) and against Corruption (Mérida Convention), together with the European Directive on the Prevention of Money Laundering and the Financing of Terrorism.

Data mining algorithms can help institutions meet these regulatory requirements efficiently and effectively, reducing the risk of sanctions and improving the entity's reputation in the market. By ensuring that all financial transactions and activities are properly monitored and analyzed, institutions can strengthen their defense mechanisms.

Transaction control is carried out between B2B (Business to Business) or B2C (Business to Consumer) from the moment the electronic invoice is requested. This allows you to track the flow of capital and have a control model in the tax deduction with solid justification of real invoices. In addition, it helps prevent collaboration with fictitious companies, legally constituted, even if they do not have real operations in their entirety, in this way illegal money is introduced into the current flow of the country's economy.

In addition to what was mentioned above, the identification of shell companies is facilitated by associating patterns in the databases with the repository of the National Tax and Customs Directorate (DIAN). This comparative analysis makes it possible to detect irregularities and discrepancies between recorded transactions and official data. Thus, fraudulent companies that try to infiltrate the financial system can be identified and eliminated. This collaboration with the DIAN not only strengthens fiscal control, but also ensures a more transparent and secure economic environment.

5. Final considerations:

The use of B2C methodologies for data mining in Colombia's financial sector changed the way users interact with banking and credit institutions, having a transversal impact on the entire population. The digitization of services and access to intuitive and secure platforms benefited all actors in contact with the financial system.

Reaching a more complete understanding of customers was one of the main benefits of properly interpreting data. The ability of the financial sector to access this information and understand customers allowed it to provide an improved service tailored to their specific needs, which strengthened the banking values and increased its ability to attract and retain customers.

- Academic authors: Researchers highlight that digitalization has improved operational efficiency and security in financial transactions, benefiting citizens and entrepreneurs by facilitating data management and reducing operational costs.

- Financial analysts: They point out that digital platforms have increased financial inclusion, allowing more people to access banking and remittance services, which has boosted economic growth, benefiting citizens, managers and entrepreneurs.

- Technology: the use of advanced technologies, such as artificial intelligence and APIs, has allowed us to develop more friendly and intuitive user interfaces, improving the customer experience and the reliability of the system, benefiting citizens, managers and entrepreneurs.

- Government representatives: digitalization has strengthened the security and traceability of monetary transactions, allowing better regulation for the safety of citizens, control of the flow of capital and people across borders, as well as greater transparency in the financial sector, benefiting citizens, States and entrepreneurs.

As reported by the Banca de las Oportunidades (BdO) financial inclusion demand survey, carried out in collaboration with the Financial Superintendence of Colombia (SFC) and the Bank of the Republic (Banrep) in 2022, in particular, the Colombian population over 18 years of age contributed to the increase in bank turnover and the opening of savings and business accounts, this has been due to the ease of creating digital accounts and the new verification regulations implemented as a result of the pandemic period in 2020, that is, recognition of biometric data, capture of photos and information in personal identification documents, such as the ID.

At another point, the industry estimated the ability to face information saturation and transform it into an advantage with different data processing models, thus opening doors to statistical analysis. The possibilities of processing and interpreting this information were endless and were part of the evolution of the sector, turning data processing into an important asset for any financial institution. The implementation of these technologies in industrial environments brought great benefits. Due to the lack of adoption of new philosophies by automation engineers, the transition to these new paradigms was carried out in a progressive manner to ensure that devices within the various levels of a value chain were not affected.

However, flexible and reusable solutions implemented through SOA -Service-Oriented Architecture-, which are products of a software architecture, should not be conceived solely as the use of a tool; They must involve the implementation of various connectivity standards, as well as organizational and cultural changes of the company, dispersing activities to achieve objectives and generating focus in each area.

With the introduction of these paradigms, production systems went from being centralized to decentralized and intelligent (API), where each of the components had the capacity to

process, communicate, control and act on the events that were carried out. In this way, each machine adapted to the changes by self-configuring and optimizing functions and resources according to the orders of the environment or configurable services. The use of API technologies allowed communication in multiple programming languages through translators, eliminating obstacles between systems.

Platforms like Trii added value to the capital market by providing access to stock investments at low amounts in real time, which consolidated digitalization as a key pillar in the future of the financial sector in Colombia. Trii, a Colombian fintech, has democratized access to the stock market, allowing more than 600,000 Colombians to easily invest in 58 local and international companies. Since its launch, the application has recorded 44,123 transactions for \$70.34 billion in the fourth quarter of 2021, with a 334% growth in the number of investors and a 412% increase in transactions by March 2022 compared to the previous year. Trii has been recognized by magazines such as Forbes and Bloomberg, and has facilitated investment through low commissions and no minimum amounts.

Colombia's financial system was significantly modernized thanks to the use of electronic payment platforms, such as digital wallets, remittance companies, new dataphone applications, mobile applications, among other emerging ones, being strengthened with the support of government institutions. These efforts not only helped reduce the use of cash, but also sought to reduce risks such as money laundering (SARLAFT) and increase tax collection. Detecting suspected fraud activities was another common application of data mining. Data mining allowed us to identify fraudulent patterns and compare the information with our own database to draw conclusions.

6. REFERENCES:

1. Arango, P. Z., González, D. V., & Valencia, J. A. (2020). Business to Consumer electronic commerce and its adoption in the consumer of the city of Manizales - Colombia. <https://www.redalyc.org/journal/5713/571361695010/html/>
2. Camejo, I. M., Medina, G. A., Beltrán, P. M., & Carrillo, K. G. (2015b). Cuban Journal of Computer Sciences Vol. 9, No. 4, October - December, 2015 ISSN: 2227 - 1899 | RNPS: 2301 <http://rcci.uci.cu> Pages 104 - 117 Editorial "Ediciones Futuro" University of Computer Sciences. Havana, Cuba rcci@uci.cu 104 Type of article: Original article Topic: Information and telecommunications technologies Received: 05/26/2015 | Accepted: 09/14/2015 Performance of mining algorithms in academic indicators: Decision Tree and DOAJ (DOAJ: Directory Of Open Access Journals). <https://doaj.org/article/6ad57c1ef32c47538edb3e319e86b65e>
3. De la Puente, Marcelo (2010). Bibliomining and data mining. Information Sciences Consulting, 15. <http://www.slideshare.net/pattsul/014>
4. 2021 FICO Consumer Survey: Identity Verification and Digital Banking. (s.f.). FICO. <https://www.fico.com/es/newsroom/fico-2021-consumer-survey-identity-checking-and-digital-banking>
5. Fernández, J. A. F., Vázquez, V. B., & Virseda, J. A. V. (2019). Risk assessment with Data Mining: the Spanish financial system. Deleted Journal, 14(3), 309-328. <https://doi.org/10.21919/remef.v14i3.349>
6. History, corporate information and strategy | Davivienda. (2024, August 14). <https://ir.davivienda.com/historia/>
7. Digital gap index 2022. (s. f.). <https://colombiatic.mintic.gov.co/679/w3-article-333031.html>

8. Lara Navarra, P., & Martínez Usero, J. Á. (2002). Electronic commerce: user loyalty. *The Information Professional - UOC*, 11(6), 408–420. <https://doi.org/ISSN 1386-6710https://doi.org/ISSN>
9. Laudon, K. C., & Traver, C. G. (2021c). *E-Commerce 2021-2022: Business, Technology and Society*, Global Edition. Pearson Higher Ed.
10. Maldonado, R., & Harris, J. (2023). Remittances to Latin America and the Caribbean in 2023: resuming previous growth. <https://doi.org/10.18235/0005270>
11. Medrano García, M. L., Díez de Castro, L. T., & San Millán Fernández, E. (2010). Youth commerce and finance through new media. *Journal of Youth Studies*, ISSN-e 0211-4364, No. 88, 2010 (Issue Dedicated to: Youth and New Media), 88, 165–181. <https://doi.org/ISSN-e 0211-4364>
12. Mejía-Neira, Á., Jabba, D., Caballero, G. C., & Caicedo-Ortiz, J. (2019). Influence of Software Engineering on Industrial Automation Processes. *Technological Information*, 30(5), 221-230. <https://doi.org/10.4067/s0718-07642019000500221>
13. Mendoza, G. Q. (2020). Legislation, risks and challenges of biometric systems. *Chilean Journal of Law and Technology*, 9(1), 63. <https://doi.org/10.5354/0719-2584.2020.53965https://doi.org/10.5354/0719-2584.2020.53965>
14. Morales, M. E. F., & Carrión, R. B. (2020). Bibliomining, data and the decision-making process. *Inter-American Journal of Librarianship*, 43(2), e18. <https://doi.org/10.17533/udea.rib.v43n2ei8>
15. Portafolio, R. (2015, January 30). Bancolombia: 140 years that history takes into account. [Portafolio.co.https://www.portafolio.co/negocios/empresas/bancolombia-140-anos-historia-cuenta-27898](https://www.portafolio.co/negocios/empresas/bancolombia-140-anos-historia-cuenta-27898)
16. Box 4: The recent digitalization of payments in Colombia - Report of the Board of Directors to the Congress of the Republic, July 2023 | Bank of the Republic. (s.f). <https://www.banrep.gov.co/es/publicaciones-investigaciones/informe-junta-directiva-congreso/recuadro4-julio-2023>
17. Zambrano-Izurieta, J. P., Mendoza-Barberán, M. G., & Del Rocío Farez-Arias, M. (2023). Functionalities of Traceability in the Development of the B2C Electronic Commerce Model. <https://www.redalyc.org/journal/6955/695574854009/>